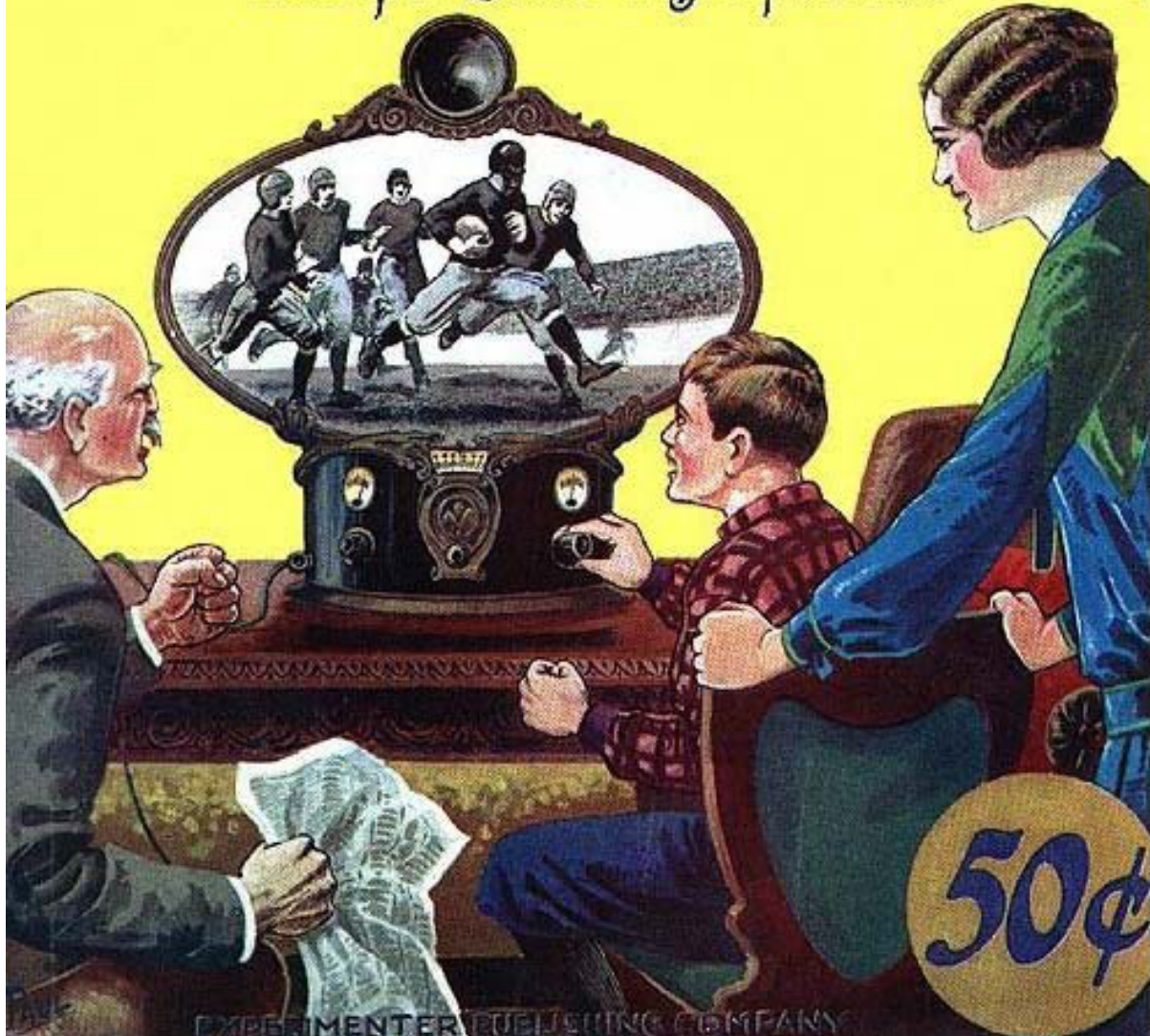


ALL ABOUT TELEVISION

INCLUDING
EXPERIMENTS
BY
H. Winfield Secor & Joseph H. Kraus



TECHNOLOGIES DE LA TELEVISION

De Benedetti

1/ Historique

Le but de la télévision est de transmettre des images animées à distance.

La première voie qui fut exploitée a été de copier le cinéma, donc de transmettre simultanément tout les points composant l'image. Chaque point lumineux était converti en signal électrique et transmis. Ce système nécessitait donc autant d'émetteurs que de points dans l'image, l'idée fut rapidement abandonnée.

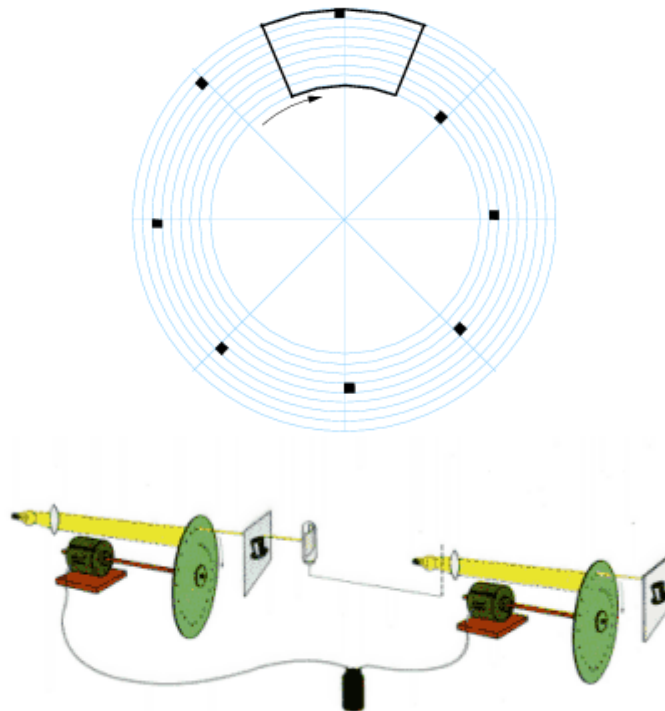
La télévision devait donc être un système séquentiel, l'image est décomposée à l'émission et recomposée à la réception.

A/ Le système mécanique.

1884, un ingénieur allemand, Paul Nipkow, étudie un dispositif à base de disque en mouvement.

Un **disque de Nipkow** est un système d'analyse d'une image destiné à sa décomposition sous forme de lignes dans le but de transmettre celle-ci par des moyens électroniques. Le système est basé sur un disque perforé tournant à 25 tours par seconde. Chaque trou, au nombre de 30 à 200, est placé à une distance décroissante du centre, ce qui permet d'analyser l'image ligne par ligne. Une cellule photoélectrique placée derrière le disque mesure les variations d'intensité de la lumière qui peuvent alors être transmises à un appareil récepteur distant.

Avec ce procédé, l'image est alors composée d'autant de lignes que le disque est percé de trous. Étant donné que l'analyse de l'image est effectuée par un objet circulaire, celle-ci n'est pas décomposée sous forme de lignes droites mais sous forme d'arcs de cercle.



En 1925, l'anglais John Logie Baird, crée la première société de télévision au monde. Quelques mois plus tard, il fera sa première démonstration publique.

En 1929, la BBC émet des programmes expérimentaux de mauvaise qualité. Seules quelques centaines de privilégiés pourront regarder ces images.



Jonh Logie Baird devant son prototype Baird (1930)



Système de réception Baird

Il faut savoir que le système Baird était un système à 30 lignes. La BBC se reposera dessus jusqu'en 1936 où elle adoptera le système EMI (*haute résolution*) à 405 lignes. Le système Baird présentait l'avantage d'être économique et facile à mettre en œuvre. La BBC avait pu utiliser ses émetteurs radio existants à faible bande passante pour diffuser les programmes. La première pièce avec diffusion simultanée de l'image et du son a eu lieu en juillet 1930.

B/ Le système électronique

En 1908, l'ingénieur britannique A. A. Campbell-Swinton a proposé un système analyseur qui utilisait un faisceau d'électrons produit par des tubes à rayons cathodiques (TRC) tant à l'émission qu'à la réception. Dès 1897, en dirigeant un faisceau d'électrons sur la surface intérieure fluorescente d'un TRC, Ferdinand Braun avait découvert qu'on pouvait faire dévier le faisceau lumineux sur la surface du tube en utilisant des champs magnétiques extérieurs. Ce fut le fondement de l'oscilloscope et, plus tard, du tube image.

Pour fonctionner, le système exigeait aussi un tube analyseur. Dans la course à la création de cet élément essentiel, il y avait trois concurrents : l'équipe britannique de Electronic & Musical Industries Ltd. (EMI), dirigée par Isaac Schoenberg, Vladimir Zworykin, un Américain à l'emploi de Westinghouse et plus tard de Radio Corporation of America (RCA), et Philo Farnsworth, un autre Américain, inventeur indépendant.



Iconoscope incorporé aux premières caméras de télévision commerciales en Amérique du Nord, vers 1940

Zworykin a démontré sa solution, l'iconoscope, en 1928,

Farnsworth a fait breveter son dissecteur d'image en 1930 et, en 1932, EMI a dévoilé le tube analyseur « emitron » qui surpassait le système mécanique de Baird. L'électronique venait de prendre le pas sur la mécanique.

En 1937, un système fondé sur l'émitron d'EMI a permis de mettre sur pied en Grande-Bretagne le premier service régulier de télévision à haute définition (pour l'époque) de l'histoire humaine. La Grande-Bretagne a établi la norme de 405 lignes et 25 images par seconde (remplacée par celle de 625 lignes en 1964



Récepteur EMI, 1938.

La Grande-Bretagne a inauguré le premier service de télévision électronique régulier au monde en 1937.

Des deux appareils américains, l'icône et le dissecteur d'image, c'est l'icône qui s'est montré le plus maniable. En 1939, la National Broadcasting Company (NBC) de RCA a diffusé des émissions expérimentales en direct de l'exposition universelle de New York.



Récepteur RCA TRK-12 lancé à l'exposition universelle de New York, 1939

Il est caractérisé par un tube vertical et un miroir qui réfléchit l'image pour la rendre visible

En 1941, les États-Unis ont fixé la norme de définition de la télévision en Amérique du Nord à 525 lignes et 30 images par seconde (norme encore en vigueur), mais la Seconde Guerre mondiale a retardé le développement commercial de la télévision. La diffusion d'émissions régulières a été entreprise seulement en 1947 aux États-Unis

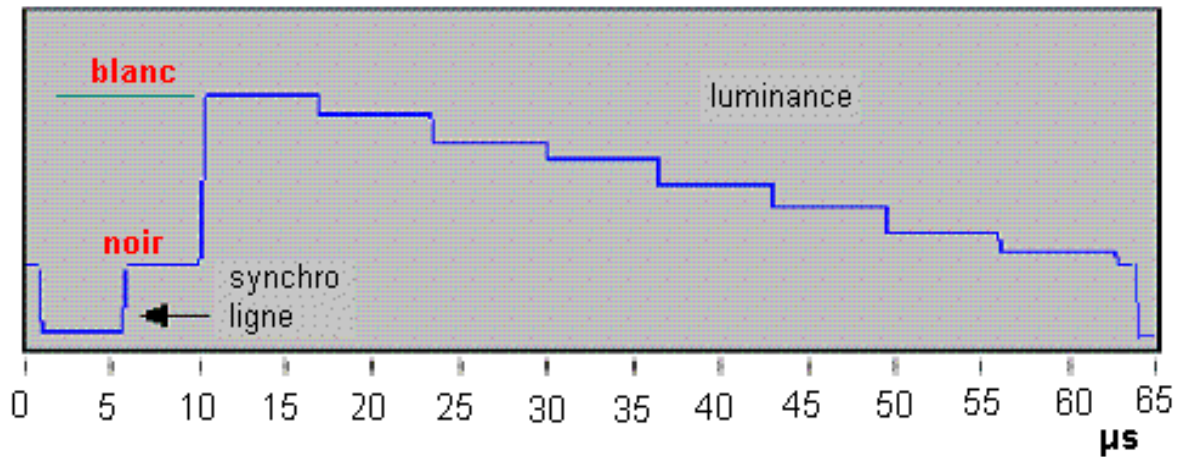
2/Principe

La télévision ou transmission à distance des images animées et des sons correspondants, utilise deux voies avec leur porteuse, l'une pour l'image, l'autre pour le son, occupant une certaine bande de fréquence, ou canal.

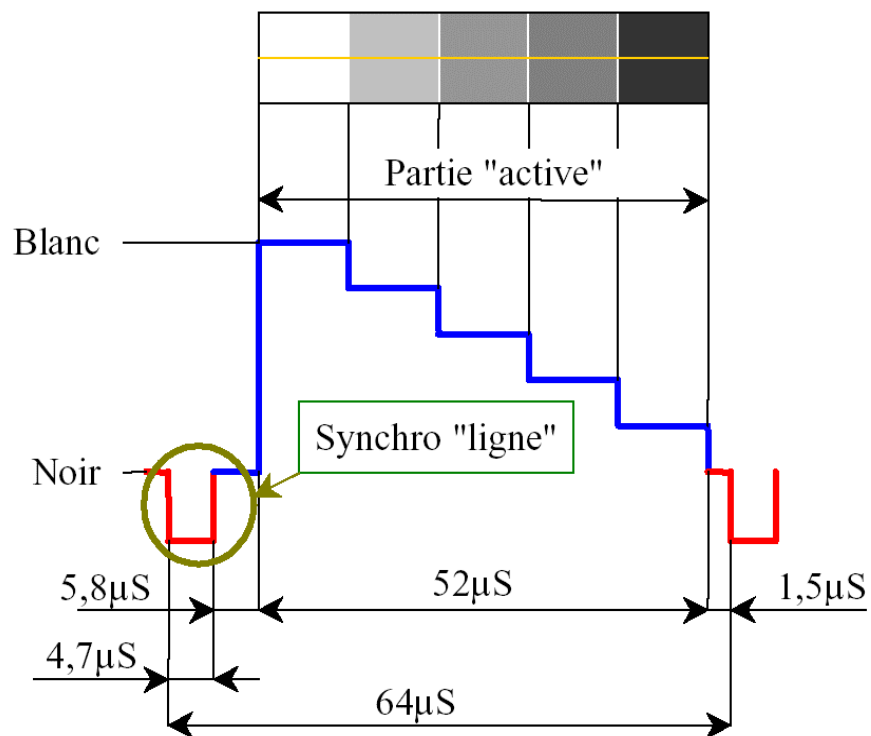
La transmission des images s'effectue en convertissant à l'émission une image optique en signaux électriques proportionnels à la brillance de chacun des points constituant l'image et en les transmettant les uns après les autres sur un canal unique.

A la réception on opère la conversion inverse

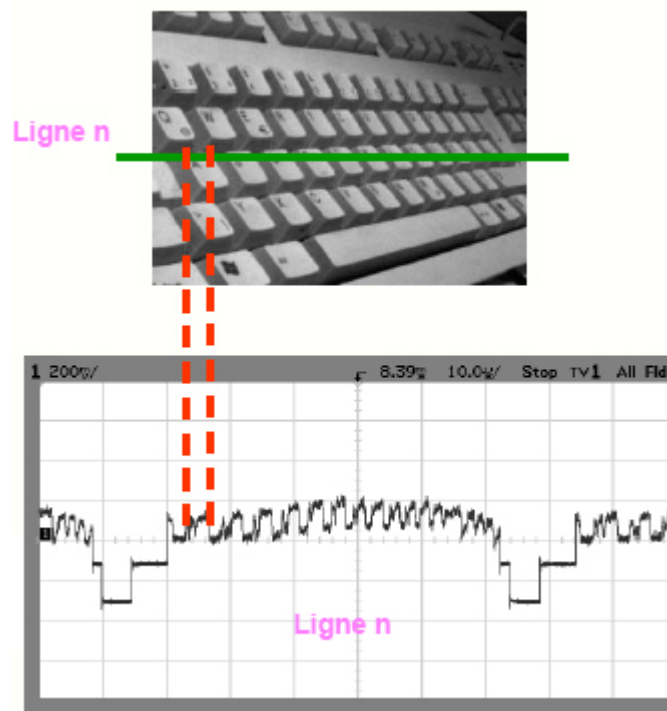
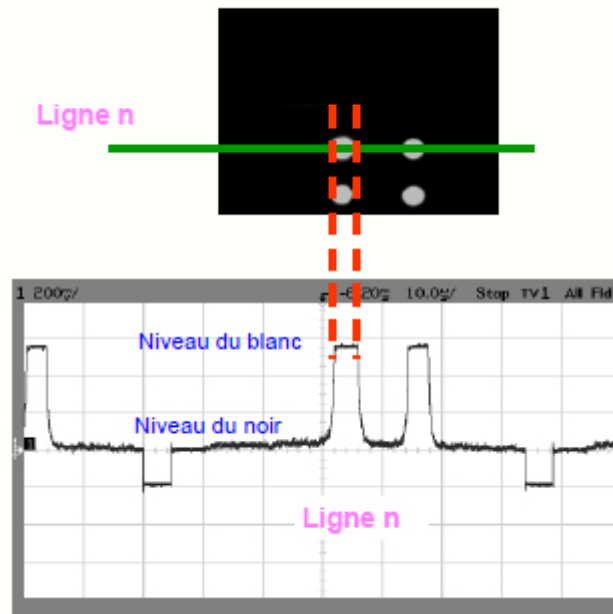
Il convient donc de transmettre, outre la variation de brillance de chacun des points en fonction du temps (vidéo) la position de chacun d'eux dans le plan de l'image (synchronisation).



Signal vidéo correspondant à une ligne



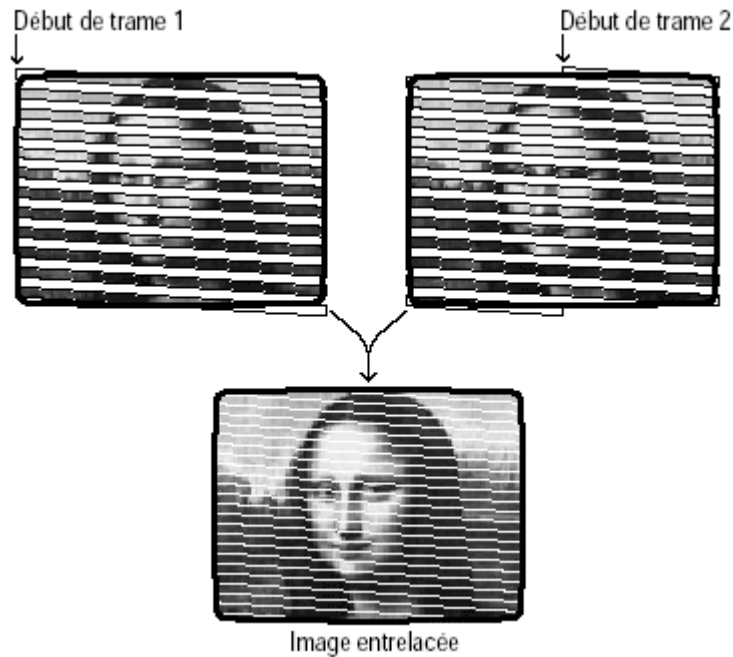
Exemples de signaux vidéo noir et blanc



A/Cycle d'analyse

La durée d'une image est fonction de la persistance rétinienne de l'œil humain (1/15 s) et de la fréquence du secteur (50Hz ou 60Hz), l'analyse compte d'une image dure 1/25 s en Europe et 1/30 s aux USA, valeur moitié de la fréquence du secteur, supérieure à la persistance rétinienne et proche de la cadence des images cinéma (24 i/s).

Pour éviter le scintillement de l'image, on analyse d'abord les lignes impaires et ensuite les lignes paires, c'est le balayage entrelacé.



L'intérêt de cette méthode est d'éviter un effet de papillonnement de l'image. En effet, lorsque le faisceau électronique balaie l'écran, les photophores s'éclairent, puis l'intensité de la lumière décroît. Donc, au moment où le bas de l'image est formé (pleine luminosité), l'intensité du haut de l'image a légèrement baissé. Or, la fréquence de rafraîchissement de la télévision analogique est de 50 Hz (Pal et Secam) ou de 60 Hz (NTSC), en raison de la fréquence du courant domestique (50 Hz en Europe, 60 Hz aux États-Unis); à cette fréquence, on peut voir une variation de contraste malgré la persistance rétinienne.

Si maintenant on effectue un balayage entrelacé, alors on balaie l'écran une ligne sur deux, et la variation de contraste a lieu entre deux lignes voisines et non plus entre le haut et le bas de l'écran. Les lignes étant très petites et resserrées, cela ne se voit pas, et l'on a un meilleur confort visuel.

Une image est décomposée en deux trames consécutives, une trame paire et une trame impaire. Une trame est une image de demi résolution verticale composée de 287,5 lignes. (sur les 625 lignes théoriques d'une image, il faut compter l'équivalent de 50 lignes pour le retour du spot en haut de l'écran, soit 575 lignes, donc 287,5 lignes par trame ; une trame commence ou se termine par une demie ligne). Une trame contient une ligne sur deux de l'image.

B/Définition

Le nombre de points par ligne détermine la définition horizontale, le nombre de lignes d'une image donne la définition verticale.

Pour effectuer un balayage entrelacé, le nombre de lignes par image doit être impair. En Europe on a adopté le système à 625 lignes et aux USA le 525.

En 625, l'oscillateur pilote fonctionne sur 31250 Hz, une division par deux donne la fréquence ligne 15625 Hz (625×25) et quatre divisions par 5 donne la fréquence de trame, 50 Hz.

En 625 lignes, définition verticale est donc de 625 points, la définition horizontale est de :

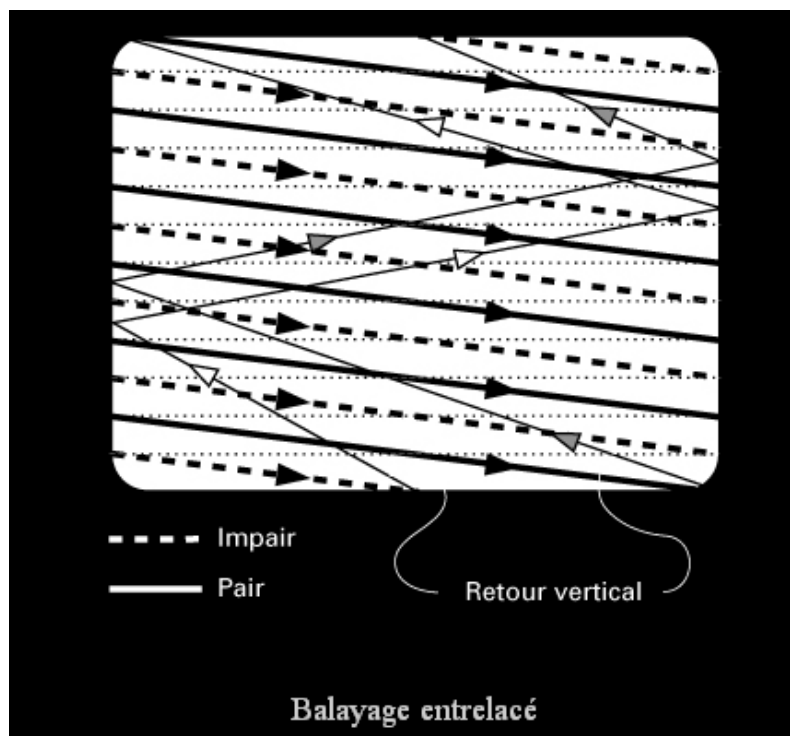
$$625 \times 4 / 3 = 830 \text{ points}$$

Si on analyse une image ayant successivement des points noirs et blancs, ils sont traduits par un signal électrique (vidéo) carré de fréquence :

$$625 \times 25 \times 830 / 2 = 6,5 \text{ MHz}$$

Ce cas étant limite, on adopte en pratique une bande passante légèrement inférieure de 5,5MHz.

En réduisant la bande passante vidéo, on diminue la définition horizontale sans agir sur la définition verticale imposée par le balayage, le point devient rectangulaire, et on perd en finesse.



C/Résumé

Fréquence image : 25 Hz
Fréquence trame : 50 Hz
Fréquence ligne : 15625 Hz

D/Synchronisation

A la fin de chaque ligne l'émetteur envoie un signal bref destiné à déclencher l'oscillateur de relaxation du récepteur, le spot d'analyse du tube cathodique doit revenir très vite au bord gauche de l'écran pour balayer la ligne suivante.

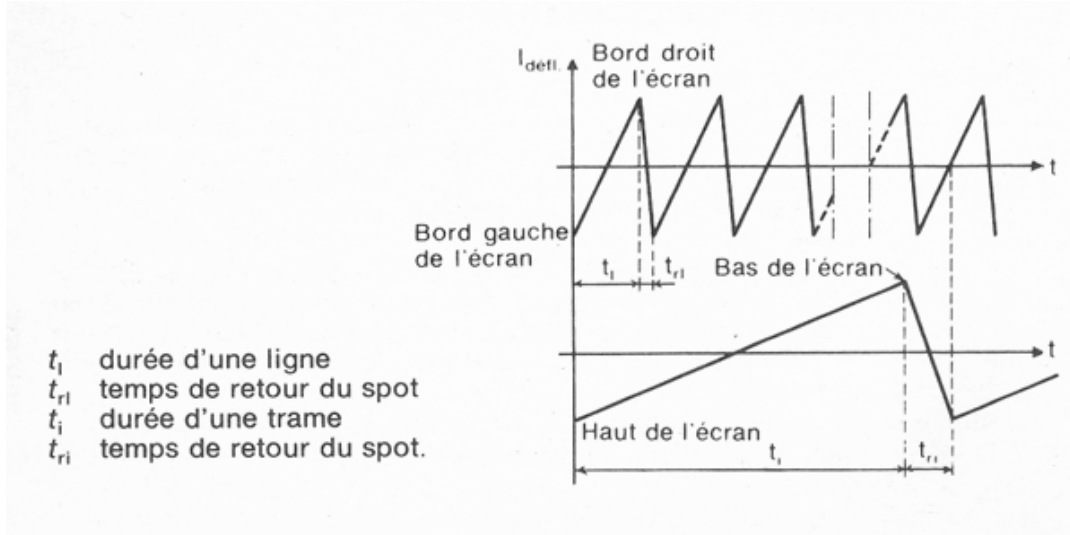
Ce signal bref s'appelle la synchro ligne.

A la fin de chaque trame, l'émetteur envoie un signal plus long destiné à faire revenir très vite le spot du bas droit de l'écran vers le haut gauche pour balayer la trame suivante.

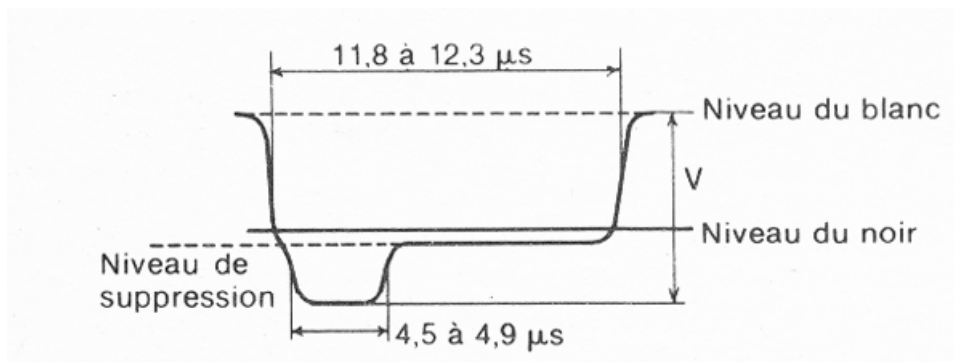
C'est la synchro trame.

La synchro ligne ne pouvant pas disparaître, la synchro trame doit contenir la synchro ligne.

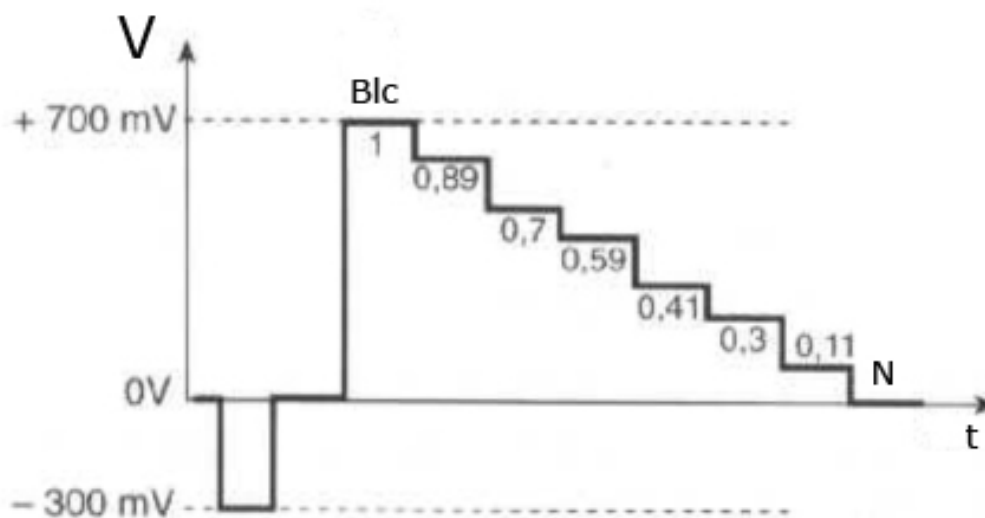
Les deux oscillateurs du récepteur, synchronisés par les « tops » de l'émetteur, fournissent les signaux en dents de scie destinés à assurer les déflexions verticales et horizontales.



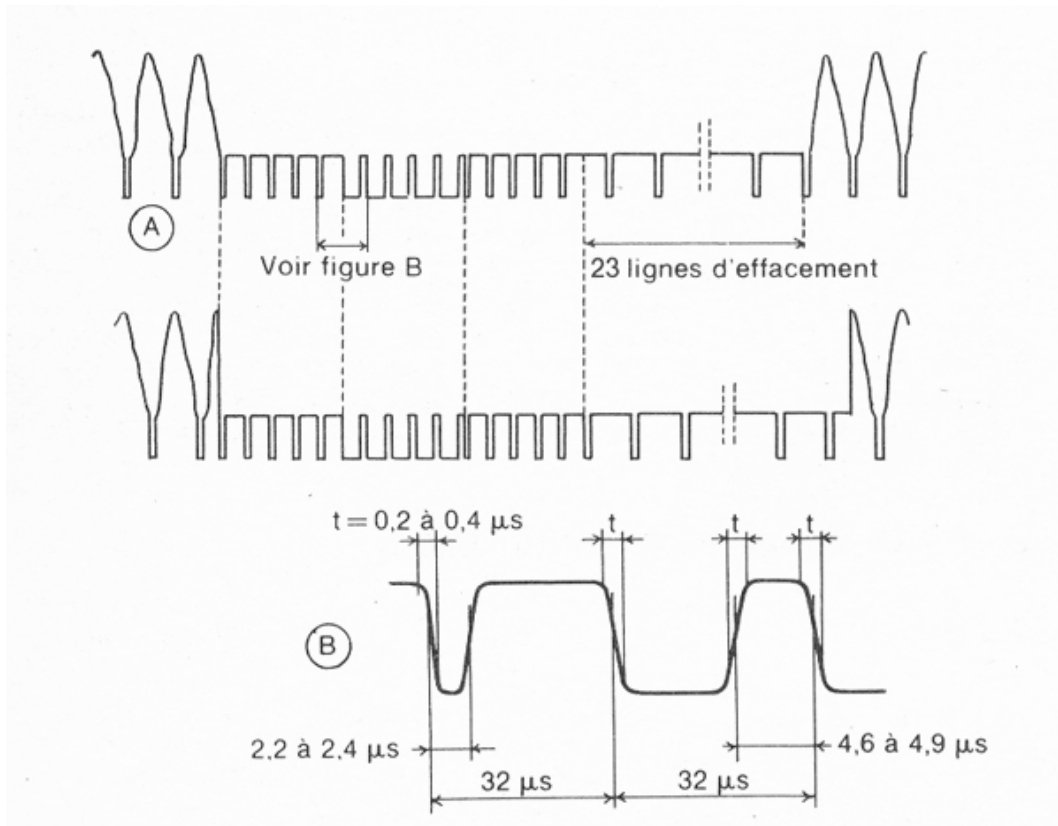
Signaux en « dents de scie »



Sync ligne



Ligne tv en N/B de steps de gris (du blanc au noir)



Sync trame

La synchro trame est composée de 5 tops synchro ligne inversés à double fréquence, précédés de 5 tops de pré égalisation et suivis de cinq tops de post égalisation, viennent ensuite 23 lignes d'effacement.

Ces lignes d'effacement, nécessaires à la remontée du spot, sont utilisées pour transporter les signaux télétexte, les anti-copie VHS, les signaux destinés au décodages de canal + (Be TV) analogique, etc...

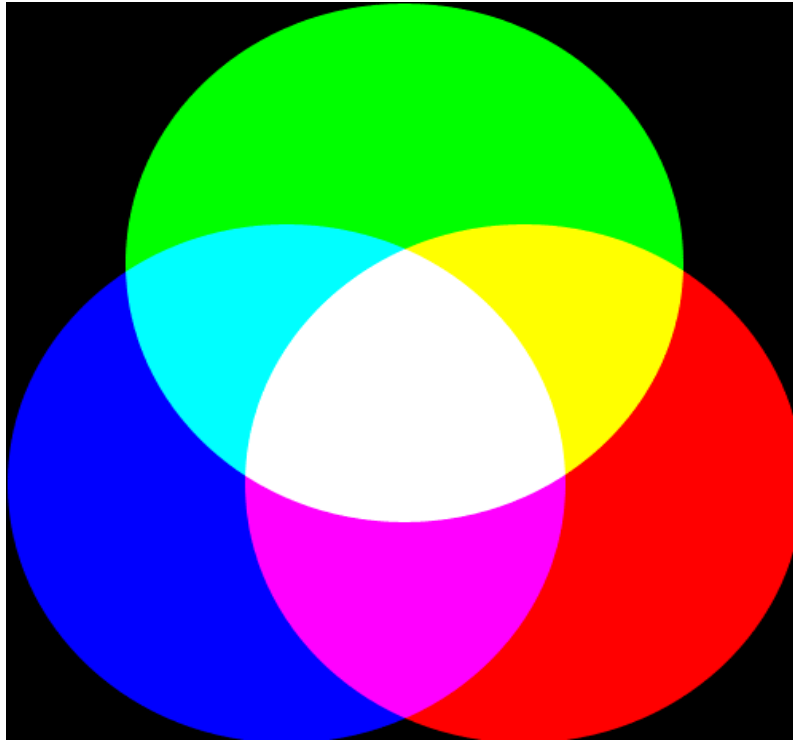
E/Colorimétrie

En télévision, on utilise la trichromie additive.

Si on illumine un écran blanc au moyen de trois projecteurs respectivement rouge, vert et bleu on obtient :

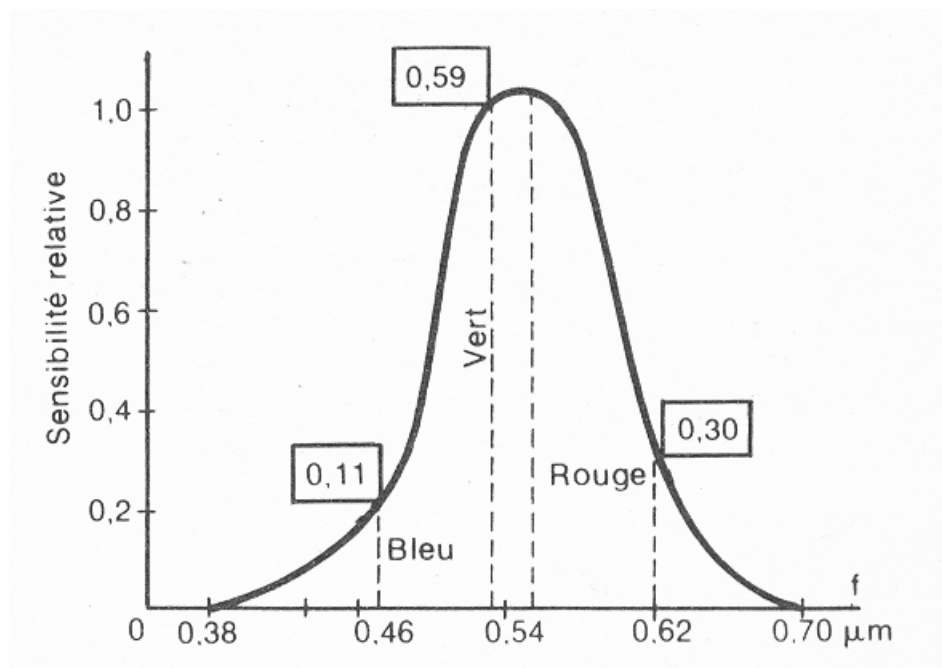
- au centre une tache blanche formée par l'addition des trois couleurs primaires
- à la périphérie, trois taches celles des primaires
- au stade intermédiaire, trois taches obtenues par les mélanges suivants :

Rouge + vert = jaune
Bleu + rouge = magenta
Vert + bleu = cyan

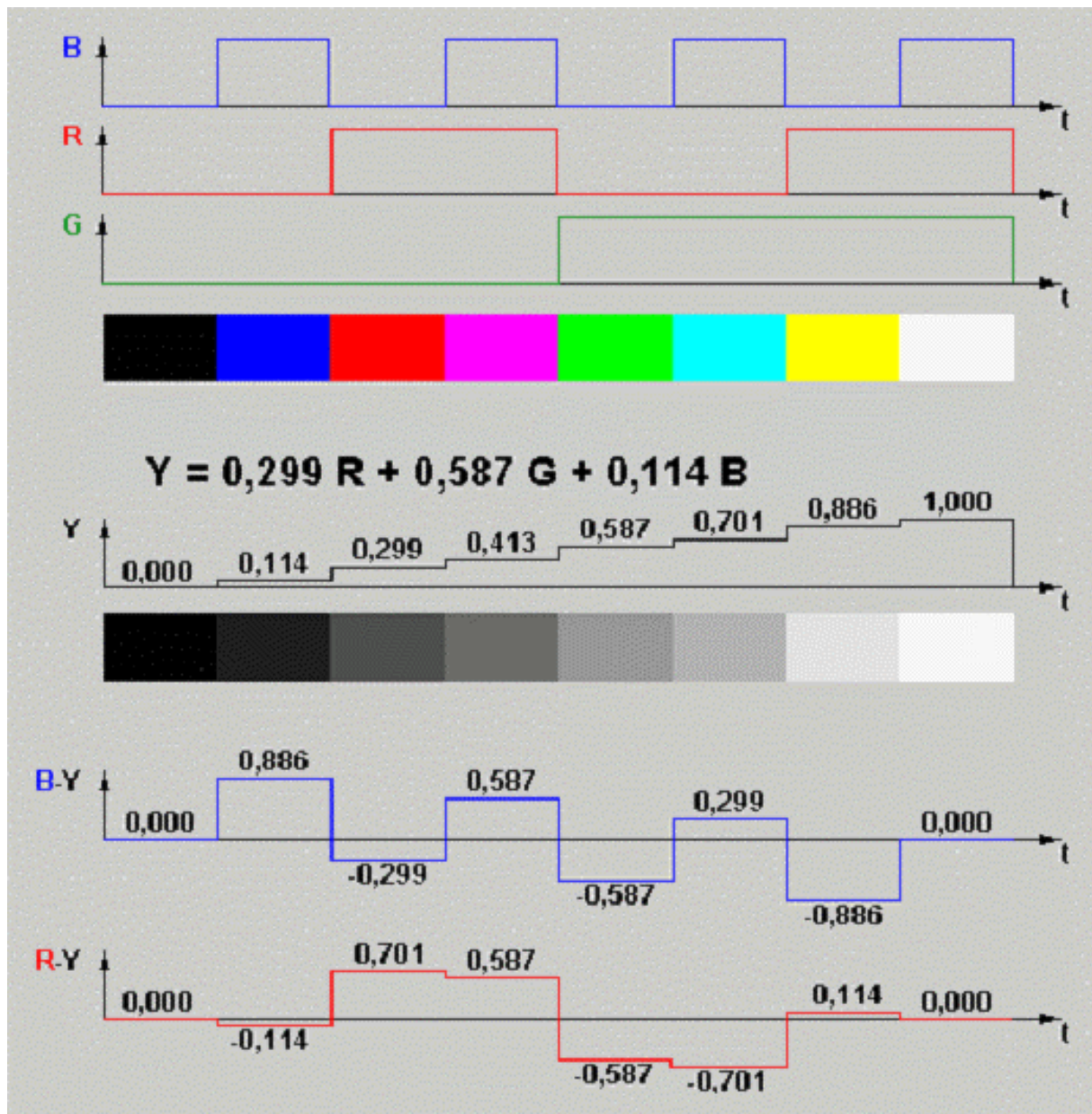


La Commission Internationale de l'éclairage a fixé les trois primaires :

rouge : 700 nm
vert : 546,1 nm
bleu : 435,8 nm

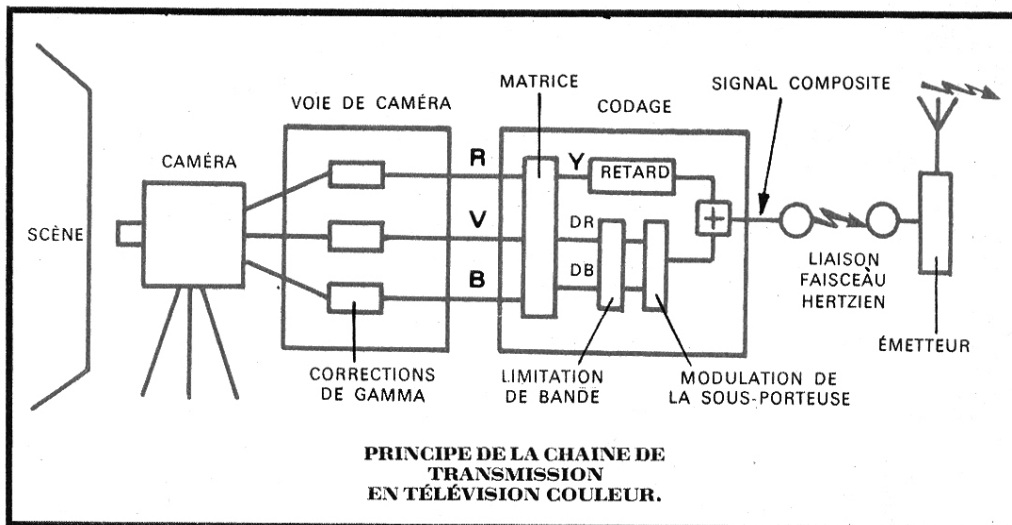


Sensibilité de l'oeil



La difficulté réside dans le fait de placer deux signaux de chrominance sur une seule sous-porteuse.

G/Systèmes PAL et NTSC



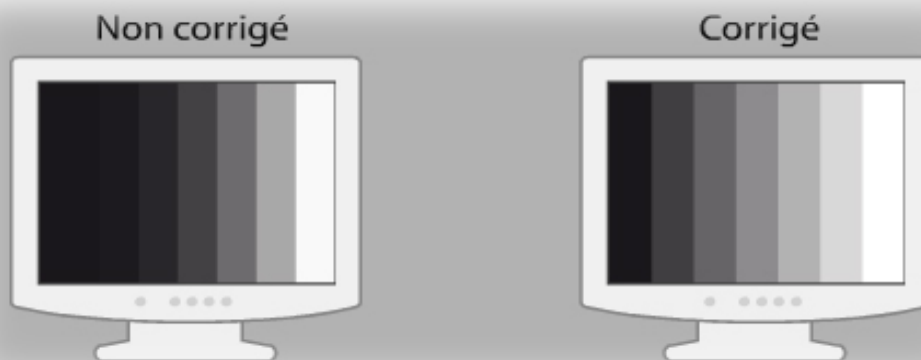
Les trois couleurs R/V/B, passent tout d'abord dans un correcteur de gamma, destiné à corriger la non linéarité tension/lumière des tubes cathodiques.

Qu'est-ce que le gamma ?

Le problème du gamma des moniteurs s'est posé dès l'apparition de la télévision dans la première moitié du XXe siècle. Les ingénieurs vidéo de l'époque furent confrontés à une restitution très sombre et contrastées des images sur les postes de télévision.

La cause de ce phénomène a été très vite cernée. La luminosité émise par les luminophores d'un tube cathodique n'est pas proportionnelle à la tension électrique appliquée dans le tube.

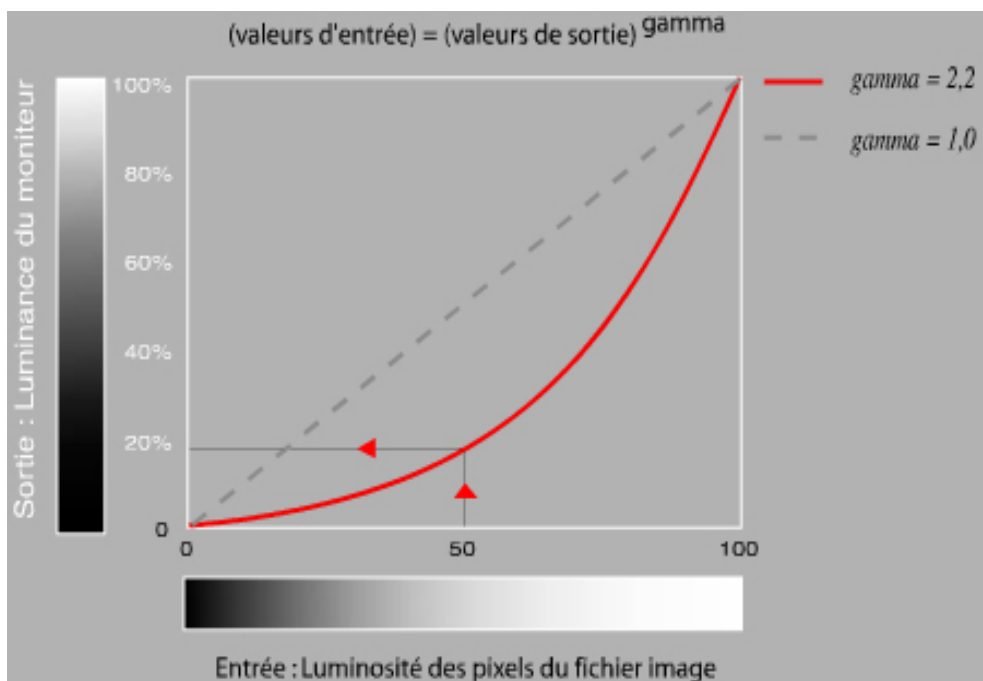
Si une tension de 1 Volt donne la luminance maximum de 100, une tension moyenne de 0,5 V ne donnera pas une luminance de 50 % mais plutôt une luminance plus faible de seulement 20 % et le phénomène ira en s'amplifiant à mesure qu'on se rapproche des valeurs proches du zéro.



En sortie, l'augmentation de la luminance est à la traîne par rapport au valeurs d'entrée. C'est cette non-linéarité qu'on appelle gamma.

Pour pallier cette distorsion, il existe 2 possibilités. L'une consiste à intervenir directement dans le récepteur TV en y introduisant un circuit électronique de compensation. Cela sous-entend un accroissement de la complexité du téléviseur grand public, des réglages supplémentaires et donc une augmentation du prix. L'autre solution consiste à effectuer la compensation de la non-linéarité du tube dès la source de production, c'est à dire dans la caméra. La correction consiste alors en une sur-amplifications des niveaux de gris correspondant à l'atténuation qu'ils subissent sur le téléviseur.

Le redressement des images dès la source est donc devenu le flux standard en vidéo, et par extension en informatique et en photographie numérique. Dans la pratique, on a donc pas à s'occuper sur notre écran de ce fameux gamma puisque toutes les images que nous manipulons ont déjà reçu une **correction du gamma** leur permettant de s'afficher correctement sur un moniteur. Cette opération effectuée à la source se nomme aussi **courbe de transfert** ou **compression du gamma**.



R,V,B deviennent R',B',V'

Les trois couleurs passent ensuite dans un matricage, on obtient : R'-Y', B'-Y' et Y' ($0.3R'+0.59V'+0.11B'$)

Les deux composantes de chrominance passent par un filtre passe bas limitant leur bande passante à 1.5MHz et ensuite on leur affecte un coefficient de pondération de 0.9 pour R'-Y' et de 0.5 pour B'-Y'.

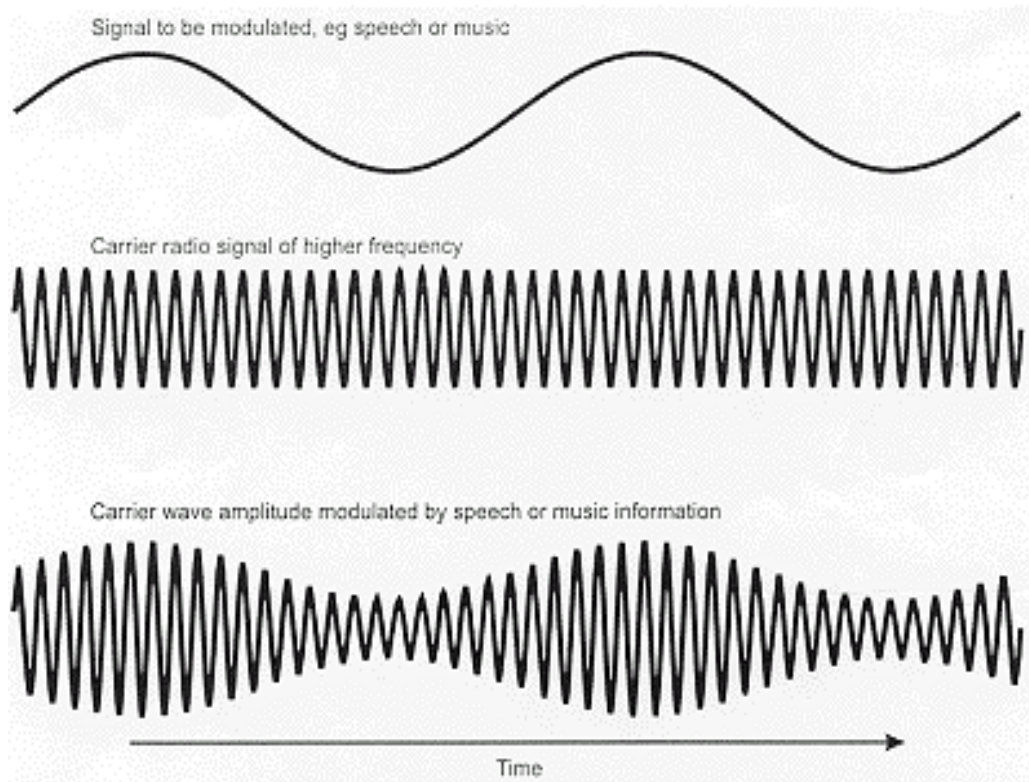
Nous obtenons enfin un signal **composant**.

Bande passante : Y=5MHz
R-Y=1.5MHz
B-Y=1.5MHz

La dernière étape consiste à ajouter ces composantes de chrominance (R-Y et B-Y) à la luminance (Y le noir et blanc) pour ne plus obtenir qu'un seul signal à transmettre.

Les deux informations de chrominance sont transmises simultanément par la modulation en quadrature de phase d'une sous-porteuse par les signaux R-Y et B-Y.

Qu'est ce qu'une modulation ?



- L'onde du haut est le signal à transmettre (celui qui va moduler)
- la seconde est la porteuse (carrier) celle qui va être modulée
- la troisième est la porteuse modulée en amplitude par l'onde du haut

l'intérêt ici est de pouvoir transmettre un signal qui a la même bande de fréquence que Y (donc non « mélangeable » avec lui) sur une fréquence unique (la fréquence de la porteuse)

Historiquement, ce sont les américains qui sont à l'origine du premier système de TV couleur : c'est en effet dès 1954, sur la base de travaux menés par les ingénieurs de la CBS (aujourd'hui sous le contrôle de Sony) et de RCA, que sont apparus les premiers équipements de prise de vues, transmission et réception basés sur un principe extrêmement ingénieux mais fort complexe édicté par le National Television System Committee qui lui a donné son nom: NTSC.

Les mauvaises langues, qui oublient trop vite que le NTSC est le père de tous les autres procédés, racontent que ce sigle signifie, en réalité, "Never Twice the Same Color" (jamais deux fois la même couleur) ce système ayant été desservi par un défaut inhérent à son concept : la dérive de la teinte des couleurs qui fait que, trop souvent, les présentateurs ressemblent à des martiens.

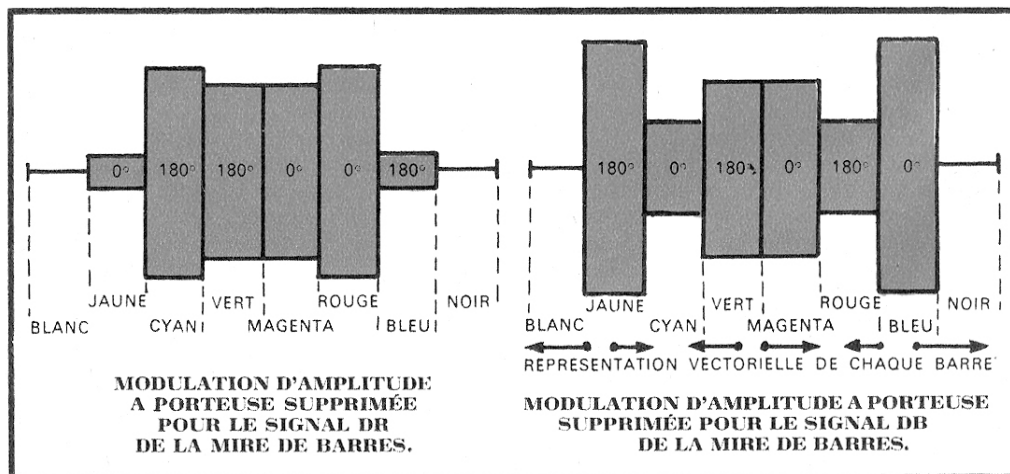
Explication : pour transmettre les deux signaux de chrominance, on utilise deux ondes secondaires (dites sous-porteuse par opposition à l'onde porteuse qui véhicule le signal TV) de fréquence 3,58 MHz en NTSC 4.43 MHz en PAL mais décalées (déphasées de 90°) entre elles.

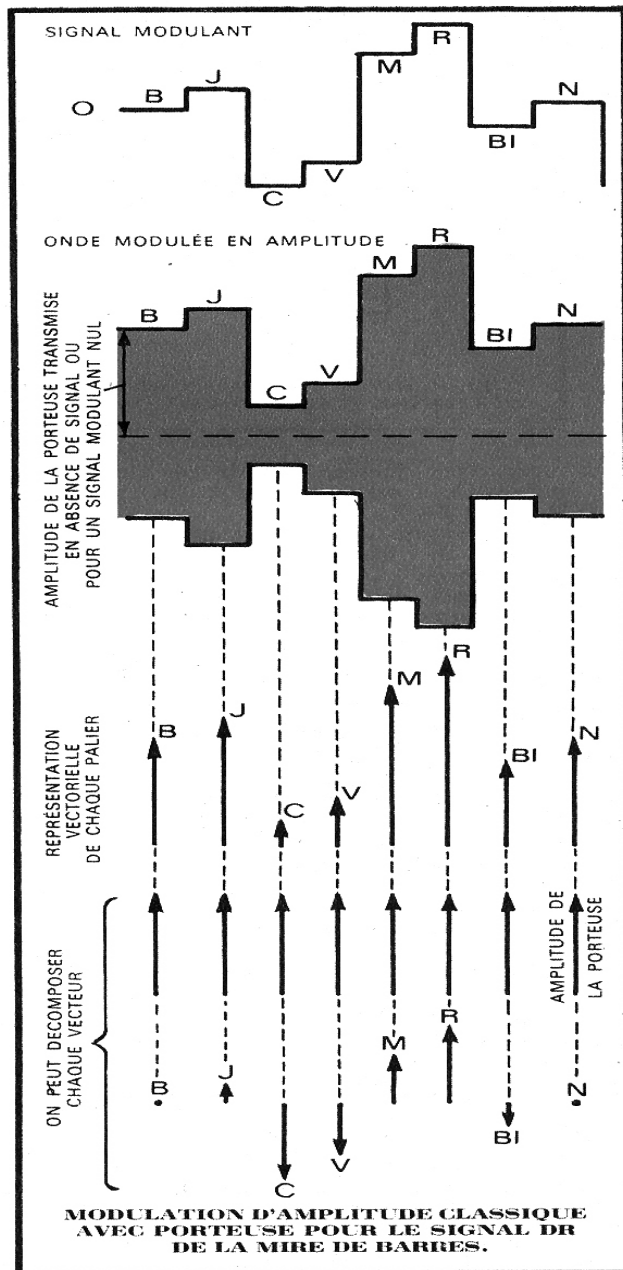
La première sous-porteuse est modulée en amplitude (très exactement modulation d'amplitude double en quadrature à porteuse supprimée) par l'un des signaux de chrominance, ensuite l'autre porteuse (déphasée de 90°)est modulée par l'autre signal de chrominance.. La modulation totale est représentée par un vecteur modulé à la fois en phase (la teinte) et en amplitude (la saturation) et variant en permanence selon les couleurs. Le signal ainsi obtenu est ensuite tout simplement ajouté au signal de luminance.

La modulation d'amplitude à porteuse supprimée

Le schéma de la page suivante représente une modulation d'amplitude classique avec porteuse ou on a choisi comme signal modulant le signal D_r (R-Y différence rouge) de la mire de barre. On voit d'après la forme de l'onde modulée, que lorsque le signal modulant est nul (noir ou blanc) on transmet la porteuse, pour une valeur positive du signal, elle augmente, pour une valeur négative, elle diminue.

Sur le schéma, on trouve aussi la représentation vectorielle de chaque palier du signal modulant et le croquis du bas montre que l'on peut considérer chacun des vecteurs comme la somme de deux vecteurs : l'un de même amplitude et de même phase que la porteuse non modulée, l'autre représentant en grandeur et en signe (positif ou négatif) la variation d'amplitude de la porteuse suivant sa modulation

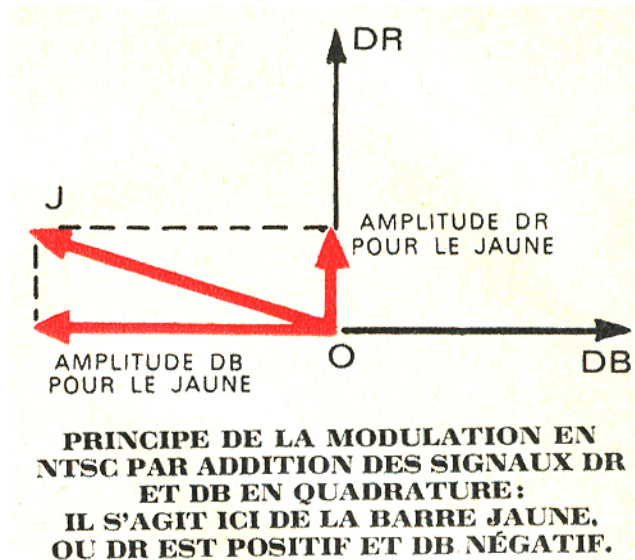




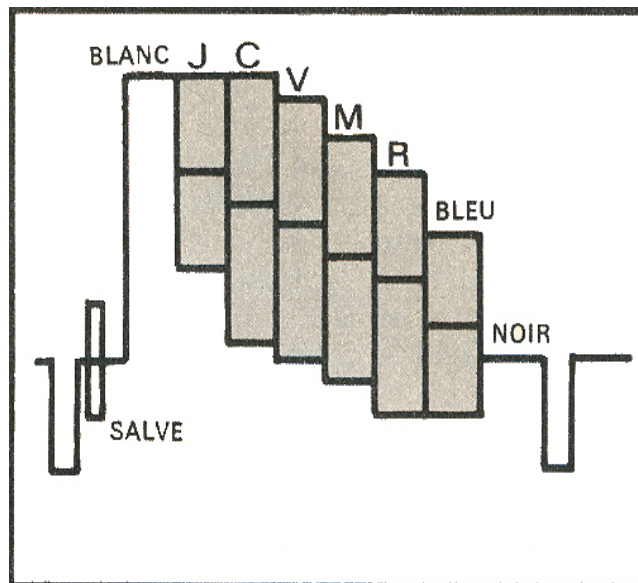
Pour les paliers positifs, l'onde haute fréquence a la phase 0° , pour les paliers négatifs, elle a la phase 180° , pour une modulation nulle (barre blanche, noire et les paliers de suppression), l'onde s'annule.

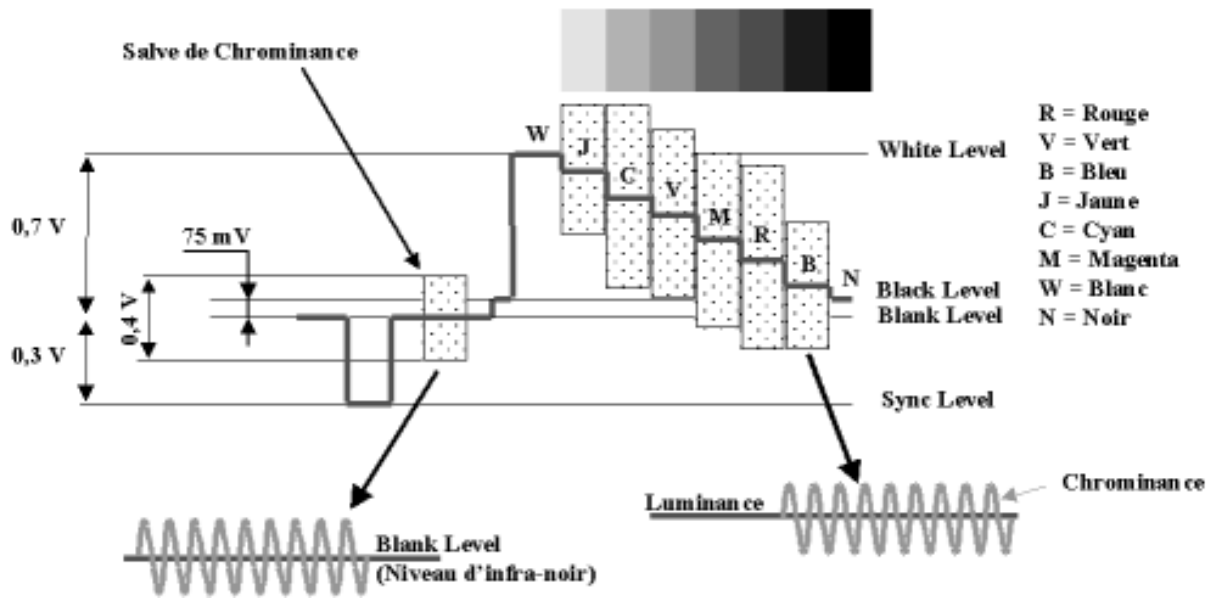
On module ensuite une porteuse de même fréquence avec le signal D_b , on impose à cette deuxième porteuse d'être déphasée de 90° par rapport à la première (en quadrature).

En additionnant simplement les deux produits de modulation, il est possible de transmettre simultanément D_r et D_b . Cela revient à faire la somme vectorielle de deux vecteurs représentatifs (schéma ci-dessous).

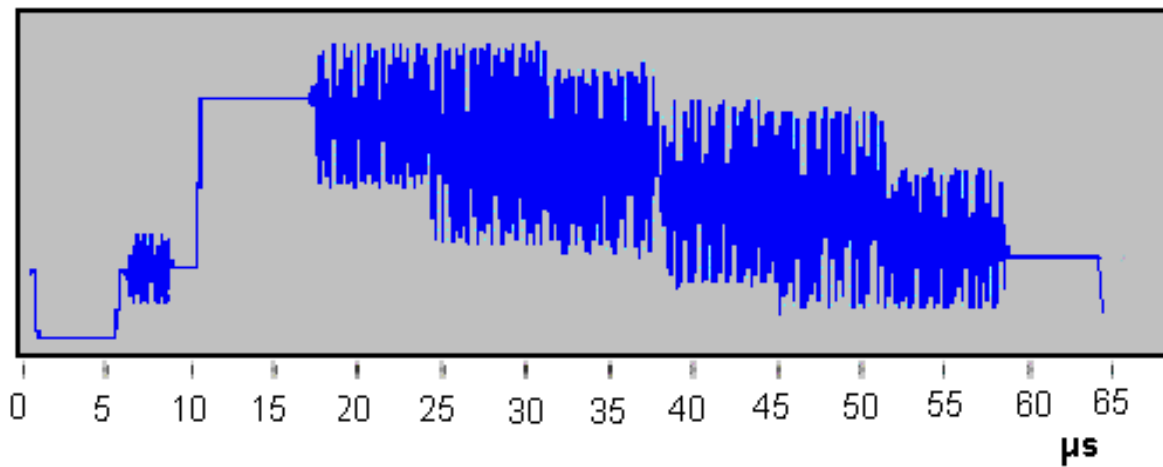


A la réception, pour démoduler une onde modulée en amplitude avec porteuse supprimée, il est nécessaire de régénérer la sous-porteuse, il faut lui donner la phase convenable et c'est dans ce but que des salves de référence (burst) sont transmises avec la phase de 180° sur les paliers arrières.

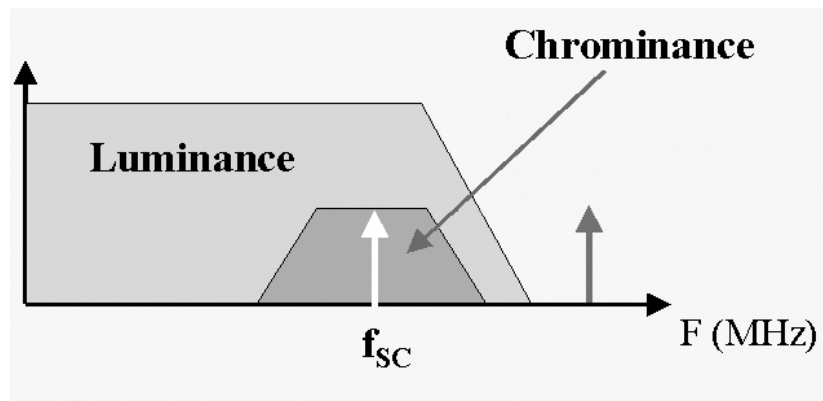




Signal composite (vue temporelle)



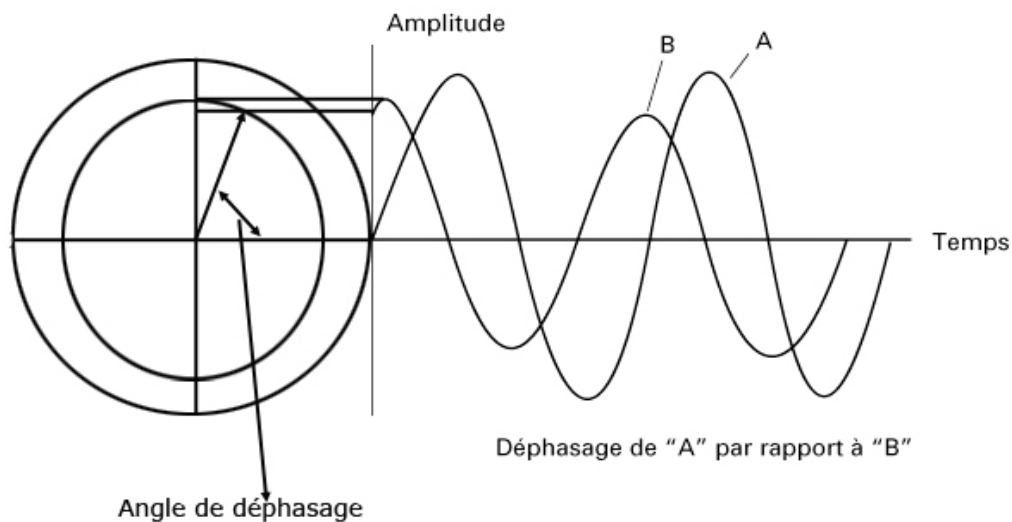
Signal composite (vue fréquentielle)



C'est cette inclusion dans les hautes fréquences N/B de la chroma qui explique l'apparition de taches colorées sur les images N/B très détaillées (ex : chemises finement rayées).

Dans un signal composite (voir schéma plus haut) on décode la chroma en comparant le 4.43MHz contenu dans la vidéo avec l'oscillateur interne (resynchronisé à chaque ligne par le burst).

La différence de phase donne la teinte l'amplitude donne la saturation.



En télévision, on utilise un vecteurscope pour visualiser les signaux de chroma dans un espace vectoriel.



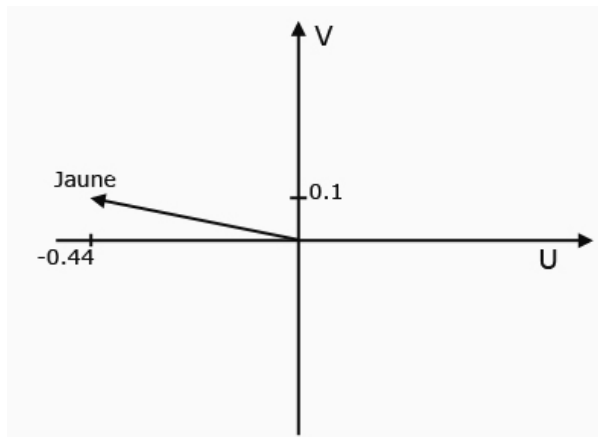
Recherche de la coordonnée d'une couleur

Prenons le jaune comme exemple.(le jaune de la mire « barres couleurs »)

$$Y_j = 0.89$$

$$V_j = (R-Y) * 0.9 \text{ donc : } (1- 0.89) * 0.9 = 0.1$$

$$U_j = (B-Y) * 0.5 \text{ donc : } (0-0.89) * 0.5 = - 0.44$$



Le système PAL

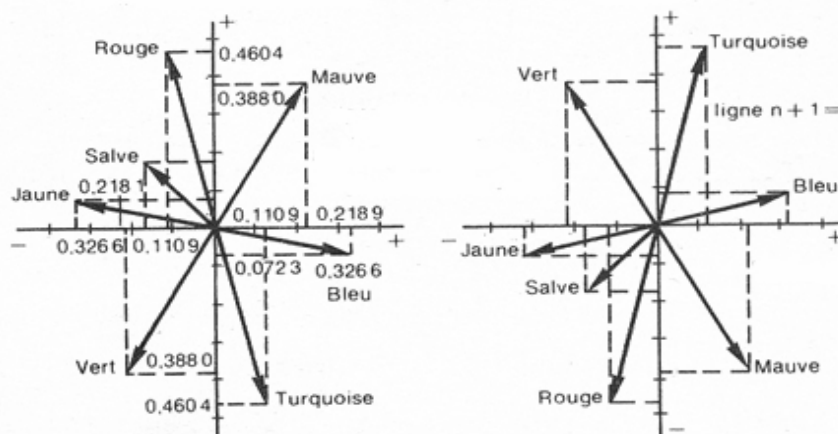
Tout ce qui a été vu précédemment est valable pour le PAL et le NTSC, seules les fréquences changent : NTSC : 30i/s , sc : 3.58MHz, PAL : 25i/s, sc : 4.43MHz.

Le problème du NTSC est qu'il est très sensible à toute variation accidentelle de phase en cours de transmission, ce qui produit des modifications de teintes à la réception (d'où son surnom : Never Twice the Same Color).

Le PAL (Phase Alterned Line) remédie à cette sensibilité, en effet R-Y voit sa phase inversée de 180° à la fréquence ligne.

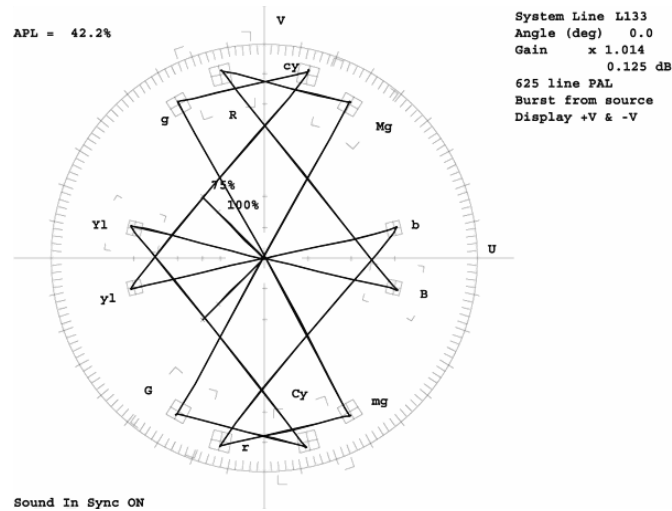
Le signal de chrominance reçu sur une ligne est mis en mémoire durant une ligne de façon à pouvoir être combiné avec le signal de la ligne suivante.

Si la transmission de ces deux composantes est affectée d'un déphasage, leur combinaison électronique permet d'éliminer le défaut.

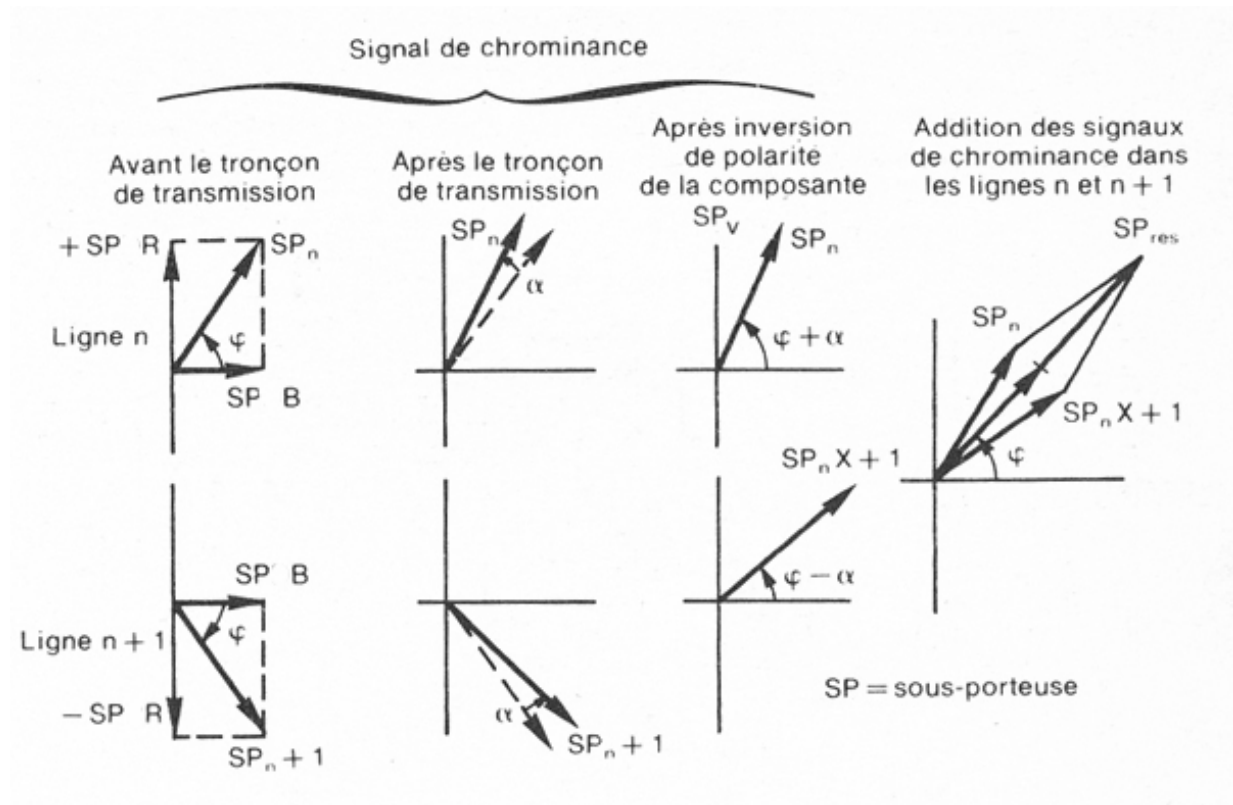


Représentation vectorielle des couleurs dans le système PAL pour deux lignes successives.

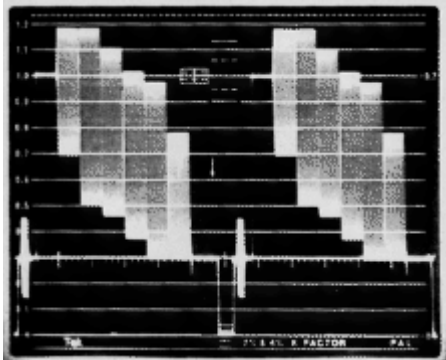
Le vectroscope PAL visualise deux lignes en même temps, ce qui donne le schéma suivant.



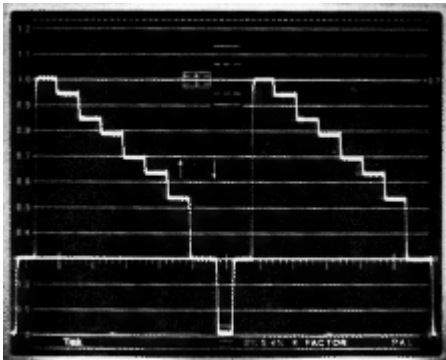
Le schéma suivant montre comment combiner les signaux de deux lignes successives pour éliminer les défauts accidentels de déphasage durant une transmission.



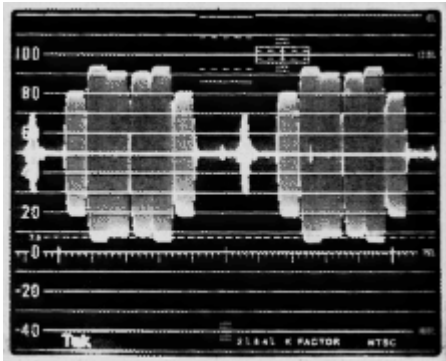
Visions à l'oscilloscope d'un signal composite



signal complet



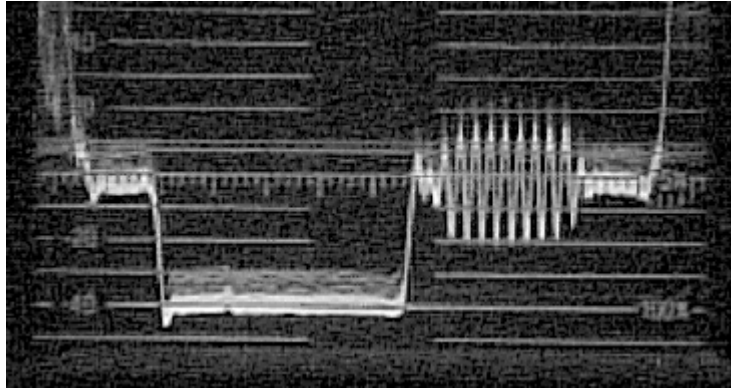
position low pass (élimination de la chroma)



Position high pass (uniquement la chroma)



vue du blanking



vue de la synchro ligne et du burst

Le SECAM






Le SECAM (pour SEquence de Couleurs Avec Mémoire) est basé sur le même principe que le NTSC. Il en diffère cependant par le fait qu'il utilise une ligne à retard de $64 \mu\text{s}$ pour ne transmettre, à chaque instant, qu'un seul des deux signaux de chrominance, en alternant séquentiellement l'un et l'autre signal de ligne à ligne. Cela est obtenu en modulant, en fréquence, une onde sous-porteuse. Afin d'éviter tout risque d'erreur, une fréquence de sous-porteuse différente est attribuée à chaque couleur (4,40 MHz et 4,25 MHz).

Le SECAM a été mis au point par l'ingénieur français **Henri de France**:

- il procure une image en couleur d'aussi bonne qualité que le PAL
- il a été adopté pour éviter la pénétration des téléviseurs étrangers (PAL, NTSC) sur le marché français
- d'où l'appellation **S**ystème **É**légant **C**ontre les **A**méricains !

standard	fréquence des sous-porteuses couleur	modulation	remarques
NTSC	<p>B-Y : $f_c = 3,579545 \text{ MHz}$</p> <p>R-Y : $f_c = 3,579545 \text{ MHz}$ (déphasée de $+90^\circ$)</p>	AM sans porteuse	<ul style="list-style-type: none"> • on transmet simultanément R-Y et B-Y • les deux sous-porteuses sont de même fréquence mais déphasées • des déphasages parasites durant la transmission faussent souvent les couleurs à l'arrivée • d'où l'appellation humoristique Never The Same Color !
PAL	<p>B-Y : $f_c = 4,433619 \text{ MHz}$</p> <p>R-Y : $f_c = 4,433619 \text{ MHz}$ (déphasée de $\pm 90^\circ$)</p>	AM sans porteuse	<ul style="list-style-type: none"> • le PAL est une amélioration du NTSC (sous-porteuse de R-Y déphasée de $+ \text{ ou } - 90^\circ$ une ligne sur deux) • on transmet simultanément R-Y et B-Y • meilleur rendu des couleurs
SECAM	<p>B-Y : $f_{B-Y} = 4,2500 \text{ MHz}$</p> <p>R-Y : $f_{R-Y} = 4,40625 \text{ MHz}$</p>	FM	<ul style="list-style-type: none"> • SECAM = Séquentiel à mémoire • on ne transmet qu'une couleur par ligne (l'autre couleur est prise de la ligne précédente) \Rightarrow pas de mélange des couleurs • bon rendu des couleurs

Différents types de connecteurs de la vidéo analogique.

connecteur	utilisation	brochage
	liaison PC – écran externe	5 signaux séparés toutes les composantes R,V,B, synchro ligne et synchro trame séparées
	équipements vidéo (TV, DVD, camera, décodeur...)	<ul style="list-style-type: none"> • signal video composite et • 2 signaux séparés (SVHS) ou 3 composantes RVB
	équipements vidéo (TV, DVD, camera, décodeur...)	2 signaux séparés : <ul style="list-style-type: none"> • luminance Y • chrominance PAL/SECAM/NTSC
	matériel professionnel	3 signaux séparés : <ul style="list-style-type: none"> • luminance Y • couleur R-Y • couleur B-Y
	équipements vidéo (TV, DVD, camera, décodeur...)	un seul signal vidéo composite <ul style="list-style-type: none"> • luminance Y + chrominance PAL/SECAM/NTSC

3/ENREGISTREMENT ET LECTURE MAGNETIQUE

A/Historique

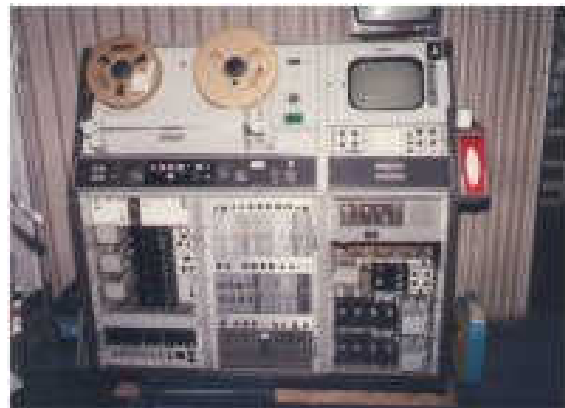
C'est avec le développement de la télévision dans les années 1950 qu'est apparu le besoin de conserver et reproduire les émissions produites. La première solution s'apparente au **kinescopage** (transfert d'un signal vidéo sur une pellicule de film) : on filmait en 16mm les écrans de télévision ! Ce procédé était lent (il fallait développer les films) et ne donnait qu'une qualité très moyenne.

Il était évident qu'une conservation sur bande magnétique, comme on la connaissait pour le son, était la solution à envisager. Pour les magnétophones à bandes, les vitesses de défilement étaient standardisées : 4,75 - 9,5 - 19 - 38 et 76 cm/sec, les trois premières étant réservées aux appareils "grand public", les vitesses supérieures au matériel de studio. La qualité de l'enregistrement magnétique étant tributaire de la vitesse de défilement de la bande, plus celle-ci est élevée, plus les nuances sonores sont bien rendues. Le principe de l'enregistrement magnétique de l'image est exactement le même que celui du son. Cependant, comme la "quantité d'information" inhérente à la composition de l'image dépasse de beaucoup celle du son, l'enregistrement d'une image

doit s'effectuer à une vitesse nettement supérieure. Pour atteindre celle-ci, on procède de préférence par un mouvement "en deux temps", consistant à déplacer la bande lentement (pour éviter son élongation, par suite des accélérations trop fortes) et à imprimer à la tête de lecture/d'enregistrement un mouvement rotatif rapide.

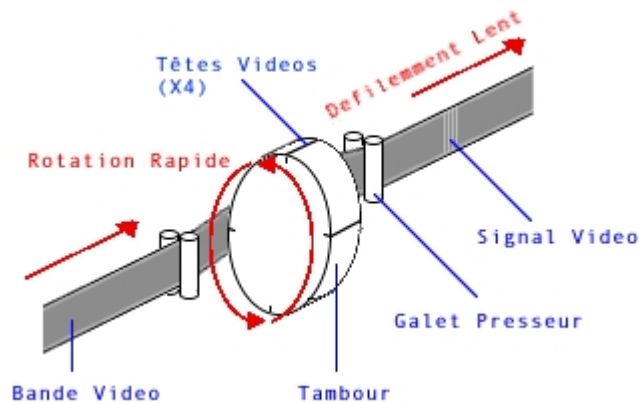
C'est la société américaine **AMPEX** qui proposa la première, en 1956 (pour des machines noir & blanc - les machines couleur furent disponibles en 1978), un système pour la télévision : le **QUADRUPLEX**, qui utilisait des bandes vidéo de 2 pouces de large (50,8 mm).

Le QUADRUPLEX utilisait 4 têtes d'enregistrement, disposées sur un tambour tournant (à 250 tours/min) perpendiculairement au sens de défilement de la bande. La bande défilait à 38 cm/sec, avec une vitesse d'écriture de 39,62 m/sec, et était maintenue incurvée en contact avec le tambour par de l'air comprimé. Le signal vidéo était donc inscrit verticalement sur la largeur de la bande. Lorsqu'une tête arrivait au bas de la bande, la suivante était prête au sommet pour écrire la suite du signal. Chaque tête écrivait 1/10 d'image à chaque rotation. Une piste de contrôle longitudinale reçoit d'une tête fixe, à chaque tour du tambour, une impulsion de synchronisation destinée à piloter les dispositifs d'asservissement de lecture. L'enregistrement du son pour lequel sont réservées deux pistes longitudinales s'effectue d'une manière classique par têtes magnétiques fixes.



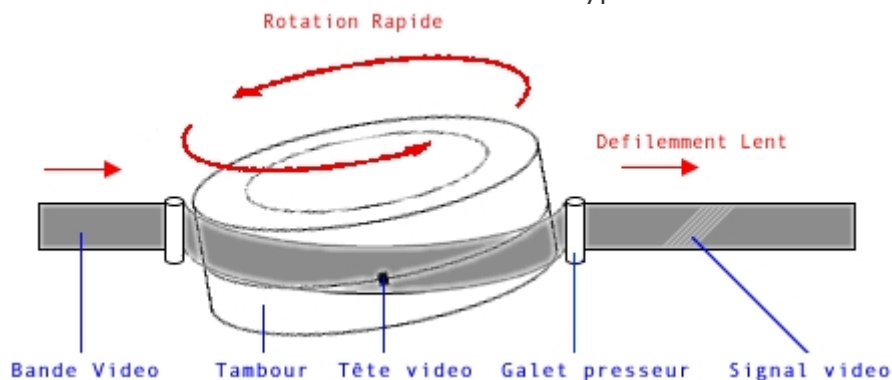
La commutation de signal entre les têtes était le point faible de cette machine : il fallait procéder à de fréquents nettoyage et re-synchronisation des têtes, ce qui occasionnait des interruptions de programmes et l'apparition du message "INTERLUDE" ! Ces machines furent supplantées par le standard 1" en 1980.

Mecanisme Quadruplex de Ampex



Ce problème fut résolu avec le développement de l'enregistrement hélicoïdal, toujours utilisé actuellement, où le tambour de têtes est fortement incliné par rapport à la bande. De cette façon chacune des deux têtes vidéo inscrit un champ (une image vidéo (frame) est composée de 2 champs (field), l'un composé des lignes paires, l'autre des lignes impaires) entier à chaque rotation. La commutation des têtes ayant lieu entre deux champs, la synchronisation est moins critique. Aux bords de la bande se trouvent des pistes longitudinales réservées au son (et plus tard au time code) qui est enregistré par une ou deux têtes fixes, indépendantes du tambour.

Mecanisme Helicoïdale type.



B/Principe

La fréquence maximale enregistrée est fonction de la largeur de l'entrefer de la tête (e) et de la vitesse de défilement de la bande (V).

La largeur de l'entrefer e correspond au pouvoir séparateur du système ; la distance séparant deux alternances sur la bande est :

$$V/2f$$

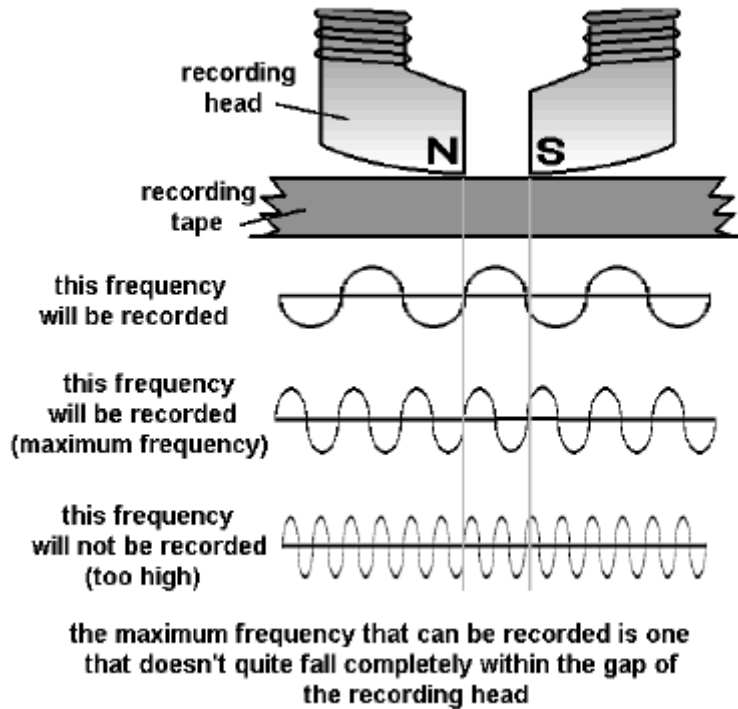
On doit donc avoir : $e < V/2f$

Avec un entrefer de $2\mu\text{m}$, pour enregistrer un signal de 6MHz, il faut une vitesse de défilement de :

$$V > e \cdot 2f = 2\mu\text{m} \cdot 2 \times 6\text{MHz} = 24\text{m/s}$$

Ce qui est beaucoup trop pour un enregistrement linéaire (par tête fixe).

On adopte donc l'enregistrement hélicoïdal. En réduisant la largeur de l'entrefer à $0.5\mu\text{m}$ et la bande passante à 3MHz, on obtient une vitesse d'analyse de la bande de 6m/s.



C/En pratique

Enregistrer un courant de fréquence f , revient à inscrire sur la bande un signal de longueur d'onde λ $\lambda = V/f$ (V étant la vitesse d'écriture)

Plage de fréquence à enregistrer : audio : BP=20KHz
Vidéo : PB= 6MHz

Or la plus petite longueur d'onde enregistrable (la plus haute fréquence) est limitée par la largeur de l'entrefer ; e. $\lambda = 1.4.e$

**Donc, en audio, nous avons : $e=5\mu\text{m}$ $\lambda=1.4 \cdot 5 \cdot 10^{-6}$
 $= 7 \cdot 10^{-6}$
 $V = \lambda.f = 7 \cdot 10^{-6} \cdot 20 \cdot 10^3 = 14\text{cm/s}$**

**En vidéo, nous avons : $e=1.5\mu\text{m}$ (par exemple) $\lambda=1.4 \cdot 1.5 \cdot 10^{-6}$
 $= 2.1 \cdot 10^{-6}$
 $V = \lambda.f = 2.1 \cdot 10^{-6} \cdot 5.5 \cdot 10^6 = 11.55\text{m/s}$**

On en arrive à la même conclusion, la seule solution est l'enregistrement hélicoïdal.

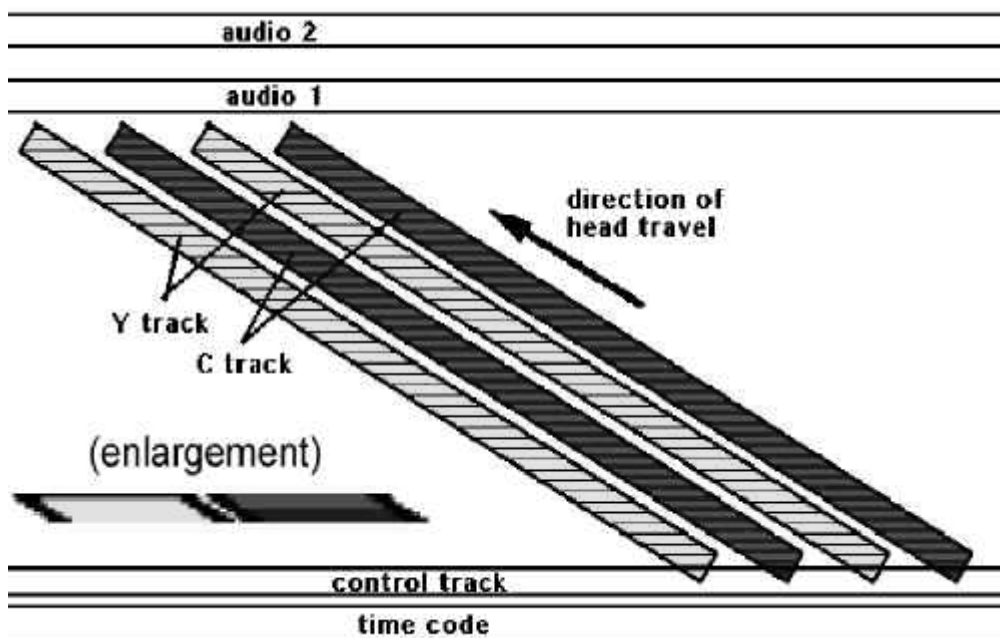
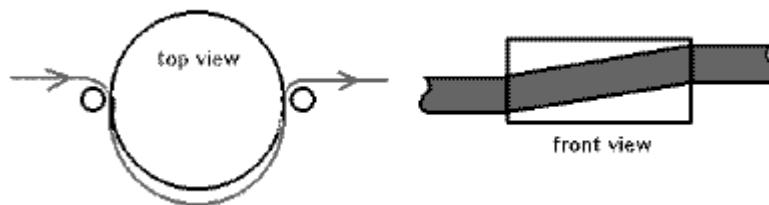
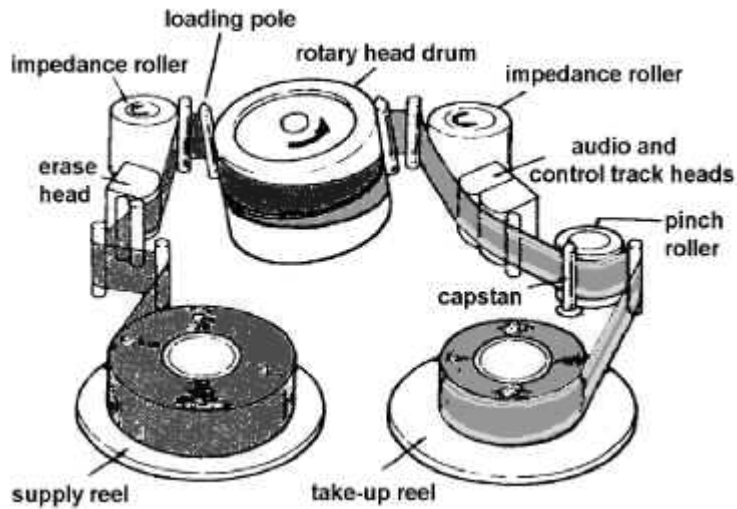
Dans lequel : $V_r = (\pi \cdot D \cdot V_t) \pm V_b$

V_r = vitesse relative entre la bande et la tête

D = diamètre du tambour porte tête

V_t = vitesse de rotation du tambour porte tête (drum)

V_b = vitesse de défilement de la bande (plus ou moins, car suivant les systèmes la bande défile dans le même sens ou en sens opposé à la rotation du tambour)



enregistrement de type bétacam

Le système hélicoïdal repose donc sur des têtes d'enregistrement / lecture disposées sur un tambour (drum) rotatif. Ces têtes inscrivent des pistes obliques sur la bande. En vidéo, chaque piste oblique correspond à une trame.

L'essentiel de la vitesse est fourni par la vitesse de rotation du tambour.

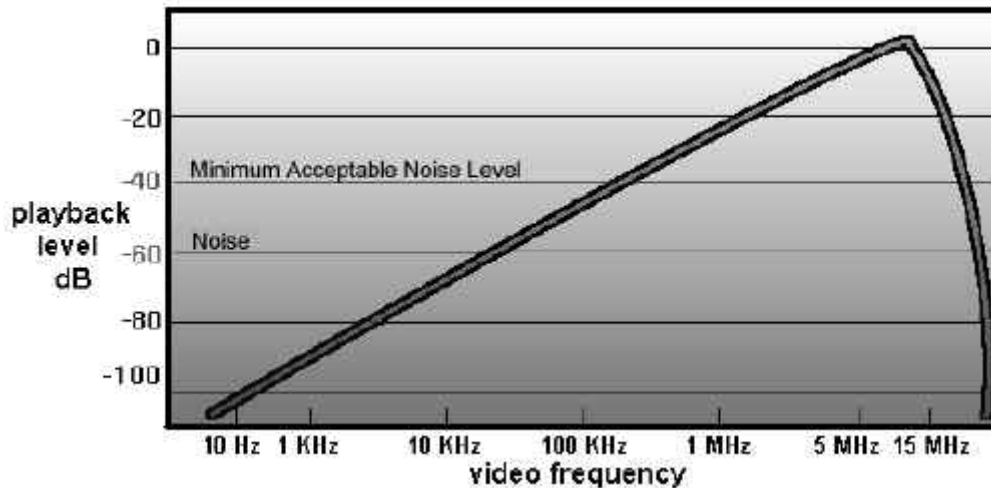
Voilà qui résout le problème de la bande passante...

Le problème suivant est lié à l'enregistrement magnétique, en effet, plus la fréquence du signal enregistré est élevée, meilleure sera la lecture en playback...et vice versa...

Or, en vidéo nous voulons enregistrer un signal qui va de 30Hz à 5 MHz.

En gros, on aura un voltage en playback 166000 plus faible à 30Hz qu'à 5MHz, exprimé en dB, nous obtenons entre la fréquence la plus basse et la plus élevée une différence de 104dB.

($20 \cdot \log 166000$)



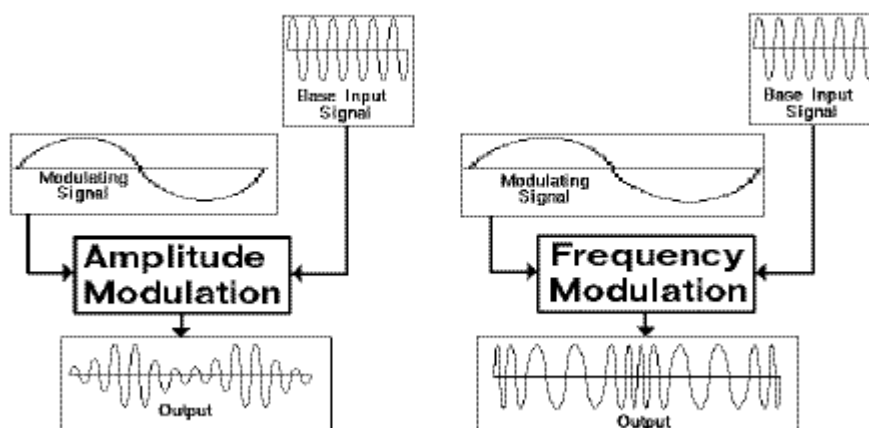
Pour résoudre ce problème, le signal vidéo est modulé en FM (fréquence modulée).

Brève explication sur les modulations

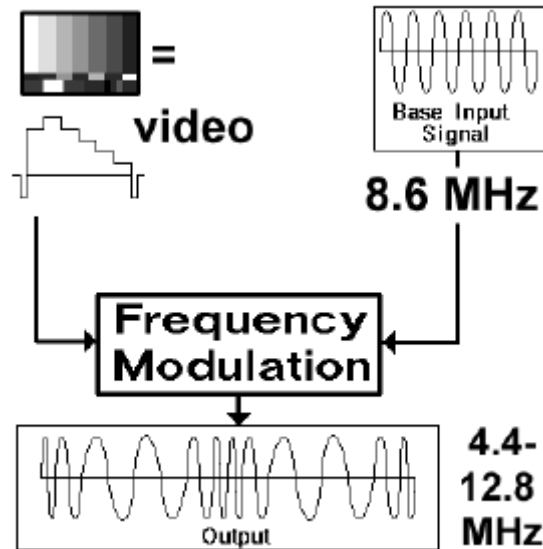
Pour moduler, on utilise un ...modulateur, ce modulateur, reçoit deux signaux ; la porteuse et le signal modulant.

Le modulant peut affecter deux paramètres de la porteuse :

- son amplitude, nous sommes alors en AM (modulation d'amplitude à fréquence constante)
- sa fréquence, nous sommes alors en FM (modulation de fréquence à amplitude constante)



Comme dit plus haut, en enregistrement vidéo, nous allons employer la FM, le signal vidéo va moduler une porteuse à 8.6MHz, le signal enregistré (sortie du démodulateur) variera donc de 4.4MHz à 12.8MHz (en moyenne).
 Ce qu'il est important de constater, c'est que la différence entre la plus basse fréquence et la plus élevée n'est plus que de 10dB (au lieu de 104dB).Ce qui est très acceptable.



A la lecture, le signal en sortie de tête est démodulé, et on récupère le signal vidéo.

Servo system

Lorsque l'on enregistre une bande vidéo, il faut placer un signal qui dira à la lecture ou placer les têtes de lecture pour un playback correct.

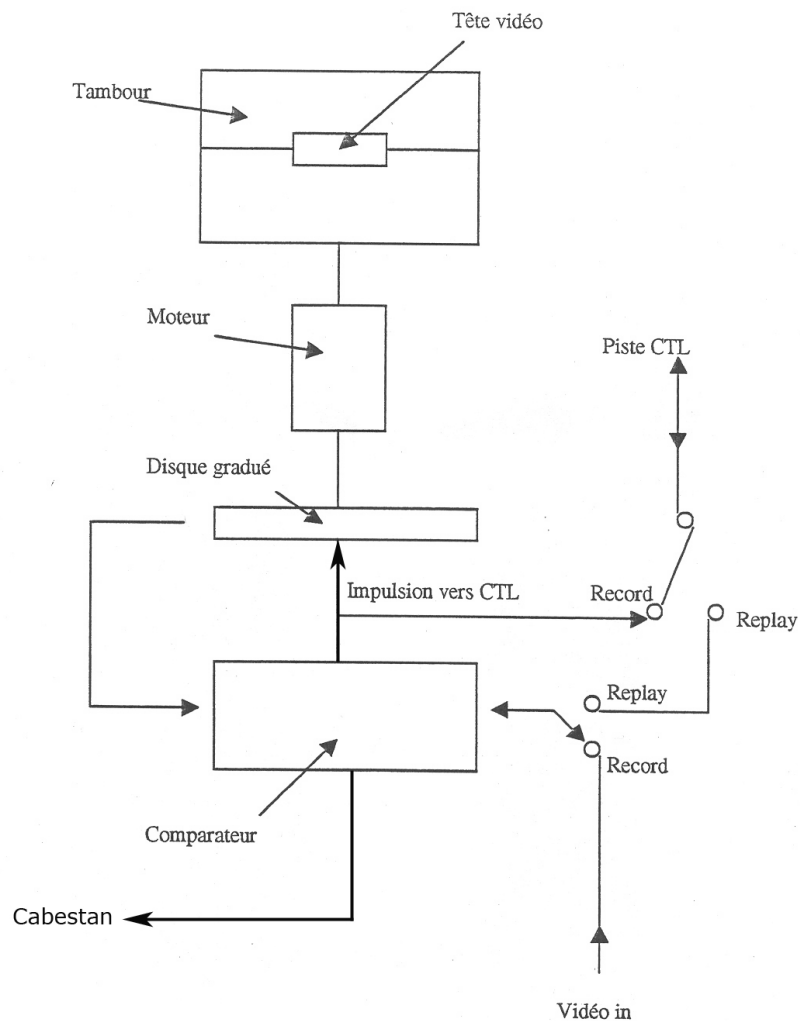
En plus de cela, le moteur du tambour porte tête et le moteur du cabestan (qui donne la vitesse de défilement de la bande) doivent tourner à une vitesse très précise, synchronisée sur le signal vidéo.

Le servo de cabestan (capstan servo) s'assure que la bande défile à vitesse constante dans le VTR (video tape recorder).

A l'enregistrement le moteur du cabestan tourne à une certaine vitesse qui est comparée au timing du signal vidéo entrant (en fait à la sync trame). En même temps une piste longitudinale est écrite sur la bande, la piste de contrôle (control track), cette piste est composée d'impulsions (une impulsion par trame, et donc par piste).

A la lecture le servo de cabestan se base sur la piste de contrôle pour s'assurer que la bande défile à la bonne vitesse.

La piste de contrôle est aussi la base du positionnement correct des têtes vidéo par rapport aux pistes vidéo.



Explication du schéma :

1/ A l'enregistrement , le moteur servant à entraîner les têtes vidéo est synchronisé sur la synchro trame de la vidéo entrante.

La position des têtes est repérée par un disque gradué, solidaire du tambour et du moteur.

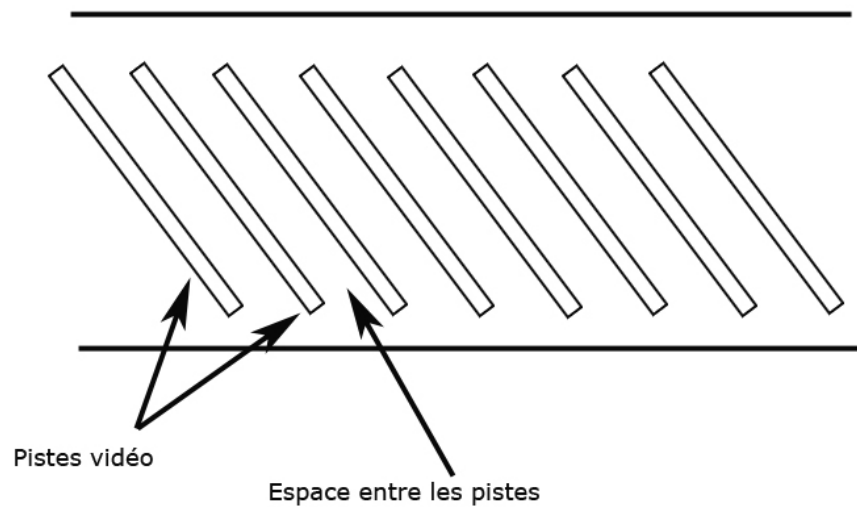
Simultanément à l'enregistrement de la vidéo sur les pistes vidéo (hélicoïdales), une impulsion est enregistrée sur la piste ctl lors d'une position particulière des têtes vidéo.

2/ Lorsque la bande est relue, l'impulsion ctl est comparée à la position du disque gradué, le comparateur agit alors sur la position du tambour et sur le cabestan pour repositionner la bande et les têtes vidéo pour obtenir la même synchronisation qu'à l'enregistrement.

L'azimut.

Pour qu'un enregistrement soit reproduit dans des conditions optimales, il faut 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, dans le cas contraire, il y a affaiblissement rapide du signal (erreur d'azimut).Donc, l'entrefer doit être perpendiculaire à l'axe de la piste.

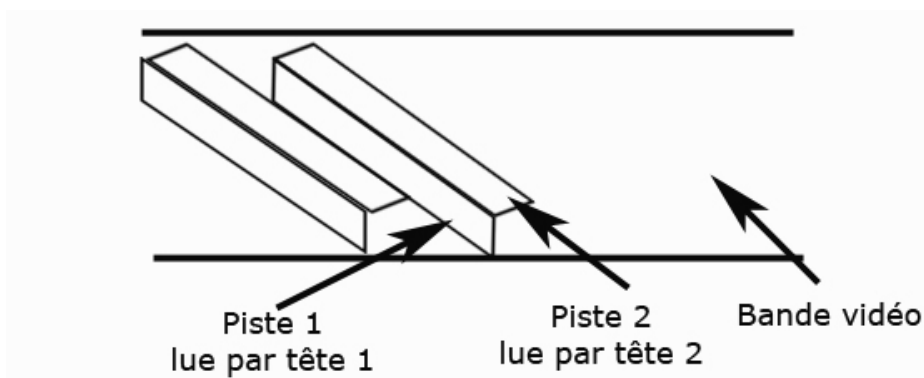
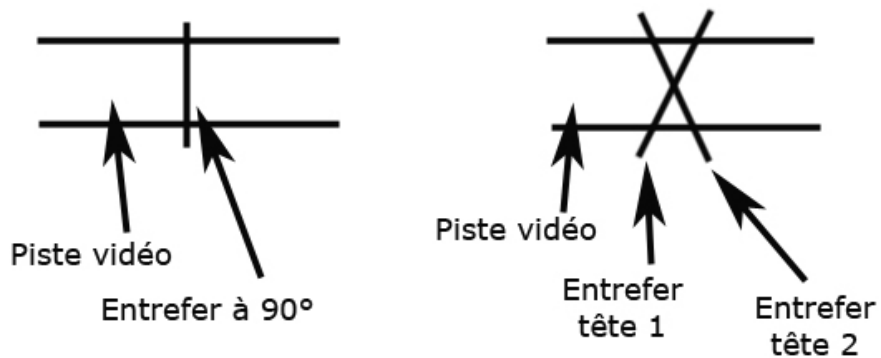
Si les pistes sont trop rapprochées, il y a risque qu'une tête lise des informations d'une piste voisine, on doit donc laisser un espace entre les pistes.



A titre d'exemple, si une piste fait $100\mu\text{m}$ de largeur, l'espace entre les pistes devra être de $40\mu\text{m}$. (le schéma n'est pas à l'échelle).

Le problème est que cet espace prend énormément d'espace sur la bande, et donc, la bande doit défiler plus vite.

Solution, exploiter l'effet d'azimut pour supprimer cet espace, les pistes contiguës ont des azimuts opposés.



Les deux têtes de lecture (et enregistrement montées sur le tambour ont des azimuts opposés, donc, même sans espace entre les pistes, il est impossible qu'une tête 1 lise des infos d'une tête 2.

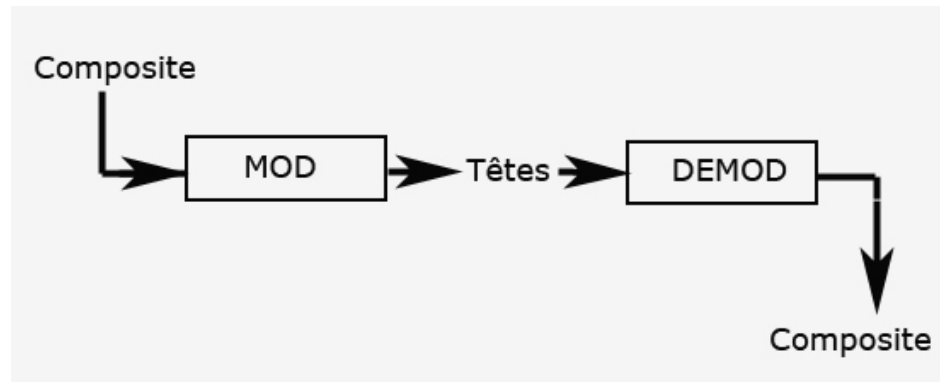
Méthodes d'enregistrement

1/ L'enregistrement direct du composite (2'', 1''B, 1''C)

Le signal composite (BP 5,5MHz) est directement enregistré en modulation de fréquence.

f1= fond de synchro f2= blanc

dans ce type d'enregistrement, la vitesse relative est très élevée, de l'ordre de 20m/s



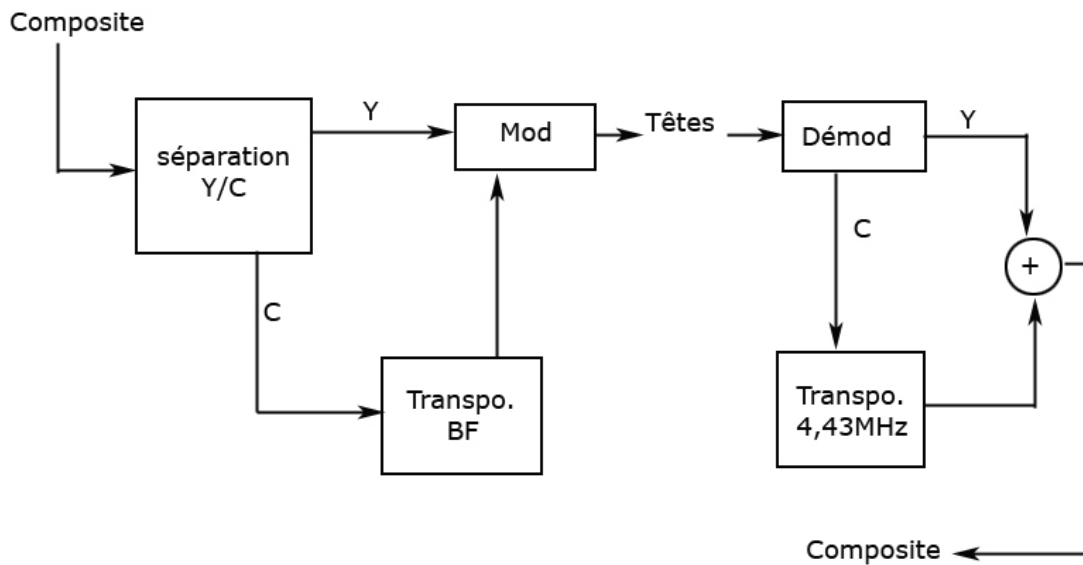
1'' VTR

2/ Procédé under color (VHS, U Matic, 8mm)

Ce procédé est techniquement nécessaire lorsque la vitesse relative n'est pas suffisante pour enregistrer toute la bande passante sans détériorer les HF, or la chrominance se trouve à 4,43MHz.

La technique consiste à transposer le signal de chroma dans les BF (en dessous de 1MHz).

- la luminance ne dépasse pas 3 MHz
- la chrominance 0,6MHz



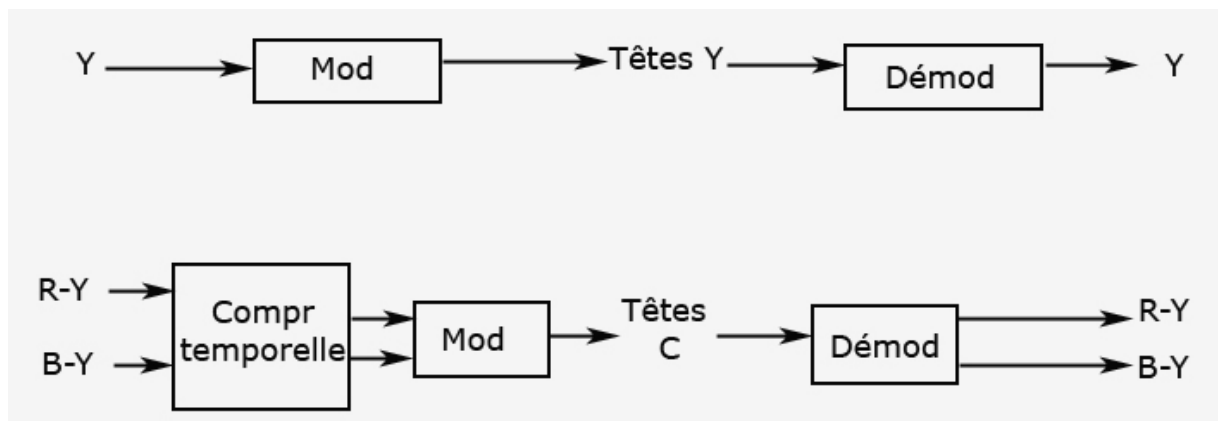
U-matic (3/4")

3/ Procédé composante.

Ici des têtes différentes sont allouées aux signaux de luminance et de chrominance.

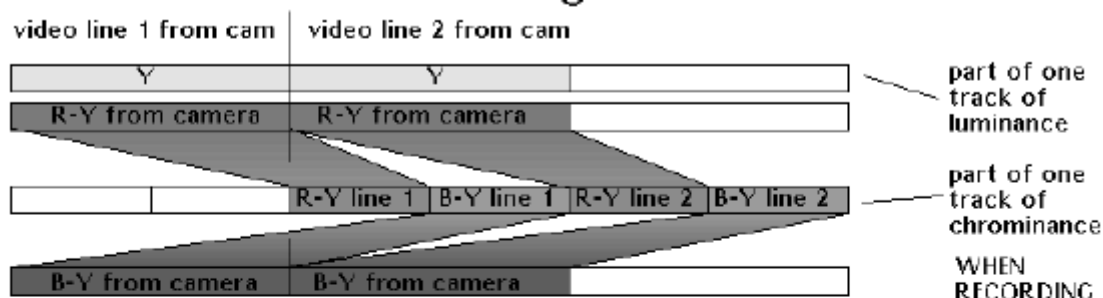
BP de la luminance = 5,5MHz

BP de la chrominance = 1,5MHz

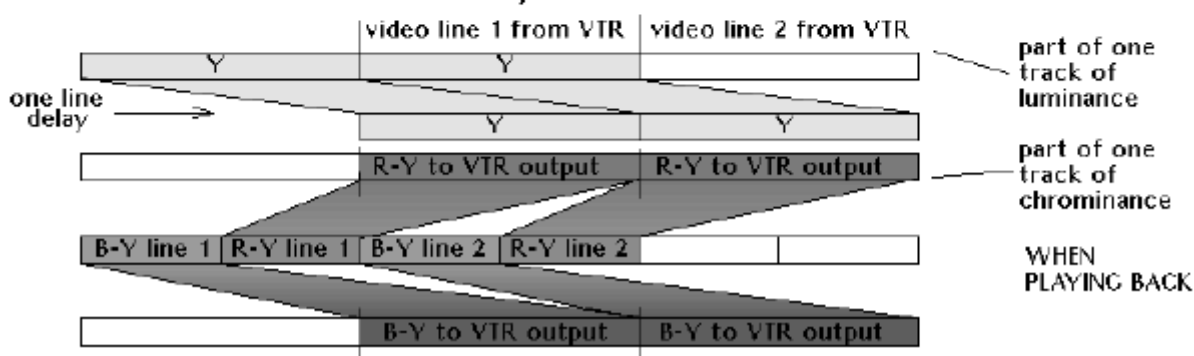


La compression temporelle des deux composantes de chrominance est nécessaire pour les faire tenir sur une seule piste vidéo, on a donc une piste pour la luminance et une piste pour la chrominance.

Betacam Recording CTDM



Betacam Playback CTDM



Ce procédé s'appelle CTDM (Component Time Division Multiplexing)

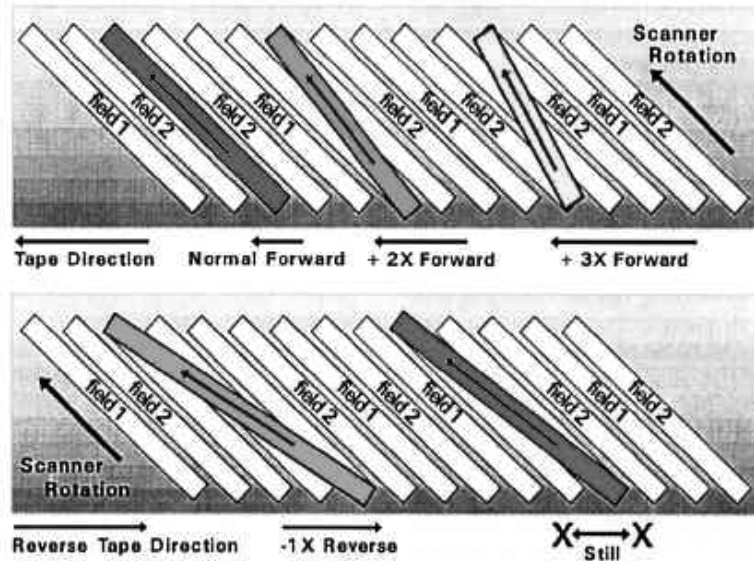
Le suivi dynamique des pistes (dynamic tracking)

Le suivi dynamique est nécessaire lorsque l'on désire avoir une image correcte à une vitesse autre que la vitesse nominale.

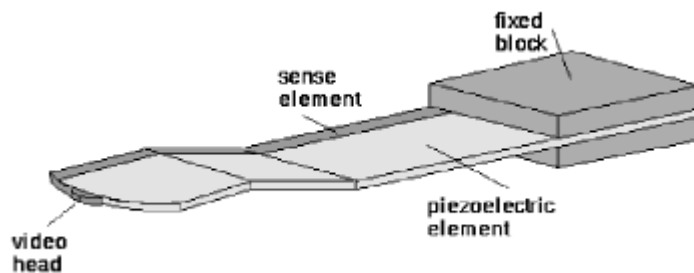
Si la vitesse de défilement de la bande varie, l'angle d'inclinaison relatif des pistes vidéo change par rapport à l'angle des têtes de lecture, qui lui est constant.

Les schémas page suivante donnent les parties de bande lues par les têtes vidéo suivant la vitesse de la bande par rapport à sa vitesse nominale.

Sur une machine sans DT, l'image sera composée de parties d'images venant de différentes pistes, séparées par du bruit.



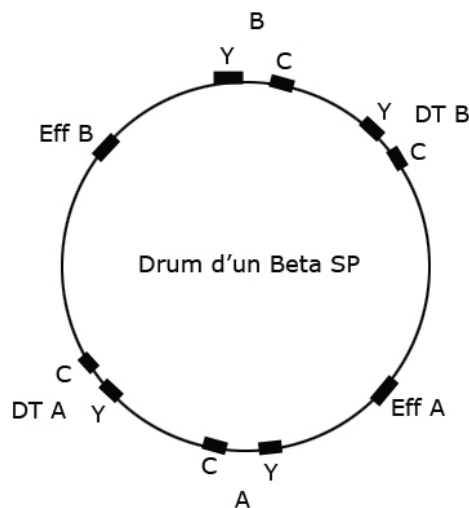
Le DT intègre un système de têtes « flottantes » placées sur un dispositif piézoélectrique.



La tension appliquée au système varie selon la vitesse de la bande, le système piézoélectrique recevant cette tension modifie donc l'inclinaison des têtes DT.

Les têtes DT prennent le relais des têtes normales lorsque la vitesse de la bande n'est plus nominale .

Le système gère aussi le saut de piste, par exemple, en X2, seule une piste sur deux est lue.



Types de bandes

Les différents types de bandes se différencient surtout par la finesse de leurs particules magnétiques, or plus les particules sont fines, plus la fréquence enregistrée pourra être élevée.

Bande oxyde : elle renferme des particules d'oxyde métallique maintenues en suspension dans un liant (VHS).

Bande métal particule : couche magnétique composée d'une poudre de fer pur (les particules de fer sont plus petites que les particules d'oxyde).ce qui permet d'augmenter la quantité d'information stockée sur une même surface (Beta SP)

Bande métal évaporé : la couche magnétique est déposée par évaporation sous vide, directement sur le support, sans liant (dix fois plus fin que le métal particule).
(DV, DVCAM). Ces bandes sont aussi moins robustes. (usure plus rapide).

Le TBC (Time Base Corrector, correcteur de base de temps)

Le TBC est un appareil qui est utilisé pour éliminer les instabilités mécaniques des magnétoscopes et pour les synchroniser avec d'autres appareils (magnétos, caméras, mix...)

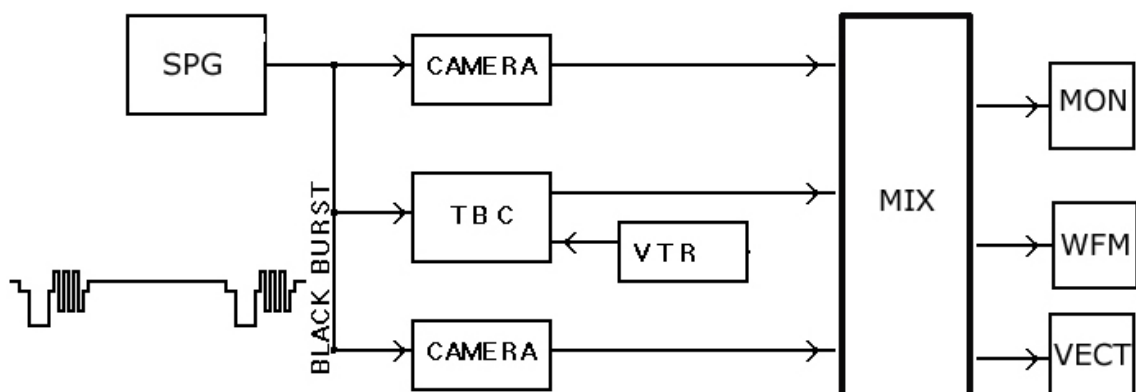
Sans TBC, il est impossible de mixer plusieurs sources vidéo.

Le TBC peut être soit externe, soit inclus dans le magnétoscope, soit inclus dans la table de mixage vidéo.

Dans sa version la plus simple, il s'agit d'une mémoire tampon qui enregistre le signal vidéo et le restitue synchrone avec les autres appareils. La synchronisation se fait à l'aide d'un waveform et d'un vecteurscope, le délai est réglable sur le TBC.

En pratique, un SPG (sync pulse generator) synchronisera tout les appareils (TBC, MIX...) en leur envoyant un signal de synchronisation appelé BlackBurst, qui est un signal vidéo noir (synchro+burst).

Ce signal rentre dans les différents appareils en Genlock in ou Ref in.



En plus de fonctions de synchronisation, la plupart de TBC offrent aussi le réglage du niveau vidéo (voltage du blanc), le niveau du noir et la saturation de la couleur.

Autre fonction importante, le DOC (drop out compensator), en reliant la sortie RF (sortie tête) du vtr au tbc, le tbc détecte l'absence de signal (puisque le signal des têtes est modulé en FM) et remplace la portion manquante par la même portion de la ligne précédente (mémorisée puisqu'il s'agit d'une mémoire tampon).

4/NUMERIQUE

A/ Principe

- Le signal analogique est celui qui représente le plus fidèlement les variations d'un phénomène physiologique.

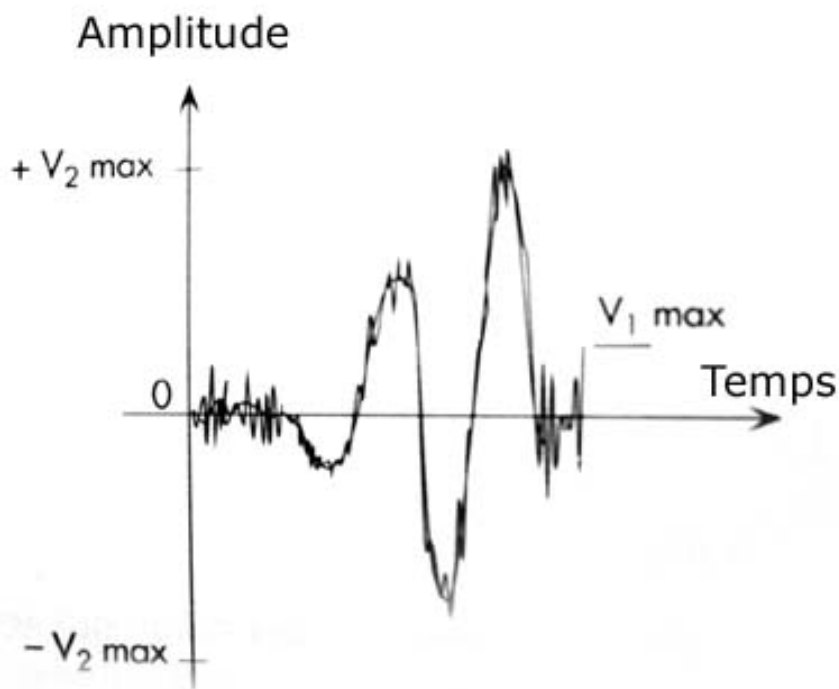
-Problème : le bruit (souffle, parasite) qui se superpose au signal lors de l'enregistrement ou diffusion...

-le bruit est quantifié par le rapport entre l'amplitude maximale du signal d'origine et celle du bruit qui lui est ajouté. (rapport signal/bruit en dB)

40 dB en VHS

60 dB en bétacam SP

S/B (rapport signal/bruit) = $20 \log V_2 \max(\text{signal d'origine}) / V_1 \max(\text{bruit})$

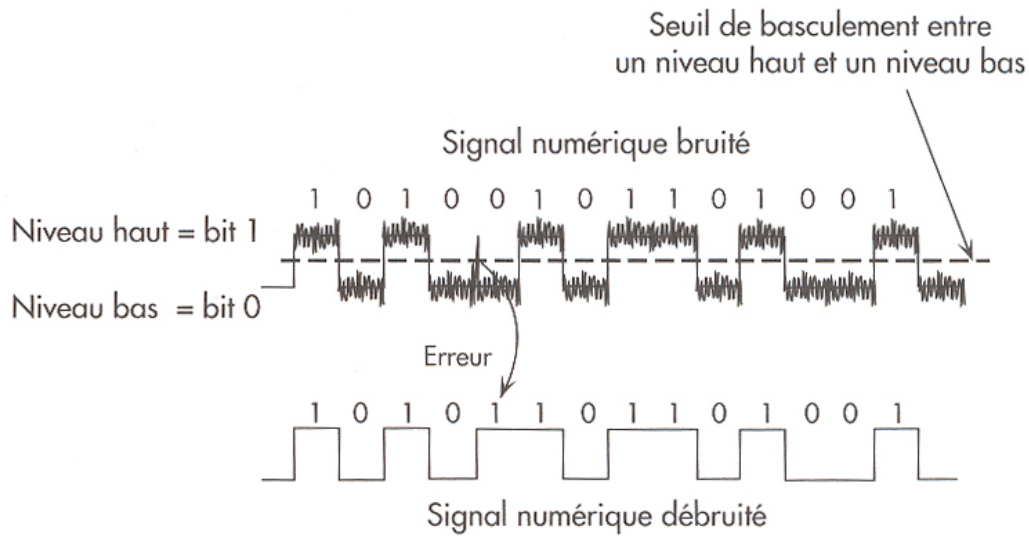


En première génération, le bruit n'est pas très gênant, le problème est qu'il s'accumule au fil des générations (copies).

En effet, lors de la copie d'un signal « soufflé » il est impossible de distinguer ce qui est signal original et ce qui est souffle puisque le signal analogique est imprévisible.

- Le langage binaire (base du numérique) se compose d'une suite de niveaux 1 ou 0, il est donc prévisible, après un 1 vient soit un 1 ou un 0, même chose après un 0. Le signal binaire peut donc être régénéré après chaque copie (élimination du souffle). Ce qui autorise un grand nombre de traitements sans pertes .

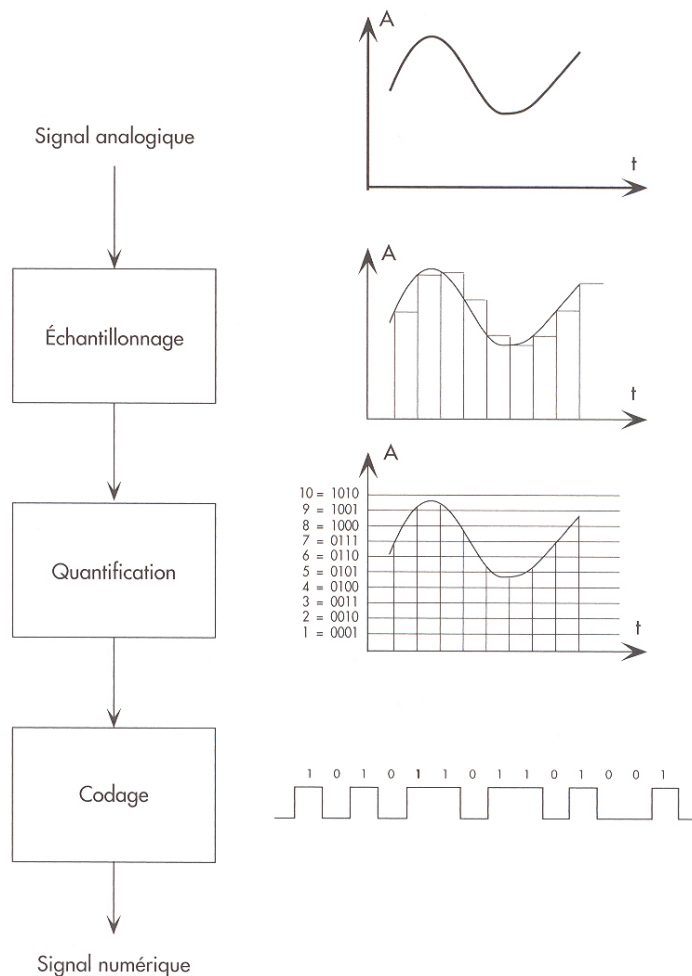
- Précisons que les capteurs ccd d'une caméra sont des dispositifs numériques traitant des signaux analogiques.

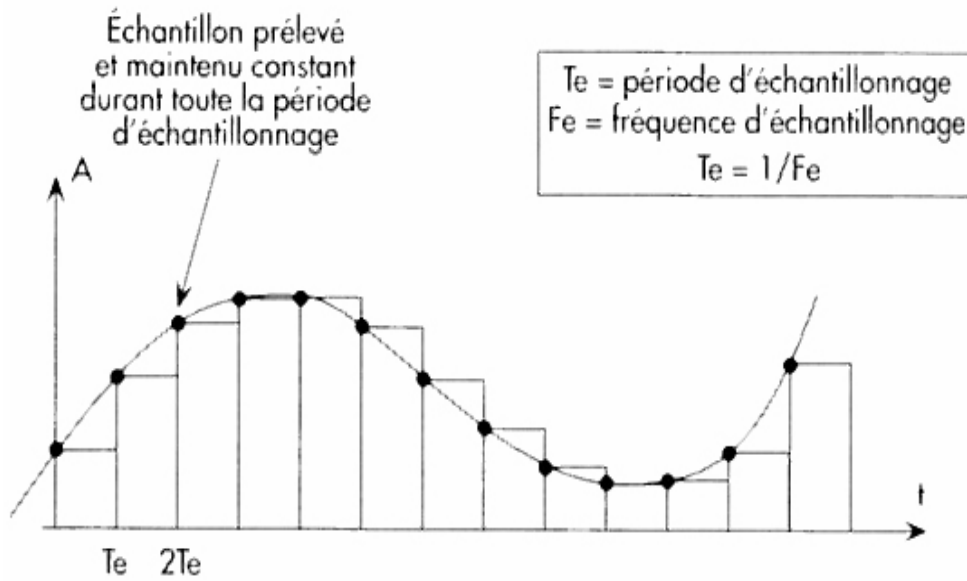


B/La conversion analogique / numérique (digital en version anglaise)

- 1/ l'échantillonnage

on prélève ponctuellement l'amplitude du signal analogique, et on garde cette amplitude jusqu'à l'échantillon suivant.

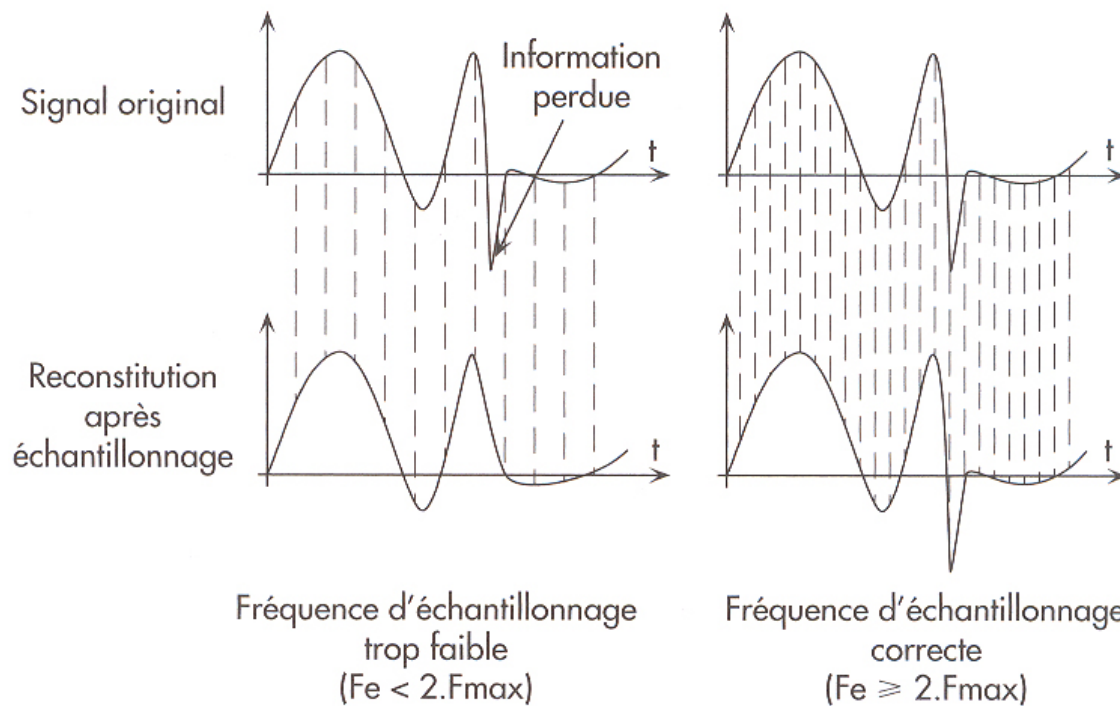




- Si l'on respecte la loi de Shannon et Nyquist, l'information après échantillonnage sera une représentation intégrale du signal original.

La loi de SN permet de déterminer la fréquence d'échantillonnage minimale à choisir.

- la fréquence d'échantillonnage doit être supérieure ou égale à deux fois la fréquence maximale à échantillonner -



- La fréquence d'échantillonnage doit donc tenir compte de la bande passante du signal vidéo qui est théoriquement de 6,5Mhz.

Ce qui donne une fréquence d'échantillonnage de 13Mhz minimum

-En outre il était souhaitable que cette norme tienne compte des différents systèmes analogiques existants (PAL, NTSC ect)

- la fréquence ligne en 625 L est de 15625 Hz
- La fréquence ligne en 525 L est de 15734 HZ

Après divers tests une fréquence d'échantillonnage de luminance égale à 13,5 MHz a été choisie.

13,5 MHz = 864 x syst 635

= 858 x syst 525 et répond à la lois de SN.

-Les signaux de chrominances sont échantillonnés à une fréquence plus faible égale à 6,75 MHz.

- Cette norme a été nommée 4 :2 :2

- ces nombres représentent le rapport entre la fréquence d'échantillonnage de la luminance et des composantes de chrominance.

- le 4 remonte au début des travaux de numérisation, lorsque l'on envisageait de numériser du composite, la fréquence d'échantillonnage était de 17,7 MHz , soit 4 x la fréquence de la sous porteuse couleur.

- le système a été abandonné, mais le nom est resté.

- le 4 :2 :2 est la norme professionnelle, il en existe d'autres, plus légères ou de haute qualité.

Déclinaisons du 4 :2 :2

- **4 :2 :2** la luminance est échantillonnée à 13,5 MHz, les composantes de chrominance à 6,75 MHz (bétacam digital)

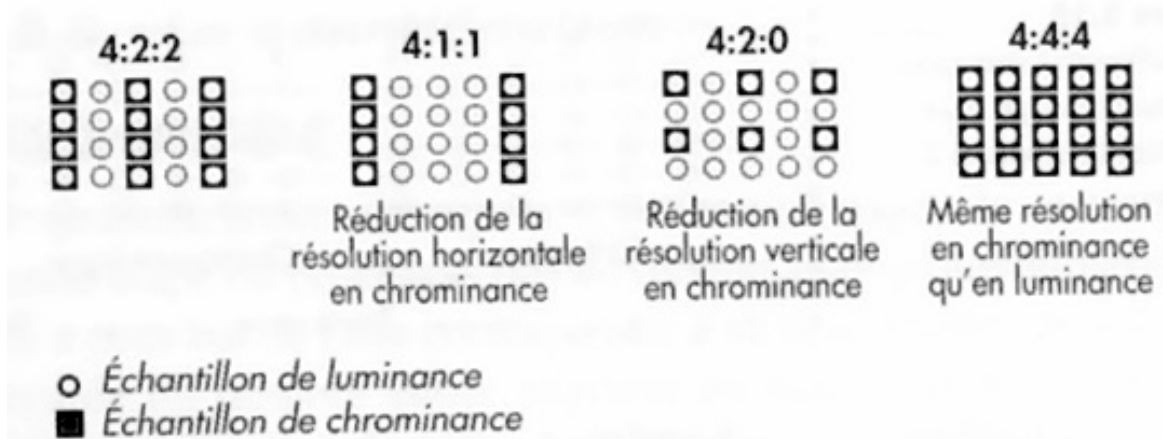
- **4 :2 :2 :4** Codage 4 :2 :2 incluant un signal de découpe échantillonné à 13,5 MHz

- **4 :1 :1** La luminance est toujours échantillonnée à 13,5 MHz, mais les composantes de chrominance sont échantillonnées à 3,375 MHz (soit ¼ de la luminance)
ex : DVCPRO25

- **4 :2 :0** Luminance échantillonnée à 13,5 MHz, chrominance à 6,75 Mhz mais une ligne sur deux.
Ex : DVD, DV, DVCAM

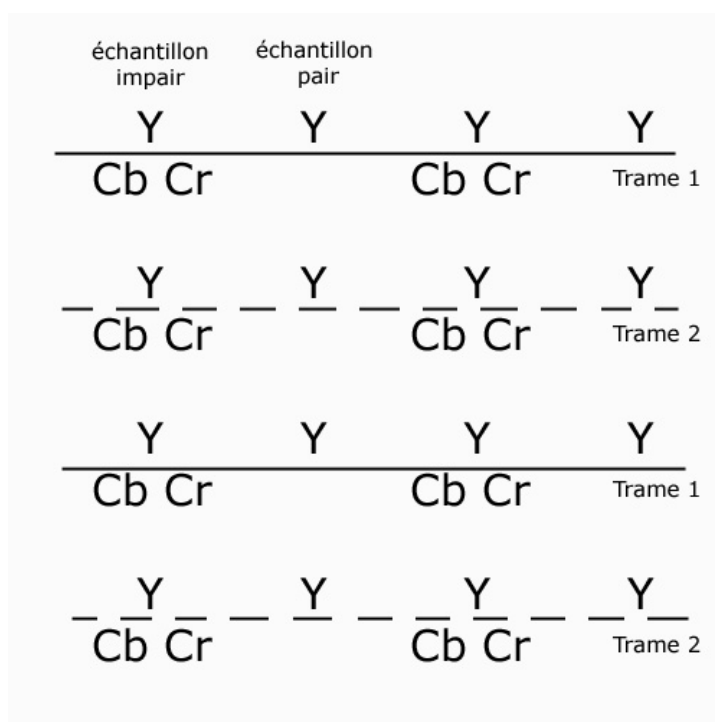
- **4 :4 :4** Toutes les composantes sont échantillonnées à 13,5 MHz, ce type de codage est employé par les stations graphiques .

- **4 :4 :4 :4** le précédent, incluant un signal de découpe à 13,5 MHz.



Structure d'échantillonnage

- La structure d'échantillonnage est la répartition des échantillons de luminance et de chrominance dans le temps et l'espace.
- Pour le numérique, on a souhaité éliminer la périodicité à 8 trames qui existait en pal, et revenir à une périodicité à 2 trames, ce qui sous entend que les images soient de structure identique.
- La structure orthogonale a été retenue après divers tests : les échantillons sont situés aux mêmes emplacements d'une ligne à l'autre et d'une trame à l'autre.



2/ La Quantification

- La quantification est la deuxième étape du processus de numérisation.
- elle a pour but de faire correspondre chaque amplitude échantillonnée à un nombre entier exprimé en base 2.

- n chiffres permettent de coder $N = 2^n$

ex codage sur 3 bits :

4	2	1
I	I	I

Dans ce cas, le nombre de niveaux possibles est de 7 (8 en comptant le 0).

- Dans cette étape le signal analogique qui peut prendre une infinité de valeurs, est converti en un signal constitué d'un nombre fini de valeurs numériques.

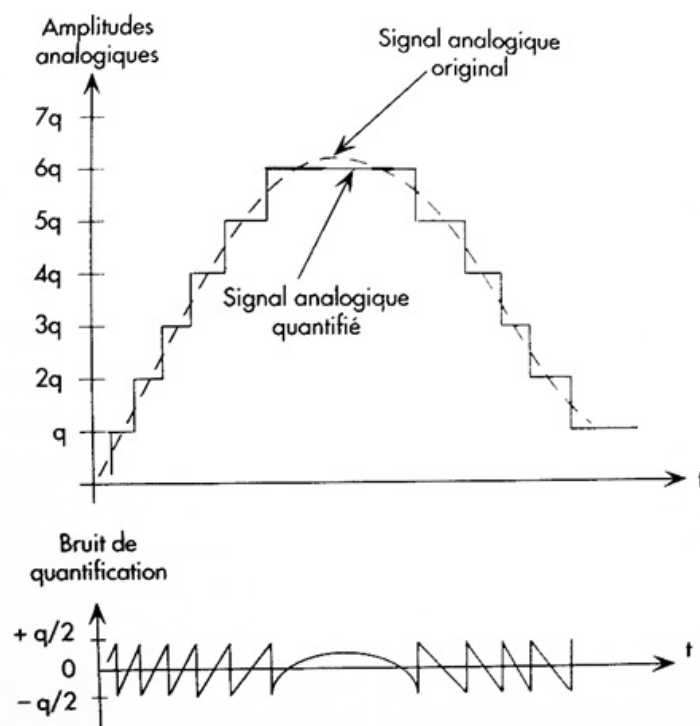
- Des erreurs d'arrondis par défaut ou par excès sont donc inévitables, car à plusieurs valeurs proches, mais différentes, correspondra une seule valeur binaire .

- la précision du signal converti sera donc directement liée au nombre de valeurs disponibles.

- en fonction de la plage à quantifier, on définit une échelle constituée d'un nombre fini d'intervalles « q » appelés « pas de quantification » ou « échelons de quantification » ou « quantum » .

- l'erreur de quantification ne dépasse jamais « $q/2$ » .

- La succession des erreurs engendrées s'appelle « bruit de quantification »



- A l'origine, la norme 4 :2 :2 spécifiait une quantification sur 8 bits des composantes du signal vidéo, ce qui permet de disposer de 256 niveaux possibles.

- Par la suite, pour les productions haut de gamme, la norme a évolué vers le 10 bits, ce qui permet 1024 niveaux.

3/ Les corrections d'erreurs

Si toute la chaîne de traitement vidéo fonctionnait idéalement, les signaux numériques pourraient être enregistrés tels quels, ce n'est malheureusement pas le cas.

- les erreurs peuvent provenir :
 - des supports d'enregistrement
 - des mécanismes d'entraînement des magnétoscopes
 - voies de transmission
 - etc...

- en numérique, une erreur est soit une inversion de données (1 devient 0), soit une absence de donnée.

- la correction par bit de parité :

Données originales

0	1	0
0	1	1
1	1	0

Ajout du bit de parité

0	1	0	1
0	1	1	0
1	1	0	0
1	1	1	

Le but est d'ajouter une valeur qui rend le nombre de « 1 » pair en ligne et colonne .

Si un nombre de « 1 » impair est trouvé, c'est qu'il y a erreur, le système permet de retrouver le bit erroné et de le corriger.

0	1	0	1
0	1	1	0
1	0	0	0
1	1	1	

Erreur en deuxième colonne et troisième ligne.

- L'entrelacement des données

La plupart des erreurs qui perturbent un signal numérique affectent en général plusieurs dizaines de bits consécutifs, c'est pourquoi a été inventé le principe du brassage de données.

- Le but est d'éloigner des données à l'origine consécutives, dans le but de disperser tout paquet d'erreur.
- Les erreurs deviennent donc plus petites et dispersées.
- Les données sont dispersées suivant un code connu du codeur et du décodeur.

C/PRINCIPES DE COMPRESSION NUMERIQUE

Pourquoi compresser ?

- Le signal numérique présente d'énormes avantages, mais consomme énormément d'espace.

Ex : une image codée en 4 :2 :2 sur 8 bits occupe un espace de 830 Kilo-octet (partie active) ce qui fait 21 Mbytes la seconde.

- autrement dit sur un CDROM de 640 MB, on pourrait stocker 30 sec de vidéo non compressée, sur un DVD de 4,7GB, 4mn.

Si l'on veut diffuser et stocker de grandes quantités de données vidéo, il est nécessaire de compresser.

Comment compresser ?

- L'art de la compression numérique en vidéo est de supprimer certaines informations de l'image et d'en simplifier d'autres.

2 types de compression.

1/ La compression sans pertes ou « lossless » effectue un traitement transparent, permettant de retrouver intégralement les données d'origine après décompression.

Malheureusement, elle ne conduit qu'à un taux de compression très faible.

2/ La compression avec pertes ou « lossy » aboutit à un taux de compression beaucoup plus élevé, mais impose de négliger certaines informations de l'image.

Si elle se fait dans des proportions raisonnables, la compression peut passer inaperçue, on parle alors de compression virtuellement transparente.

Si la compression est élevée, elle entraîne des distorsions et artéfacts plus ou moins visibles.

Défauts typiques de la compression.

- effet de bloc : des blocs de pixels deviennent visibles (aspect de mosaïque) taux de compression excessif.
- effet de halo : une frange apparaît sur les contours. Débit trop faible
- effet de blurring : détails moins nets. Compression trop élevée.
- bruit de quantification « effet vitre sale » , le bruit de quantification du à la digitalisation s'est accumulé lors de plusieurs cycles de conversion A/D et D/A.
- Mosquito noise : du bruit apparaît sur des transitions d'éléments en mouvement, sous la forme de points noirs. Quantification sur un nombre de bits trop faible.

Les normes de compression vidéo

Il existe deux formes de compression en vidéo : la compression spatiale ou JPEG et la compression temporelle ou MPEG.

1/ Le JPEG et MJPEG

Le JPEG est à la base une norme de compression pour images numériques .

Le MJPEG est une adaptation pour la vidéo, celle-ci n'étant qu'une succession rapide de photos. (M = motion).

Ce système n'ayant jamais été normalisé, les fabricants ont développé des systèmes incompatibles entre eux.

Depuis est apparu le système DV qui utilise une compression de type MJPEG, mais dont l'algorithme à été normalisé.

La compression JPEG se décompose en 6 étapes :

a/ la décomposition en blocs :

La première étape consiste donc à découper l'image en macroblocs 8 pixels sur 8. Chaque bloc comprend donc 64 nombres variant de 0 à 255 (en 8 bits).

b/ transformée en cosinus discret (DCT)

Appliquer une DCT à ces macroblocs revient à les faire passer du domaine temporel au domaine fréquentiel.

- dans le domaine temporel les valeurs sont décrites les unes à la suite des autres, c'est l'état initial du signal.
- dans le domaine fréquentiel, le signal est décrit par son spectre, par ses composantes spatiales, représentant tout les motifs géométriques qu'il est possible de représenter avec un bloc de 64 valeurs.

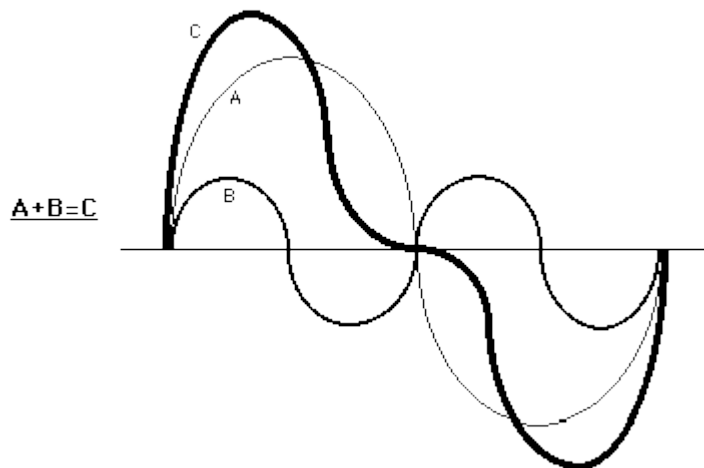
Le passage du domaine temporel au domaine fréquentiel mérite une petite explication...

L'image étant quelque chose de compliqué, nous commencerons par quelque chose de simple, le son (pas sur que les profs de son confirment...).

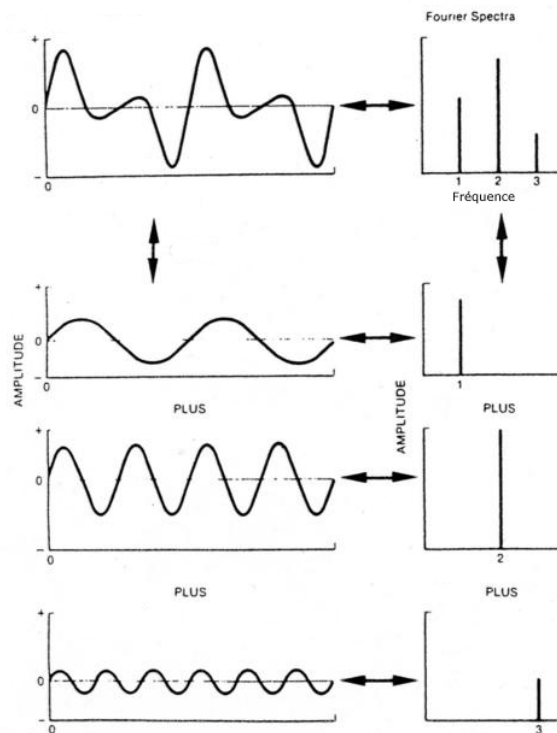
La DCT est basée sur la loi de Fourier (1801).

Par cette loi, il est démontré mathématiquement qu'un mouvement périodique complexe se décompose toujours en une somme de sinusoïdes

Exemple : Addition de deux sinusoïdes en phase, la deuxième ayant une fréquence double et une **amplitude moitié moindre** relativement à la première



Ces différentes sinusoïdes peuvent aussi se représenter sous forme fréquentielle (spectre fréquentiel) : Exemple avec 3 sinusoïdes :

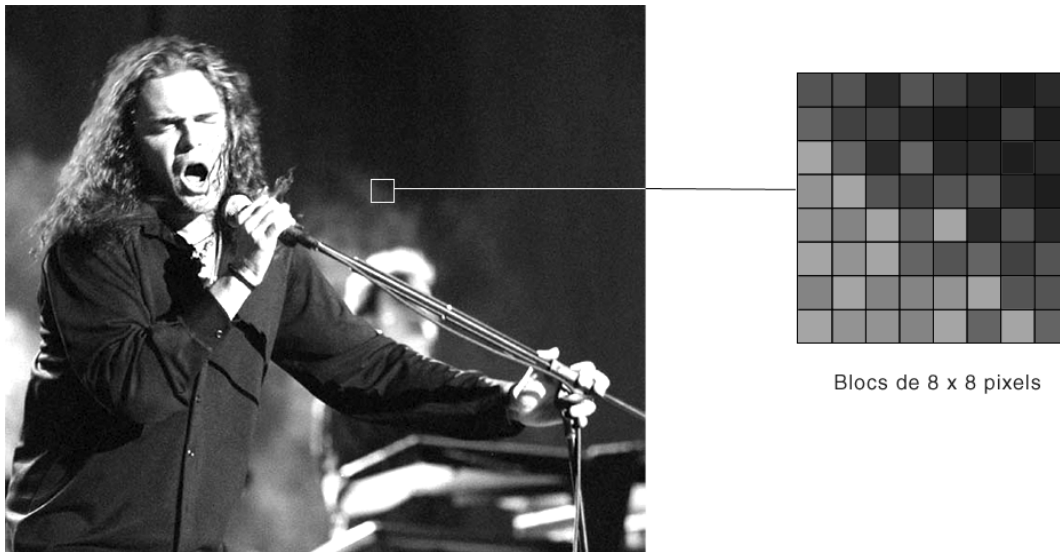


On voit donc dans le schéma de la page précédente qu'un signal complexe peut être décomposé en une série de fréquence simple que l'ont peut visionner sous forme de spectre fréquentiel.

On fonctionnera de la même façon en image (avec des blocks de 8 X 8 pixels) ce qui nous donnera une vision spectrale en deux dimensions.

En image une haute fréquence représente du détail fin, une basse fréquence des plages uniformes.

Donc chaque canal de l'image (y, r-y b-y) sera divisé en block de 8/8 pixels.



Pour une image noir et blanc 8 bits, chaque pixel a une luminosité comprise entre 0 et 255. La mémoire allouée à chacun de ces blocs est donc de :
 $64 \text{ (pixels)} \times 8 \text{ (bits)} = \mathbf{512} \text{ bits.}$

Chaque bloc sera converti en fréquence par la DCT. L'encodeur ne sait pas à l'avance quelles fréquences sont présentes dans le bloc, il doit donc couvrir toutes les fréquences possibles avec un bloc de 64 pixels.

Ce qui veut dire que chaque bloc est encodé comme une combinaison linéaire de 64 motifs dont les fréquences croissent du haut vers le bas et de la gauche vers la droite .

Comme nous encodons un bloc de 64 pixels et que nous avons 64 motifs possibles, chaque motif se verra assigné d'un coefficient qui représentera l'importance de ce motif dans le bloc analysé.

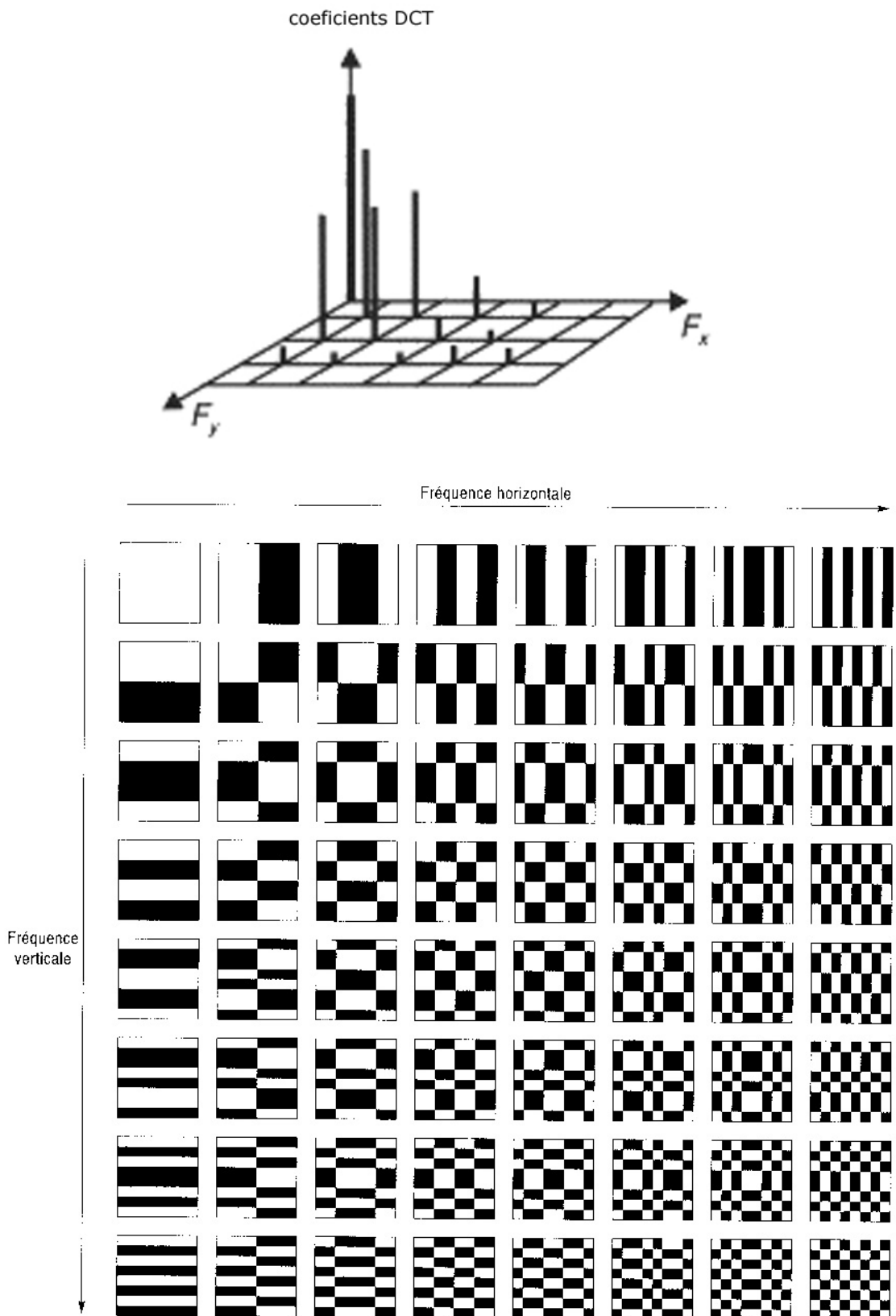


Figure 3.4 : Illustration de la contribution de chacun des coefficients à l'aspect du bloc.

Notons que la DCT ne compresse pas, elle ordonne fréquentiellement, elle facilite la compression, elle est de plus totalement réversible.

La compression JPEG se base sur le fait que dans une image courante, il y a en général beaucoup plus de basses fréquences que de hautes fréquences, les coefficients les plus élevés se trouveront en général en haut et à gauche.

La DCT a donc la propriété remarquable de concentrer l'énergie du bloc sur peu de coefficients situés dans l'angle supérieur gauche.

Pour les petits curieux, voici la formule mathématique d'une DCT et d'une DCT inverse :

DCT

$$Img(x,y) \rightarrow F(u,v) \quad F(u,v) = \frac{2}{N} c(u).c(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} Img(x,y) \cdot \cos\left[\frac{\pi}{N} u \left(x + \frac{1}{2}\right)\right] \cdot \cos\left[\frac{\pi}{N} v \left(y + \frac{1}{2}\right)\right]$$

DCT inverse

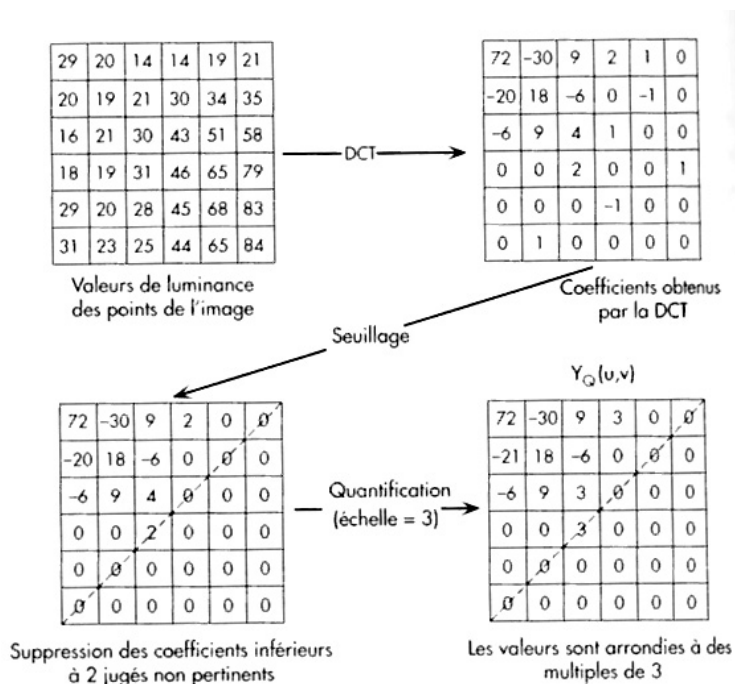
$$F(u,v) \rightarrow Img(x,y) \quad Img(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} c(u).c(v).F(u,v) \cdot \cos\left[\frac{\pi}{N} u \left(x + \frac{1}{2}\right)\right] \cdot \cos\left[\frac{\pi}{N} v \left(y + \frac{1}{2}\right)\right]$$

Vu comme ça, c'est évident..

c/ Quantification des coefficients DCT.

La vision humaine étant moins sensible aux détails fin qu'aux plages uniformes, on va « sous coder » les fréquences spatiales élevées.

Les valeurs situées dans les hautes fréquences seront arrondies, codées sur peu de bits, parfois même éliminées.

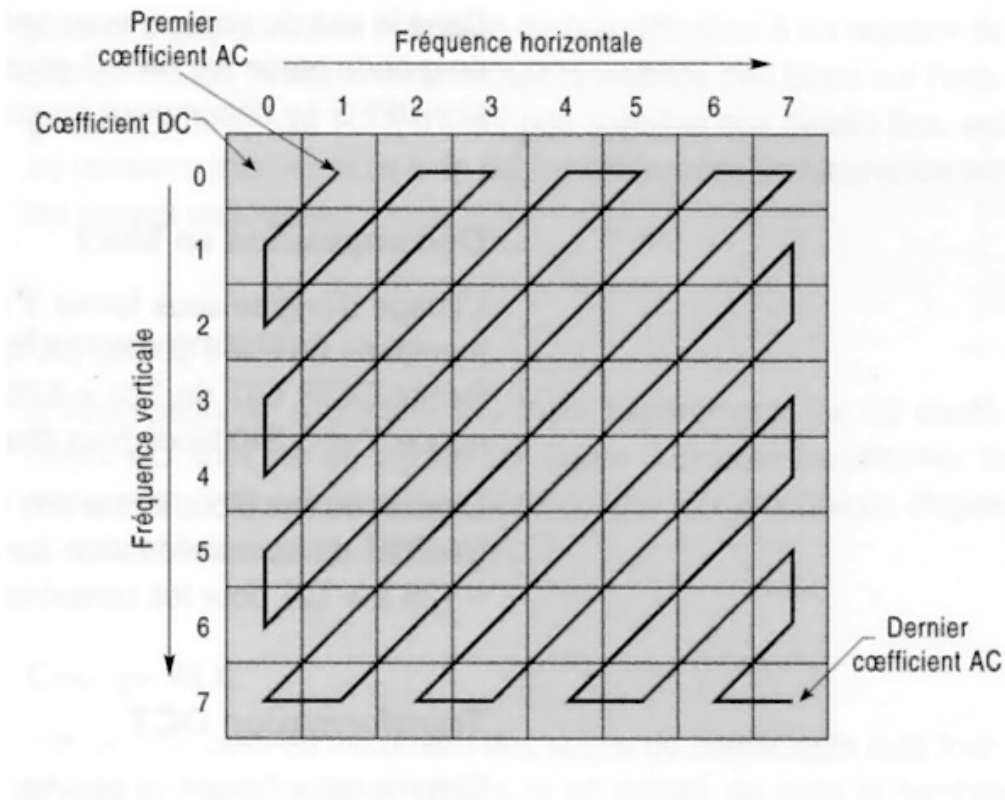


Dans l'exemple ci-dessus, après la DCT, tout les coefficients inférieurs à deux ont été supprimés, ensuite, les coefficients sont arrondis à des multiples de trois (les valeurs 2 et 3 sont des exemples plus ils sont élevés, plus l'image sera détériorée).

Cet exemple illustre une façon de compresser en JPEG, il existe de nombreux codecs utilisant des algorithmes différents.

d/ Balayage en zigzag de la matrice.

La matrice obtenue après quantification est balayée en zigzag, de sorte que les coefficients soient réarrangés sous forme d'un vecteur.



e/ Codage à longueur variable .

Le vecteur formé contient des suites plus ou moins longues de coefficients de valeurs identiques, parfois nulles.

Plutôt que de les coder systématiquement, le codage à longueur variable (RLE) formate le vecteur sous la forme de paires de données.

La première indique la valeur du coefficient, la seconde le nombre de fois ou il est répété.

f/ Codage entropique.

Le principe du codage entropique est d'attribuer le plus court au coefficients les plus fréquents, et le code le plus long aux coefficients statistiquement les moins fréquents.

Ce codage est réalisé après une analyse statistique de l'image compressée.

NB : En image fixe, on code ensuite sur le nombre de bits nécessaires par images, en MJPEG, on code sur le nombre de bits nécessaires à la séquence.

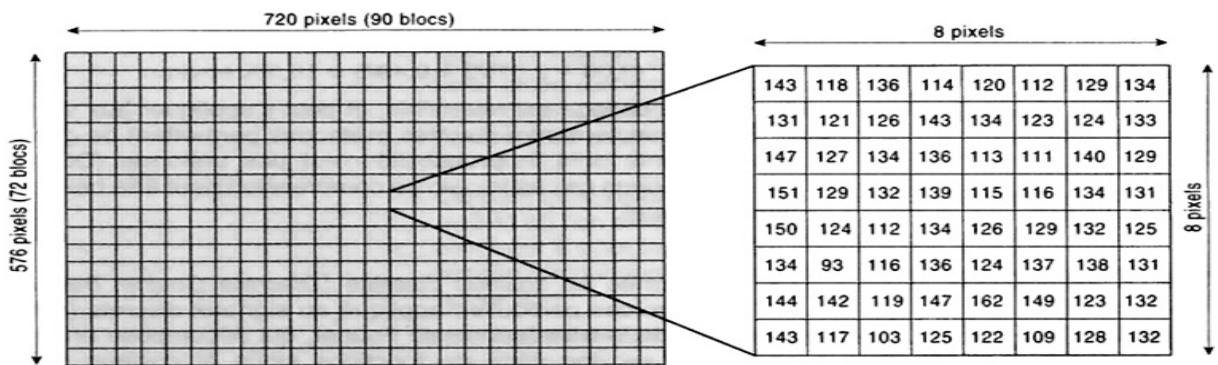
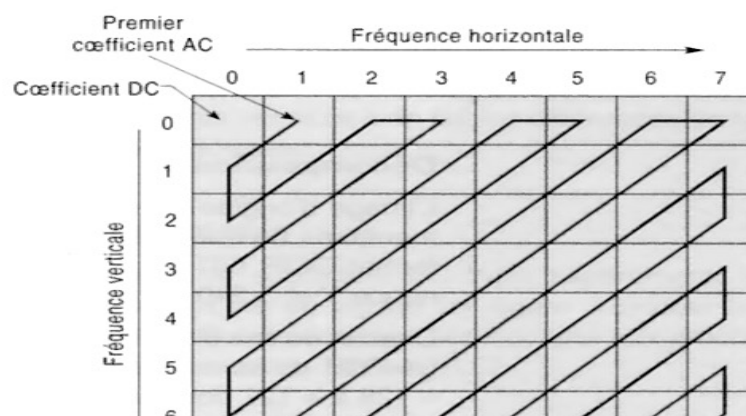
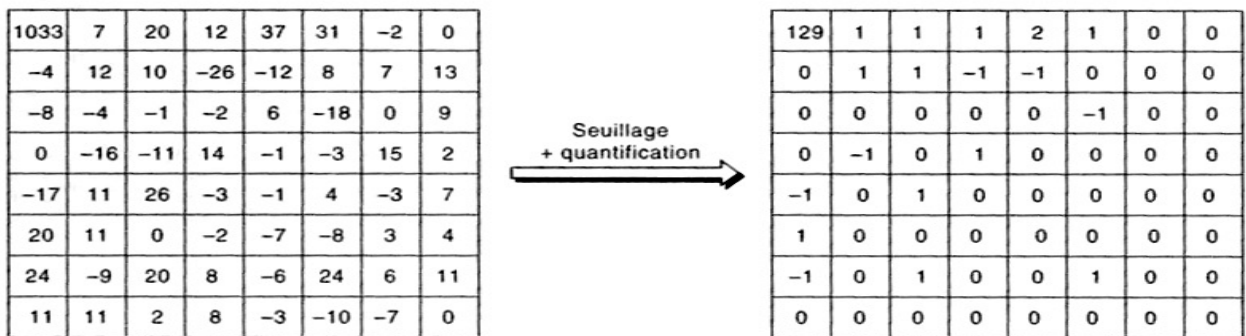
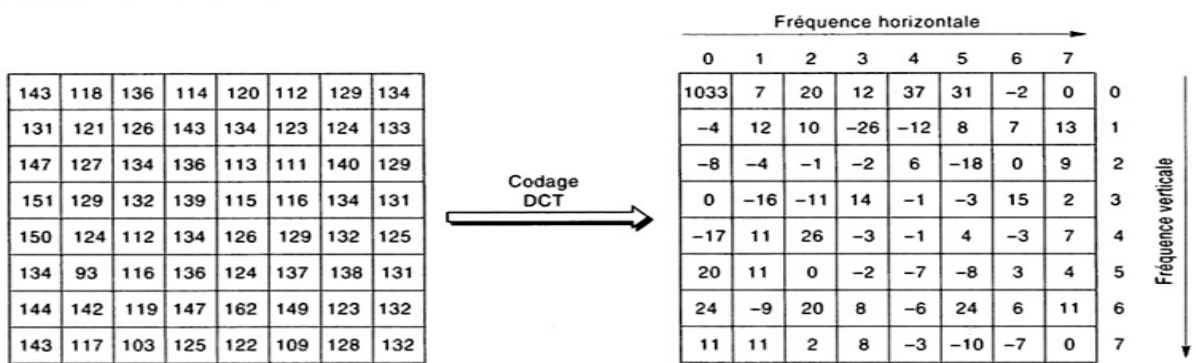


Figure 3.2 : Découpage en blocs de 8 × 8 pixels (les valeurs représentent l'intensité lumineuse d'un pixel).



exemples :



image JPEG compressée à 1/10 (division par 10)



la même image compressée à 1/100

On remarque dans l'exemple ci dessus qu'une compression trop élevée touche aussi les basses fréquences et détériore l'image.

2/ La compression temporelle ou MPEG

Dans la compression spatiale en vidéo (MJPEG), chaque image était compressée, indépendamment les unes des autres. Chaque image contient donc l'entièreté des informations qui lui sont nécessaires.

La compression temporelle (MPEG) tient compte des redondances entre images.

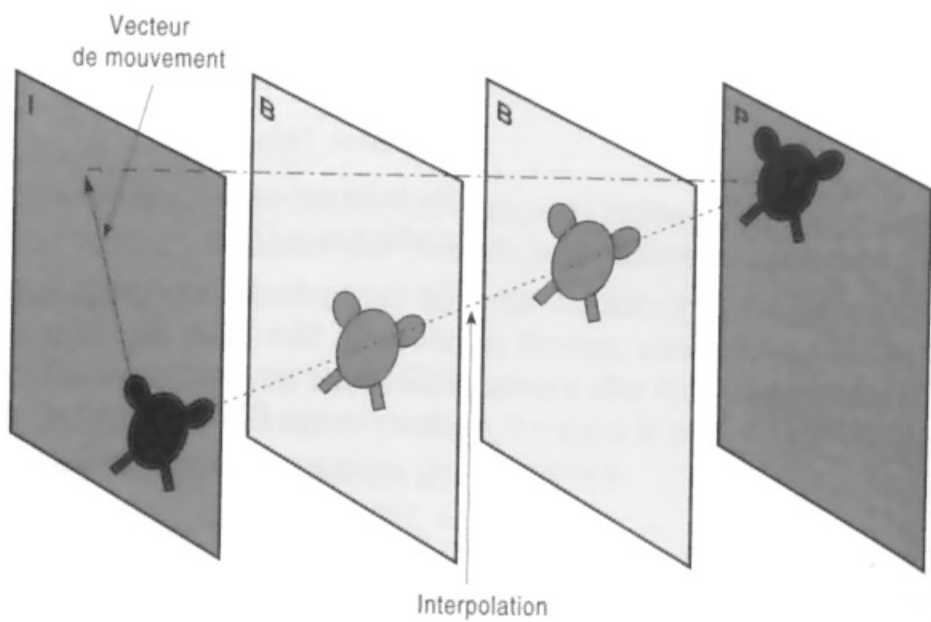
Une séquence de type MPEG se compose de trois types d'images :

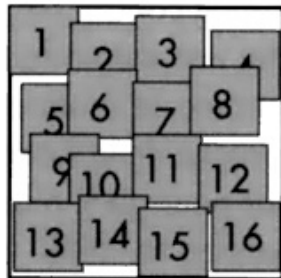
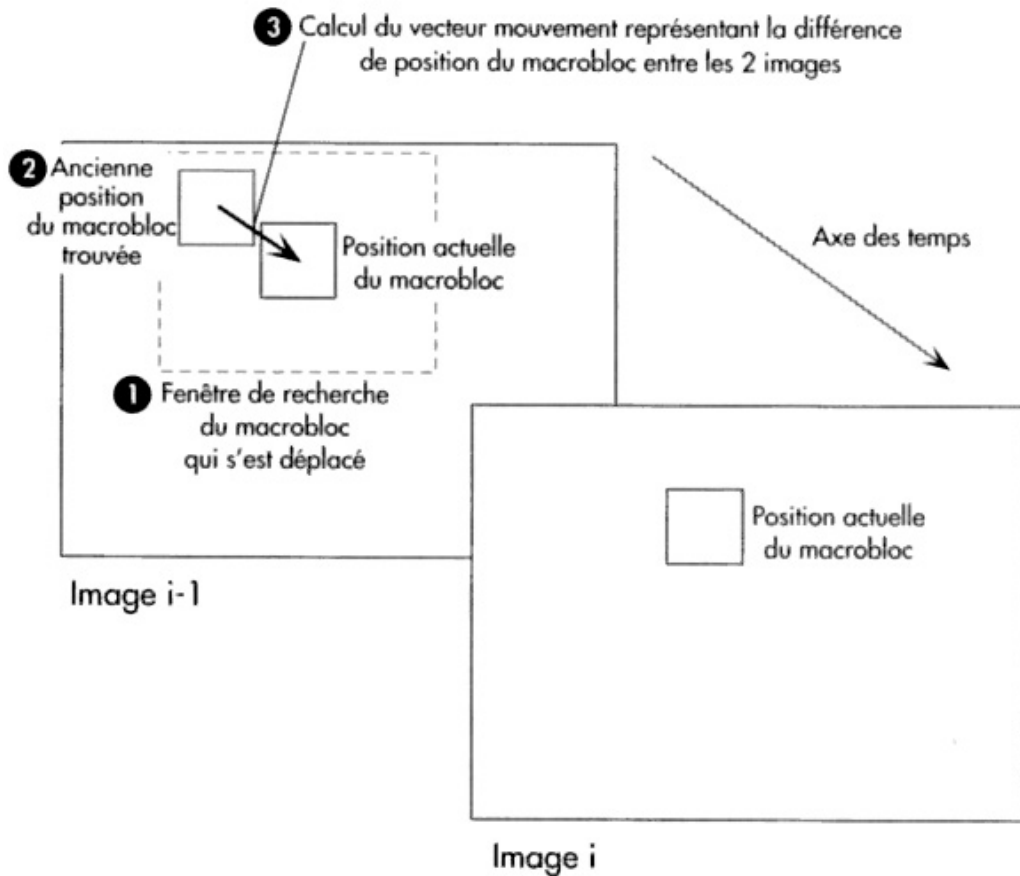
Les images I (intra)

Une image I est codée en mode intra-image en JPEG, elle contient toutes les informations nécessaire à son décodage.

Les images P (prédites)

Comme leur nom l'indique, les image P sont prédites à partir des images I, à l'aide des vecteurs de mouvements, c'est-à-dire le calcul du déplacement des éléments de l'image. Une image P est typiquement trois fois moins lourde qu'une image I.





Macroblocs de l'image précédente i-1 utilisés pour prévoir la nouvelle image i

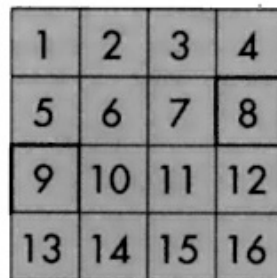


Image i prédite, après application des vecteurs mouvement sur les macroblocs de l'image i-1

Les images B (bidirectionnelles)

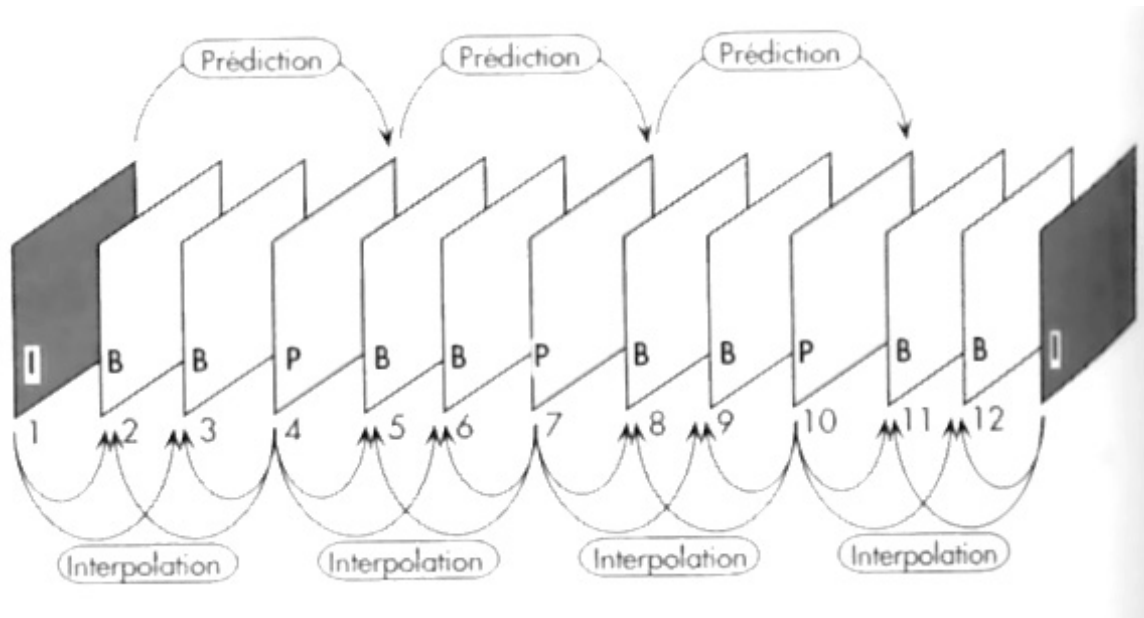
Elles sont calculées bi-directionnellement à partir des images I et des images P à l'aide des vecteurs de mouvement.

Une image B est typiquement trois fois moins lourde qu'une image P

Une séquence MPEG composée des trois types d'images précitées s'appelle un **GOP** (group of pictures).

Un gop commence par une image I et se termine par l'image précédant l'image I suivante.

Un GOP peut se résumer à une seule image I, d'une séquence IP, IB, ou IPB.



Dans l'exemple ci-dessus, nous avons un GOP à trois types d'image, IP et B.

Dans un GOP, nous avons deux paramètres : M étant l'intervalle entre deux images P (ou entre I et P) et N, qui est l'intervalle entre deux images I (égale à la longueur du GOP).

Dans cet exemple, nous avons un M=3 et un N=12.

Le principal inconvénient du système MPEG est que l'on n'a pas accès à chaque image de la séquence, en effet une image B dépendant en partie de données situées après elle, on ne pourra pas couper à cet endroit en montage.

Du fait de cet inconvénient, le MPEG est surtout un format de diffusion (DVD, VCD, Bouquets numériques etc....).

En diffusion, le GOP est de 12, en DVD de 15, en Betacam SX (acquisition broadcast) de 2.

Il est évident que plus le GOP est long, plus la compression est importante et plus les vecteurs de mouvement devront être précis.

Après calcul des images intermédiaires à l'aides des vecteurs de mouvement, elles sont comparées aux images réelles pour générer des données d'erreur de prédiction. Ces erreurs sont ajoutées aux vecteurs de mouvement à l'encodage.

Jusqu'ici, nous avons vu le principe général de MPEG, nous allons maintenant nous intéresser aux différents MPEG.

1/ MPEG 1

Apparue en 1992 la norme MPEG1 était destinée à compresser de la vidéo en faible résolution.

Le but était de pouvoir stocker 74 min de vidéo (audio comprise) sur un CDROM (650Mbytes à l'époque), seul support disponible en 1992.

Le résultat ne devait pas dépasser en qualité le VHS.

Pour arriver à cela, le débit vidéo ne doit pas dépasser 1,15Mbits/sec (rappel :21Mbytes/sec en 4 :2 :2 non compressé).

Avant même la compression, MPEG1 part d'une image de format SIF (Source Intermediate Format) caractérisée par une réduction de moitié des résolutions horizontales et verticale.

En clair, il utilise des images dans lesquelles une trame sur deux est supprimée (la restante est ensuite doublée à la lecture) et un point sur deux est supprimé à chaque ligne, en luminance comme en chrominance.

Nous avons donc finalement une image faisant 360 pixels par 288 lignes, avec une fréquence trame de 25 Hz.

Le débit SIF avant compression est de 31,5 Mbits/sec.

Ce format pourrait être appelé du 2 :1 :0....

En ce qui concerne la compression MPEG proprement dite, MPEG1 reprend la théorie vue précédemment, avec un M=3 et un N=12.

2/ MPEG 2

MPEG 2 fut normalisé en 1994, il s'agit ici d'une famille de standards qui reprend les caractéristiques de base de MPEG1, mais en les adaptant à toutes les exigences de l'industrie audiovisuelle.

En effet, MPEG2 est indépendant des applications et des supports, c'est une « boîte à outils » concernant aussi bien la haute définition, le broadcast, l'amateur et l'internet.

Comparaison avec MPEG1 :

- image d'entrée jusqu'à la haute définition
- traitement du balayage entrelacé (50 trames/sec) et du balayage progressif (50 images/sec)
- codage hiérarchique permettant de transmettre différents niveaux de qualités pour une même image, avec compatibilité descendante entre les niveaux.
- compatibilité avec MPEG1 (un lecteur DVD peut lire le format VCD qui est du MPEG1).

Profils et niveaux de MPEG2

Je comparais plus haut MPEG2 à une boîte à outils, ces outils sont en fait des profils et des niveaux.

Un profil correspond à une approche quantitative, il définit le jeu d'outils de compression utilisé.

Un niveau détermine le nombre de lignes et de pixels qu'il est possible de coder (résolution).

Les profils MPEG2

- Le profil simple : il traite l'image 4 :2 :0 et n'utilise que des image I et P
- Le profil principal (main profile ou MP) : c'est le profil de base en diffusion grand public (DVD, bouquets numériques). Il reprend les outils du profil simple avec son traitement 4 :2 :0, mais il utilise l'interpolation bi directionnelle (images B).
- Le profil d'échelonnabilité SNR (Signal Noise Ratio) : il possède tout les outils du profil principal, mais il est capable de séparer les données en deux parties.
Une partie pour le signal de couche de base qui contient les données relatives à une qualité d'image donnée.
Une seconde partie, signal de couche supérieure, qui ajouté à la couche de base améliore rapport signal/bruit.
Exemple : une image tv MPEG2 peut être transmise sous la forme image MPEG1+info supplémentaires pour en faire une MPEG2.
- Le profil d'échelonnabilité spatiale : il possède tous les outils du profil précédent, plus un, constituant une nouvelle méthode de division des données en fonction de la résolution de l'image.
Il consiste toujours en un signal de couche de base et un signal de couche supérieure, mais ce dernier améliore ici la résolution.
Exemple : TVHD=TVstandard+info.
- Le profil élevé (high profile) : dispose de tous les outils du précédent, mais il peut coder une image 4 :2 :2 .
- Le profil 4 :2 :2 : reprend les outils du profil principal, mais travaille en 4 :2 :2. Il ne rentre pas dans la catégorie des profils échelonnables.

Les niveaux MPEG2

- Le bas niveau (low level) : il n'accepte que le format SIF (MPEG1)
- Le niveau principal (main level) : c'est le niveau de base en 4 :2 :2
- Le haut niveau 1440 : c'est le niveau correspondant à une image haute définition à 1440 points par ligne.
- Le haut niveau (high level) : niveau correspondant à une image HD , mais à 1920 points par ligne.

Vu ce qui précède, 24 combinaisons profil/niveau sont possibles, en fait, actuellement, 2 combinaisons couvrent toutes les applications courantes.

La première est le Main profile @ Main level noté MP @ ML , c'est la norme qui prévaut dans le domaine grand public comme le DVD ou les bouquets de transmission numériques.

La seconde est le 4 :2 :2 Profile @ Main level ou 4 :2 :2 P @ ML, utilisée en professionnel avec un débit maximal de 50 Mbits/sec.

Une troisième est prévue pour la HD, 4 :2 :2 P @ HL, avec un débit de 300Mbits/sec.

rofiles^P	Simple	Main	4 : 2 : 2	SNR	Spatial	HIGH
Niveaux						
High		4 : 2 : 0 1920x1152 80Mbits/sec IPB	4 : 2 : 0-4 : 2 :2 1920x1152 300Mbits/sec IPB			4 : 2 : 0-4 : 2 :2 1920x1152 100Mbits/sec IPB
High 1440		4 : 2 : 0 144x1152 60Mbits/sec IPB			4 : 2 : 0 1440x1152 60Mbits/sec IPB	4 : 2 : 0-4 : 2 :2 1440x1152 80Mbits/sec IPB
Main	4 : 2 : 0 720x576 15Mbits/sec IP	4 : 2 : 0 720x576 15Mbits/sec IPB	4 : 2 : 0-4 : 2 :2 720x576 50Mbits/sec IPB	4 : 2 : 0 720x576 15Mbits/sec IPB		4 : 2 : 0-4 : 2 :2 720x576 20Mbits/sec IPB
Low		4 : 2 : 0 352x288 4Mbits/sec IPB		4 : 2 : 0 352x288 4Mbits/sec IPB		

La technologie « the mole » (la taupe)

Une suite de compression MPEG, décompression, recompression etc.. n'est pas transparente, en effet, à chaque recompression les paramètres utilisés dépendent du codeur et des M et N utilisés. Ce genre de traitement détruit rapidement le signal.

Pour éviter cela, à la première compression, les paramètres utilisés sont stockés sur un fichier, à la compression suivante, il suffit de reprendre les paramètres dans ce fichier pour que la compression soit transparente.

3/ MPEG 4

Le MPEG 1 était destiné aux CDROMS, MPEG 2 à la télévision, MPEG 4 est lui destiné à l'informatique et aux télécommunications.

Les différents types de compression vus jusqu'à présent considéraient les images comme une succession de pixels, ce qui fait que le degré d'interactivité était minime.

MPEG 4 est au contraire une compression orientée objet, une image est composée d'un fond, d'un logo, d'un sous titre, d'un personnage etc., au-delà de cela, MPEG 4 peut compresser une image traditionnellement, mais avec beaucoup plus d'efficacité, parce qu'utilisant des techniques plus récentes, il suffit de comparer le poids et la qualité d'une séquence en DIVX (dérivé du MPEG 4) avec la même séquence en DVD (MPEG 2).

MPEG4 détecte donc les différents objets compris dans une image, et les soumet à la compression la mieux adaptée. Un sous titre ne nécessite pas le même débit qu'un personnage principal. Ces différents objets peuvent en outre être transmis sur des canaux différents, certains d'entre eux peuvent aussi être enlevés (chaque chaîne TV peut ainsi apposer son propre logo).

MPEG 4 introduit aussi la notion de sprites, qui sont des objets de taille supérieure à la scène filmée (arrière plan, décor statique). Ces sprites sont transmis une fois pour toutes en début de séquence, et son ensuite manipulés par des mouvements de zoom, recadrages, panoramiques etc....

Par ailleurs, MPEG 4 répond au principe d'échelonnabilité, ce qui élimine le problème de devoir coder le signal différemment pour chaque bande passante. Le flux de données est séparé en une couche de base qui peut être reçue par tout les décodeurs, et plusieurs couches secondaires, chacune améliorant la résolution spatiale ou temporelle de l'image. Par exemple, la couche de base serait destinée aux téléphones portables, et les couches secondaires fourniraient les informations nécessaires à la réception optimale sur un téléviseur.

Les profils de MPEG-4

- Le profil simple : Il utilise un schéma de compression emprunté à MPEG-2 et ne vise que des applications à très bas débits, du genre téléphonie mobile ou agendas électroniques (débit de 32 à 384 Kbits/sec).

Il ne code que des images I et P, ses vecteurs de mouvement opèrent sur des macroblocs de 16x16 pixels.

- L'advanced simple profile : peut traiter des images I P et B avec un débit maximal de 8 Mbits /sec . C'est le profil utilisé en vidéo standard (et en DIVX).

- Le studio profile : vise les applications broadcast, de la résolution standard à la haute définition. Il supporte le 4 :2 :2 et le 4 :4 :4 , son débit peut varier de 50 Mbits/sec à 1,2 Gbits/sec, il est adapté à la haute définition et au cinéma numérique.

MPEG-4 se décompose en une suite de normes, les parties, qui spécifient un type de codage particulier. Dans chaque partie plusieurs profils (collection d'algorithmes) et niveaux (contraintes quantitatives) sont définis. Un consortium industriel désirant utiliser MPEG-4 choisit une ou plusieurs parties de la norme et, pour chaque partie, il peut sélectionner un ou plusieurs profils et niveaux correspondant à ses besoins.

A titre d'information, les différentes parties de MPEG-4 sont nommées ci-après :

- La Partie 1 décrit la synchronisation et le multiplexage de la vidéo et de l'audio.
- La Partie 2 est un codec de compression pour les signaux vidéo. L'un des nombreux profils de ce codec est l'ASP (*Advanced Simple Profile*).
- La Partie 3 est une norme de compression pour le codage perceptuel et les signaux audio; elle spécifie notamment le format audio AAC.

- La Partie 4 décrit les procédures pour les tests de conformité.
- La Partie 5 fournit des logiciels de référence des autres parties de la norme.
- La Partie 6 décrit le *Delivery Multimedia Integration Framework* (DMIF).
- La Partie 7 fournit des implémentations optimisées (cf. partie 5)
- La Partie 8 décrit les méthodes de transport du MPEG-4 sur IP.
- La Partie 9 fournit des implémentations matérielles des autres parties à titre d'illustration.
- La Partie 10 est une norme avancée de compression vidéo appelée aussi H.264 ou AVC (*Advanced Video Codec*), et qui comporte une extension appelée SVC (*Scalable Video Coding*).
- La Partie 11 spécifie la description de scène et moteur d'application.
- La Partie 12 spécifie le format de fichier ISO Base media.
- La Partie 13 fournit les extensions de gestion et de protection de la propriété intellectuelle (IPMP).
- La Partie 14 spécifie le format de fichier MP4.
- La Partie 15 spécifie le format de fichier du codec AVC (partie 10), sur base de la partie 12.
- La Partie 16 fournit l'extension du cadre d'animation (AFX).
- La Partie 17 spécifie le format de sous-titrage *Timed Text*.
- La Partie 18 spécifie la compression et transmission de polices de caractères.
- La Partie 19 décrit le flux de texture synthétisé.
- La Partie 20 spécifie la représentation "allégée" de description de scène (pour applications mobiles).
- La Partie 21 spécifie MPEG-J GFX.
- La Partie 22 spécifie le format *Open Font*, basé sur OpenType.

La partie 10 est très intéressante et pleine de promesses...

La norme est habituellement appelée H.264/AVC (ou AVC/H.264 , H.264/MPEG-4 AVC ou MPEG-4/H.264 AVC)

Le codec H.264/AVC est adapté à une très grande variété de réseaux et de systèmes (par exemple, pour la diffusion de la télévision, le stockage HD DVD et Blu-ray, le streaming RTP/IP , et des systèmes de téléphonie).

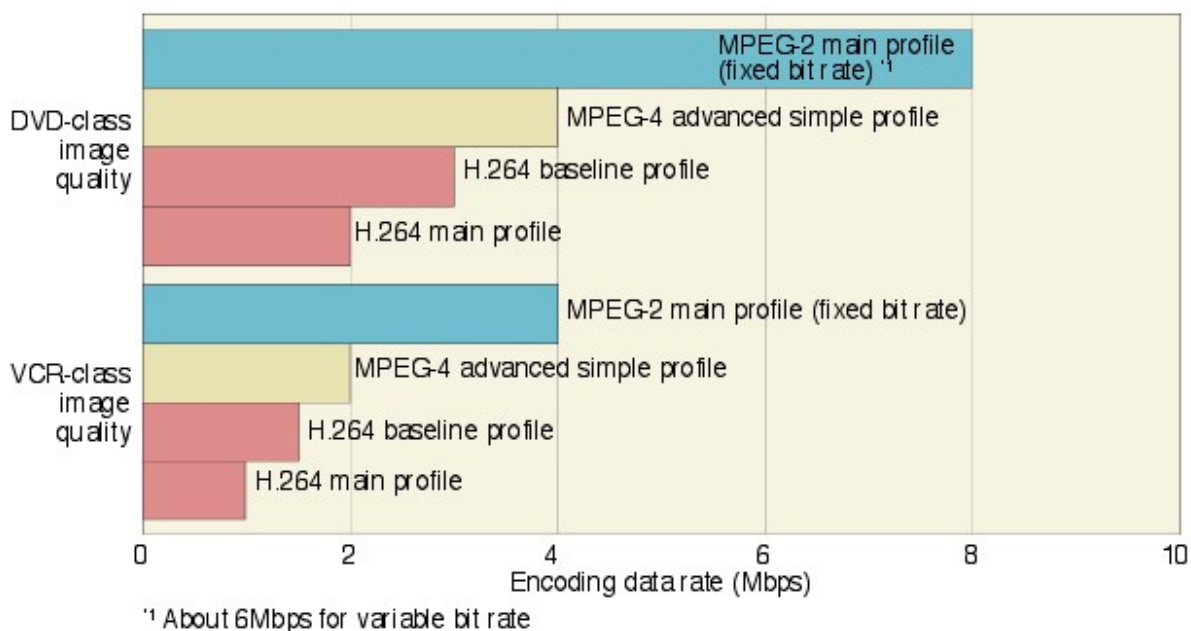
H.264/AVC(MPEG-4 Part 10) comprend de nombreuses techniques nouvelles qui lui permettent de compresser beaucoup plus efficacement les vidéos que les normes précédentes (H.261, MPEG-1, MPEG-2, MPEG-4 Part 2/ASP) et fournit plus de flexibilité aux applications dans un grand nombre d'environnements réseau. Dans ces fonctionnalités principales sont inclus :

- Une compensation de mouvement pouvant être effectuée par rapport à plusieurs images de référence déjà codées. Le choix de l'image de référence intervient au niveau macroblock . Ceci permet d'utiliser dans certains cas jusqu'à 32 images de référence (contrairement aux précédentes normes, qui étaient limitées à une ou dans le cas d'*images B* conventionnelles, à deux) et jusqu'à 4 références différentes pour un même macroblock. Cette fonctionnalité particulière permet habituellement des améliorations modestes au niveau du débit et de la qualité dans la plupart des scènes. Mais dans certains types de scènes, comme par exemple les scènes contenant des flashes rapides et répétitifs ou des scènes réapparaissant fréquemment, il permet une réduction du débit réellement significative.
- Une compensation de mouvement pouvant utiliser 7 tailles de blocs différentes (16×16, 16×8, 8×16, 8×8, 8×4, 4×8 4×4) permet une segmentation très précise de zones se déplaçant.
- Une précision au quart de pixel pour la compensation de mouvement, permettant une description très précise du déplacement des zones en mouvement. Pour la

chrominance, la précision de la compensation de mouvement se fait même au huitième de pixel.

- Un filtrage anti-blocs ((en)deblocking filter), effectué dans la boucle de codage et opéré sur les blocs 4×4, permettant de réduire les artefacts caractéristiques du codage avec transformation en bloc.
- Une transformée entière effectuée sur des blocs de taille 4×4 pixels (proche de la DCT classique). Pour les nouveaux profils issus des extensions FExt, une transformée supplémentaire de taille 8×8 a été ajoutée.
- Une prédiction spatiale sur le bord des blocs voisins pour un codage « intra » (plutôt que la seule prédiction sur les coefficients continus présente dans MPEG-2)
- Un codage arithmétique ((en)CABAC: Context-adaptive binary arithmetic coding), qui est une technique sophistiquée de codage entropique qui produit d'excellents résultats en termes de compression mais possède une grande complexité (non disponible dans les profils *baseline* et *extended*).
- Un processus automatisé simple de prévention contre la création accidentelle de faux codes de démarrage. Il s'agit de séquences binaires spéciales qui sont placées au sein des données, permettant un accès aléatoire au flux de données ainsi qu'une resynchronisation en cas de perte temporaire du flux.
- La numérotation des images permet la création de sous-séquences (permettant une scalabilité temporelle par l'inclusion optionnelle d'images supplémentaires entre d'autres images) ainsi que la détection et la dissimulation de la perte d'images entières (qui peuvent se produire en cas de perte de paquets réseau ou d'erreurs de transmission).
- Le comptage de l'ordre des images permet de conserver l'ordre des images et du son dans des images décodées isolément des informations de minutage (ce qui permet à un système de transporter, contrôler et/ou changer l'information de minutage sans affecter le contenu des images).

Ces techniques, ainsi que plusieurs autres, aident H.264 à dépasser significativement les standards précédents, dans une grande variété de circonstances et dans une grande variété d'environnements d'application. H.264 peut fonctionner souvent nettement mieux que la vidéo MPEG-2 en obtenant la même qualité avec un bitrate diminué de moitié, voire plus.



Les applications de MPEG-4

Le simple profile est utilisé en streaming, il est supporté par real player, quick time...

L'advanced simple profile par le système DIVX.

Et le studio profile par un nouveau système haute définition développé par SONY, le HDCAM SR, conçu pour le cinéma numérique.

Comparaison des divers systèmes de compression utilisés en acquisition.

M-JPEG

Le M-JPEG, notamment utilisé par le Betacam digital, est une famille de systèmes propriétaires, chaque fabricant a développé le sien, tous incompatibles entre eux. Seul point commun, les images sont compressées indépendamment les unes des autres à 25 images/sec, c'est donc une compression intra-image.

Une compression M-JPEG est jugée sans perte jusqu'à un taux de 3 :1.

DV

Le DV est à l'origine un M-JPEG, mais avec deux améliorations appréciables. Tout d'abord, DV repose sur un algorithme normalisé, tous les DV de différents constructeurs sont compatibles entre eux. Deuxièmement, son codage est optimisé, ses tables de quantification sont calculées par zones. Ce qui lui permet avec une compression 5 :1 d'avoir la même qualité qu'un M-JPEG à 3 :1.

DV fut à l'origine développé pour les systèmes amateurs, mais par la suite, son algorithme fut utilisé pour des applications industrielles et broadcast (DVCAM, DVCPRO, DVCPRO25 etc..). Mais à partir de ce moment, adieu la compatibilité...

MPEG-2 422

Utilisé par un seul système, le Betacam SX, MPEG2-422 est le seul type de compression utilisant les redondances inter-image.

Il n'utilise que des images I et B, le problème du montage a été résolu avec des magnétoscopes hybrides utilisant un disque dur pour décompresser les images B si elles se trouvent au point d'insertion.

MPEG-IMX est aussi un système utilisant MPEG2-422, mais en intra-image.

	Signal vidéo	Algorithme de Compression	Taux de Compression	Débit vidéo en Mbits/sec
Digital Beta	4 :2 :2 10 bits	M-JPEG	2 :1	99
MPEG-IMX	4 :2 :2 8 bits	MPEG2 422 Intra image	3,3 :1	50
D9	4 :2 :2 8 bits	DV	3,3 :1	50
DVCPRO50	4 :2 :2 8 bits	DV	3,3 :1	50
Beta SX	4 :2 :2 8 bits	MPEG2 Inter-image	10 :1	18
DVCPRO25	4 :1 :1 8 bits	DV	5 :1	25

DV, DVCAM	4 :2 :0 (625/50) 4 :1 :1 (525/60) 8bits	DV	5 :1	25
MICROMV	4 :2 :0 8bits	MPEG2 Inter-image	10 :1	12

Le MICROMV développé par SONY est un système numérique amateur très compact. Il se base sur une compression MP@ML avec un ratio de 10 :1, il a un GOP de 12 images, ce qui réduit considérablement la précision de montage (0,5 sec).

Ce type de compression lui permet un débit vidéo de 12Mbits/sec.

C'est aussi le seul système dans lequel l'audio est compressé (6 :1).

L'audio est enregistré avec la vidéo sur une piste hélicoïdale.

Toutes ces réductions permettent d'utiliser des cassettes de taille 30% inférieures aux cassettes DV.

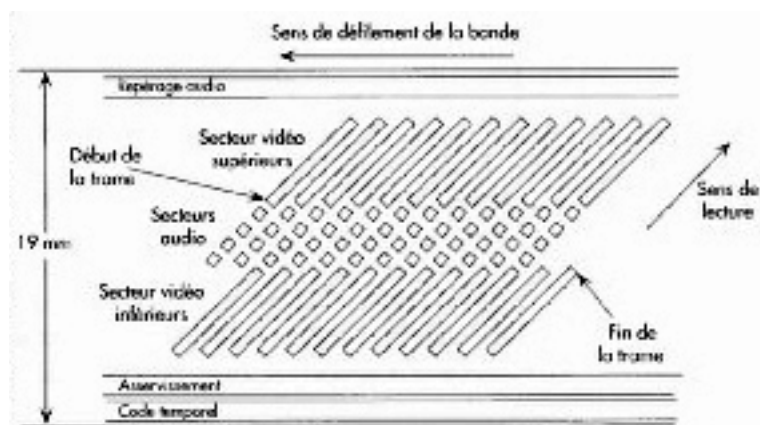
Les formats d'enregistrement numériques.

Le D1

- 4 :2 :2 sur 8 bits
- Non compressé
- 4 canaux audio à 48Khz sur 20 bits
- Débit vidéo : 172 Mbits/sec
- Débit audio : 8 Mbits/sec
- données : 47 Mbits/sec
- Débit total : 227 Mbits/sec

En conclusion, nous avons par trame un débit de 4,5 Mbits/sec (227/50), ce qui demanderait dans un système 1 piste = une trame (comme en analogique) des piste d'une longueur de 2m avec un tambour de 70 cm de diamètre.

D'où nécessité de segmenter les pistes.



1 trame = 12 pistes = 24 secteurs

les 4 pistes audio numériques sont enregistrées entre les secteurs vidéo.

50 lignes correspondent donc à 4 secteurs enregistrés par 4 têtes différentes.

- vitesse de rotation du tambour 150 Hz
- pistes sans azimut, nécessité d'une piste de garde de 5 micromètres (piste de 40 micromètres)

- redondance de 100% du son

Le D2

- système numérique composite non compressé, n'est plus utilisé.

Le D3

- Comme le précédent, mais sur une bande d'un demi pouce (les autres nécessitent une bande $\frac{3}{4}$ de pouce

Le **D4** n'existe pas, le 4 portant malheur dans la culture nipponne...

Le D5

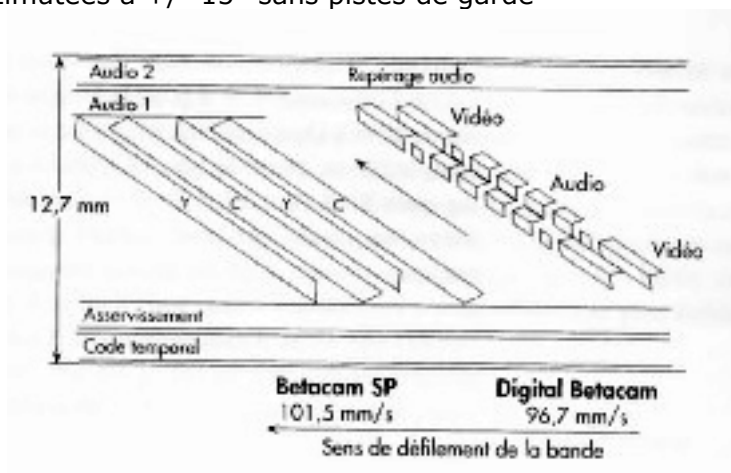
- 4 :2 :2 sur 10 bits
- Bandes $\frac{1}{2}$ pouce
- peut enregistrer du 16/9 avec un échantillonnage à 18 MHz (mais sur 8 bits).
- débit total 288 Mbits/sec
- débit vidéo 218 Mbits/sec en 4/3, 234 Mbits/sec en 16/9
- n'existe qu'en magnéscope de studio (pas de cam)

Le betacam digital

- 4 :2 :2 sur 10 Bits
- compressé avec un algorithme MJPEG propriétaire à 2 :1
- 4 canaux audio numériques à 48 Khz sur 20 bits

structure d'enregistrement :

- une trame = 6 pistes hélicoïdales
- rotation du tambour à 75 Hz
- Pistes azimuthées à +/- 15° sans pistes de garde



Ces magnétoscopes sont compatibles en lecture avec le betacam SP, ce qui nécessite pas moins de 18 têtes sur le tambour, 4 pour l'enregistrement, 4 de lecture avec DT, 4 de mode confiance, 2 d'effacement, 4 de lecture analogique avec DT.

Le DV

Le DV est un système numérique grand public développé par un consortium composé de Matsushita, Philips, Sony et Thomson.

- 4 : 1 : 1 en 525/60
- 4 : 2 : 0 en 625/50
- 8 bits
- algorithme de compression de type MJPEG à 5 : 1
- deux canaux audio à 48 KHz sur 16 bits ou quatre canaux audio à 32 KHz sur 8 bits
- débit vidéo avant compression 124 Mbits/sec
- débit vidéo après compression 25 Mbits/sec
- débit total : 41,85 Mbits/sec

D'où vient le DV ?

La recherche sur la vidéo numérique et le développement de produits numériques ont commencé il y a maintenant de nombreuses années. A la fin des années 80 ils étaient en grande progression. « Juillet 93 : Matsushita, Philips, Sony et Thomson annoncent qu'ils adopteront un cahier des charges commun pour développer le premier format d'enregistrement vidéo numérique destiné au grand public. » Ils sont directement rejoint par Hitachi, JVC, Mitsubishi, Sanyo, Sharp et Toshiba. Cela allait mettre fin à la guerre des formats que l'on connut avec le S-VHS, le Hi-8 et le Betamax. Ce consortium historique dans le monde de l'audiovisuel réunit directement une cinquantaine d'autres compagnies active dans ce secteur et mena à l'introduction du premier DVC (Digital Video Cassette) destiné au consommateur.

Initialement appelé DVC il devint finalement DV pour « Digital Video. » Le DV offre une qualité d'image assez remarquable, puisque globalement comparable à celle du Betacam SP, la référence broadcast durant de nombreuses années.

Alors que la vidéo numérique débuta tout d'abord comme technologique réservée au professionnel et qu'elle continue de gagner en importance dans ce secteur, la seconde partie des années 90 a vu son extension chez le « prosumer »¹ et le « consumer » grâce à l'apparition du nouveau format numérique DV.

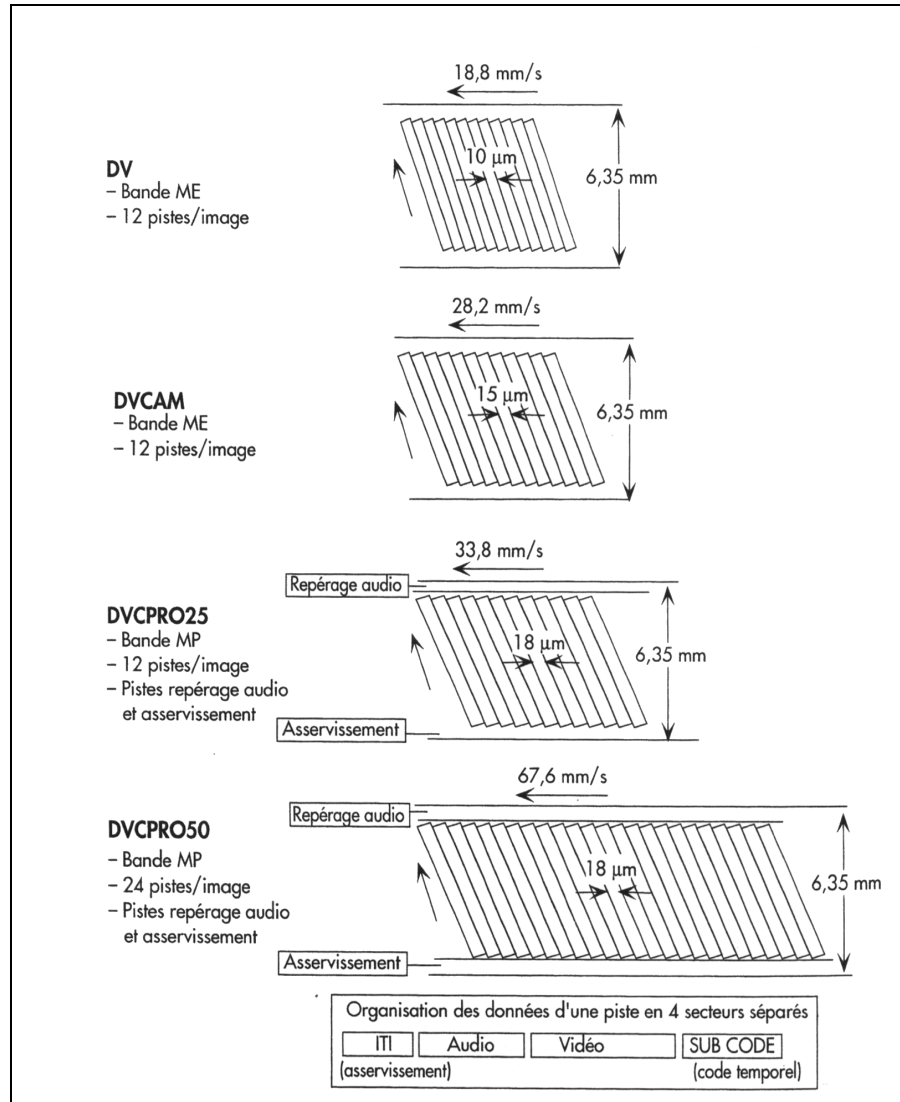
Des variantes « professionnelles » du format DV furent inventées. Et finalement on eut droit à notre guerre des formats. DVCPRO fut lancé par Panasonic en 1995, suivit un an plus tard par la DVCAM de Sony. Ces deux formats possèdent le même type d'encodage vidéo et audio que le DV standard destiné au consommateur, la différence se joue dans la largeur des pistes, la vitesse de bande ainsi que des composants plus robustes et plus coûteux tels les objectifs ou encore les CCD utilisés.

¹

L'algorithme MJPEG standard à été optimisé dans le système DV, en effet celui-ci inclus une détection de mouvement entre deux trames d'une même image.

- si il y a peu de mouvement, la compression est réalisée sur l'ensemble de l'image sur des blocs de 8X8 pixel en exploitant la redondance inter-trame, ce qui permet une meilleure efficacité de la compression.

- si la différence est trop importante entre les deux trames, la compression s'effectue par trame sur des blocs de 8X4 pixels.



- largeur de piste 10 micromètres
- bande de type métal évaporé
- 12 pistes par image en 625, 10 pistes en 525
- pas de piste longitudinale, tout est enregistré sur les pistes hélicoïdales (TC, asservissement, repérage audio...)
- pistes azimuthées à +/- 20°
- une piste = 4 secteur
- tambour à deux têtes, rotation à 150 Hz
- vitesse de défilement de bande = 2cm/sec

L'interface IEEE1394

Cette interface a été développée par Sony et Apple, elle permet de connecter les équipements DV entre eux, de connecter des équipements DV à des ordinateurs ou des périphériques à des ordinateurs (jusqu'à 63 périphériques).

Il s'agit d'une interface sérieuse comprenant 3 paires de conducteurs, une paire pour les informations, une paire pour les signaux de commande et une paire pour l'alimentation. Il faut noter qu'il existe des interfaces de ce type ne comprenant pas la paire alimentation, cela se traduit par un connecteur plus petit.

- le débit d'une IEEE1394 est de 400 Mbits/sec sur 4,5m
- le débit d'une IEEE1394b est de 800Mbits/sec sur 100 m

Elle porte le nom de Firewire chez Apple et i.Link chez Sony.
IEEE signifie : Institute of Electrical and Electronic Engineers.

DVCAM

Version industrielle du DV chez Sony (4:2:0 et 4:1:1)

- La vitesse de la bande est multipliée par 1,5 par rapport au DV
- Largeur de piste : 15 micromètre

Le système est donc plus robuste que le DV, plus adapté au tournage industriel.

DVCPRO25 (ou D7)

Version industrielle du DV chez Panasonic

- uniquement en 4 :1 :1
- utilisation d'une piste d'asservissement longitudinale
- utilisation d'une piste de repérage audio longitudinale
- largeur de piste 18 micromètres
- utilisation de bande « métal particule » (plus résistante aux drops)

DVCPRO50

Destiné à la production haut de gamme, celui-ci fonctionne en 4 :2 :2.

- compression à 3,3 :1
- débit vidéo : 50 Mbits/sec (d'où son nom)
- largeur de piste : 18 micromètres
- une image est segmentée en 24 pistes

BETACAM SX

Le Betacam SX est un système d'acquisition en 4 :2 :2 se basant sur une compression MPEG2-4 :2 :2 (MPEG2 4 :2 :2P@ML)

- Le codage se fait sur des groupes de deux images IBIBIBI
- par ce procédé, le débit est égal à 70% d'une compression intra-image à qualité égale
- pour pouvoir monter à l'image, les B sont reconstituées avant l'editing quand cela s'avère nécessaire (à l'aide de têtes avancées sur le tambour).
- compression 10 :1
- débit vidéo : 18Mbits/sec
- 4 canaux audio à 48 khz sur 16 bits

- une image est égale à 12 pistes
- rotation du tambour à 75 Hz
- largeur de piste : 32 micromètres
- utilise des bandes MP (metal particule)

Le MPEG-IMX (ou D10)

Positionné sur le marché entre le Betacam digital et le Betacam SX. Premier magnétoscope possédant (en option) une carte réseau avec adressage TCPIP. Les données sont converties en fichiers MXF (Material eXchange Format) et ce pour n'importe quel contenu (Beta SP, Beta D, Beta SX). Les données peuvent ensuite être transférées via un connecteur RJ45 standard vers n'importe quel équipement compatible.

- Compression MPEG2 4 :2 :2 à 3,3 :1
- compression **intra-image**
- Débit 50 Mbits/sec
- compatible en lecture avec tous les systèmes Betacam
- une image est égale à 8 pistes
- 4 canaux audio à 48 KHz sur 24 bits (ou 8 canaux à 48 KHz sur 16 bits)

L'interface SDI (Serial Digital Interface)

- Débit : 270 Mbits/sec
- Véhicule le signal 4:2:2 décompressé avec 4 canaux audio

L'interface SDTI (Serial Data Transport Interface)

- Débit : 270 Mbits/sec-
- Véhicule les données numériques compressées

Le XDCAM

Ligne d'équipement d'acquisition lancé par Sony en 2004 à enregistrement sur disque optique.

Le signal enregistré peut être du DVCAM ou du MPEG-IMX.

Une version haute définition, le XDCAM HD, emploie une compression de type MPEG2.

Une autre version, le XDCAM EX, utilise comme support une carte flash.

Le disque optique est de type Blu Ray à laser bleu, la capacité de stockage est de 23,3Go ce qui donne une durée d'enregistrement de 90 min en DVCam, 75 min en IMX30, 57 min en IMX40 et 45 min en IMX50.

Le XDCAM encapsule les flux vidéo DVCAM ou IMX dans des fichiers MXF, ce qui garantit une compatibilité avec d'autres fabricants.

Le P2

Ligne d'équipement d'acquisition lancé par Panasonic, basé sur le DVCPRO. Il utilise comme support une carte compact flash. Il encapsule les flux vidéo dans des fichiers MXF pour assurer une compatibilité avec d'autres constructeurs.

Une carte P2 est un assemblage de 4 cartes SD (secure digital), il en existe de différentes capacité avec un maximum de 64 Go (en 2009) La durée d'enregistrement est de 256 min en DVCPRO25, 128 min en DVCPRO50 et en AVCINTRA50, 64 min en DVCPROHD et en AVCINTRA100.

Les caméras de type P2 sont équipées de 4 slots pouvant recevoir des cartes p2 et ses cartes se relaient sans interruption de flux.

Le principal avantage de ce système est la disparition de toute pièce mécanique source de nombreuses pannes.

Le DVD

Le DVD (Digital Versatil Disc) est un disque optique à haute densité, conçu pour stocker de la vidéo, de l'audio ou tout autre donnée informatique.

Le format DVD a été prévu afin de fournir un support de stockage universel alors que le CD était originalement prévu en tant que support audio uniquement.

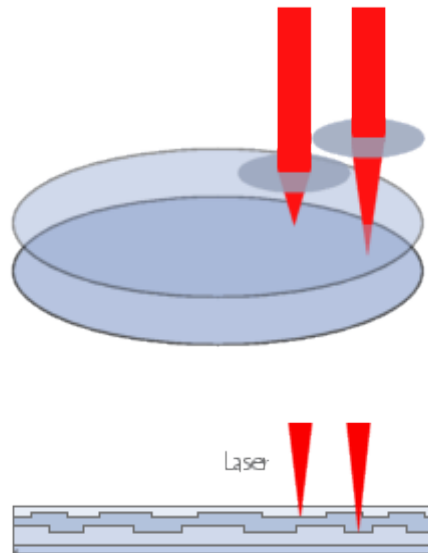
Le format DVD était porté originalement (dès le 15 septembre 1995) par un consortium de dix sociétés du monde multimédia (Hitachi, JVC, Matsushita, Mitsubishi, Philips, Pioneer, Sony, Thomson, Time Warner et Toshiba). A partir de 1997, un nouveau consortium, baptisé «DVD Forum» a succédé au consortium initial.

Un DVD peut facilement être confondu avec un CD dans la mesure où les deux supports sont des disques en plastique de 12 cm de diamètre et de 1.2 mm d'épaisseur et que leur lecture repose sur l'utilisation d'un rayon laser. Toutefois, les CD utilisent un laser infrarouge possédant une longueur d'onde de 780 nanomètres (nm) tandis que les graveurs de DVD utilisent un laser rouge avec une longueur d'onde de 635 nm ou 650 nm. De plus, les lecteurs de CD utilisent généralement une lentille dont la focale vaut 0,5, alors que les lecteurs de DVD sont basés sur une lentille ayant une focale de 0,6. Ainsi, les DVD possèdent des alvéoles dont la taille minimum est de 0,40 μ avec un espacement de 0,74 μ , contre 0,834 μ et 1,6 μ pour le CD.

Les DVD existent en version «simple couche» et «double couche» (en anglais «Dual Layer», noté DL). Ces derniers sont constitués d'une couche translucide semi réfléchissante à base d'or et d'une couche réflexive opaque à base d'argent, séparées par une couche de liaison (bonding layer). Pour lire ces deux couches le lecteur dispose d'un laser pouvant changer d'intensité en modifiant sa fréquence et sa focale :

- avec une intensité faible le rayon est réfléchi sur la surface dorée supérieure,
- avec une intensité plus élevée le rayon traverse la première couche et est réfléchi sur la surface argentée inférieure.

La couche inférieure possède toutefois une densité moindre. De plus l'information y est stockée «à l'envers» sur une spirale inversée, afin de limiter le temps de latence lors du passage d'une couche à une autre.



Par ailleurs, les DVD existent en version simple face ou double face, à la manière des disques vinyles. Dans le second cas, l'information est stockée de part et d'autre du support.

On distingue généralement 4 grandes familles de supports DVD, possédant différentes capacités selon leurs caractéristiques physiques :

Type de support	Caractéristiques	Capacité
CD		650Mo
DVD-5	simple face, simple couche	4.7 Go
DVD-9	simple face double couche	8.5 Go
DVD-10	double face, simple couche	9.4 Go
DVD-17	double face, double couche	17 Go

DVD-R et DVD-RW

Le DVD-R est enregistrable une seule fois, le RW un millier de fois, capacité de 4,7Go.

DVD+R et DVD+RW

Physiquement similaires aux précédents, mais présentent plusieurs avantages :

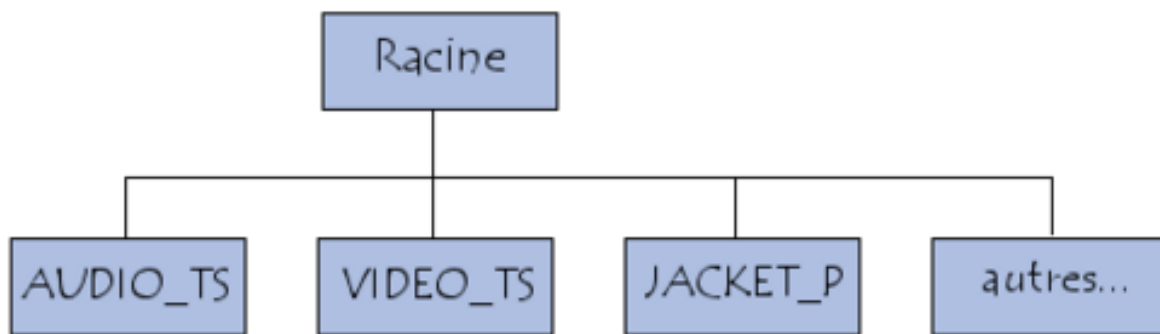
- ils permettent la visualisation des séquences vidéo sans finalisation du disque.
- Ils autorisent une écriture en multisession.
- Permettent de créer un index une fois la vidéo enregistrée.

Codage

Le DVD video utilise une compression MPEG2 **MP@ML** en 4:2:0 avec un GOP de 15.

Structure d'un DVD video

Un DVD-Video possède une organisation hiérarchique de ses répertoires permettant de contenir les données vidéo et audio. Il repose habituellement sur la structure suivante :



Le répertoire principal, nommé VIDEO_TS (pour Video Title Sets), a pour vocation de contenir les fichiers du DVD Vidéo. Le répertoire AUDIO_TS concerne les DVD-Audio mais sa présence est parfois demandée par certains lecteurs DVD de salon.

JACKET_P contient les images des jaquettes du DVD. Il est enfin possible d'y adjoindre d'autres répertoires, pouvant être lus sur un ordinateur.

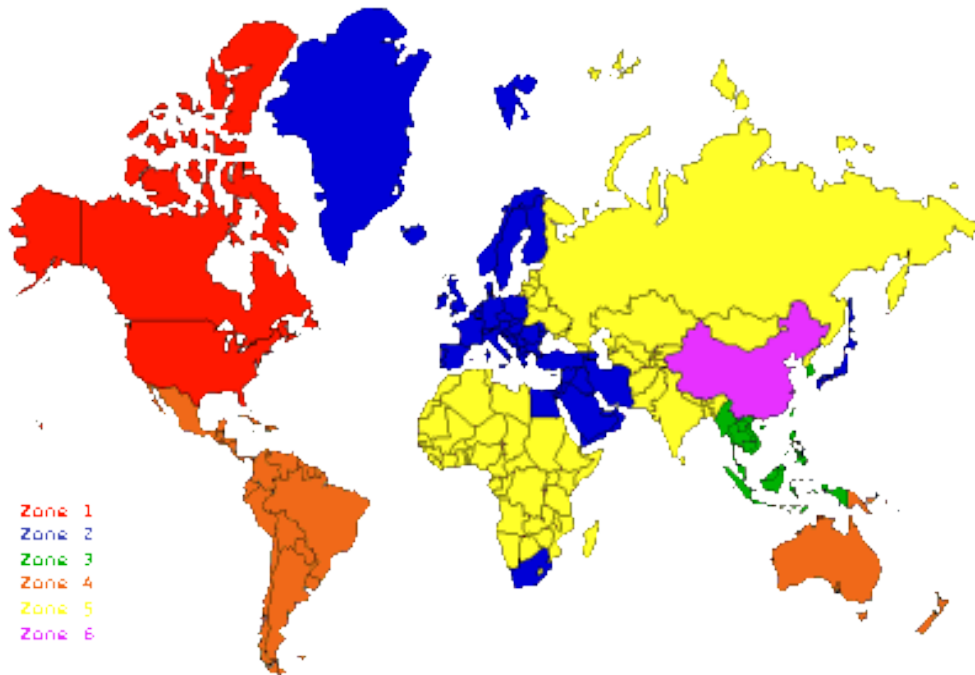
Les «titres vidéo» correspondent à des films, des vidéos ou des albums. Un titre est composé d'un «ensemble d'objets vidéo» (VOBS, Video Object Block Sets), chacun composé :

d'un «fichier de contrôle» (appelé VTSI, pour Video Title Set Information), et contenant les données de navigation.

d'un ou plusieurs objets vidéo (VOB, Video Object Block). L'objet vidéo (VOB) est l'élément de base du DVD. Il contient des données vidéo, audio et des images multiplexées, au format MPEG2. Ainsi, un fichier .VOB peut être lu par un lecteur vidéo logiciel en changeant son extension en «.MPG». Les spécifications du DVD imposent que chaque fichier VOB ne dépasse pas un giga-octet. Chaque VOB est lui-même composé de «cellules» (Cells), représentant les différents clips vidéo ou audio composant le VOB : par exemple des chapitres vidéo ou les chansons d'un album.

Zones

Les DVD Vidéo sont conçus pour ne pouvoir être consultés que dans certaines régions du monde : il s'agit du découpage en zone (prévu initialement pour limiter la diffusion des copies illicites). Il est ainsi théoriquement impossible de lire un DVD d'une zone en étant situé dans une autre. Néanmoins, la quasi-totalité des lecteurs de DVD pour ordinateurs et une grande partie des lecteurs de salon peuvent être «dézonés» grâce à des utilitaires.



Le Blu-ray

Le disque Blu-ray est un média optique conçu pour le stockage de vidéos haute définition et de données. Il présente les mêmes dimensions qu'un CD ou un DVD standard.

Ce format a fini par s'imposer après une « guerre de standard » l'opposant au format HD DVD, soutenu essentiellement par Microsoft et Toshiba. Ce format avait l'avantage d'être moins cher, mais se contentait d'un espace de stockage un peu plus réduit. À la suite de l'annonce par Toshiba d'abandonner en février 2008 le développement des produits HD DVD, l'ensemble des autres partenaires ont abandonné le format, laissant Blu-ray s'imposer en tant que standard du stockage haute définition.

Son nom provient de l'utilisation d'un laser bleu (en fait plutôt violet) pour la lecture ou l'écriture des données. Ce laser a une longueur d'onde de 405 nm, alors que les lasers présents dans les lecteurs CD ou DVD classiques utilisent respectivement des lasers infrarouges proches (780 nm) ou rouges (650 nm). Le laser bleu, avec sa longueur d'onde plus faible, permet de lire ou écrire des « trous » plus petits sur un même espace, ce qui permet par là même d'augmenter le nombre de sillons du disque et donc de stocker un plus grand nombre d'informations.

Sa capacité courante est de 25 Go en simple face et 50 Go en double face. Trois standards sont définis : le BD-R (disque enregistrable), le BD-RE (réinscriptible) et le BD-ROM (lecture seule).

Des disques BD-ROM à 4 couches sont également possibles : il existe ainsi des disques de 100 Go utilisant 4 couches de 25 Go, et de même de 200 Go utilisant 6 couches de 33,3 Go. Pour le moment, tous les lecteurs du marché ne sont pas capables de les lire. À terme, il pourrait exister jusqu'à 10 couches (par face).

Compression

Le blu-ray spécifie trois formats de compression pour la vidéo : MPEG-2, H.264/AVC, SMPTE VC-1. Le codec MPEG-2 est le standard actuellement utilisé pour les DVD. En utilisant ce codec, il est possible de placer deux heures de vidéo haute définition sur un disque simple couche, car le MPEG-2 haute définition nécessite un taux de transfert d'environ 25 Mbit/s (chaque seconde de film utilise donc 3 Mo). En comparaison, le codec H.264/AVC (qui est une évolution du MPEG-4) et le codec SMPTE VC-1 (dérivé des développements de Microsoft pour son codec inclus dans Windows Media Player 9) permettent de stocker jusqu'à 4 heures de film sur le même support.

Le Blu-ray est aussi utilisé dans le broadcast puisqu'il est le support d'enregistrement du système XDCAM de Sony (en cartouche de protection)

La Haute Définition

La **Haute Définition (HD** ou en anglais **High Definition**) désigne une classification d'équipements, appareils, normes de télédiffusion et vidéo. Elle est généralement associée au format d'image « large » ou 16/9.

Ce sigle est utilisé pour désigner un ensemble de normes vidéo numériques.

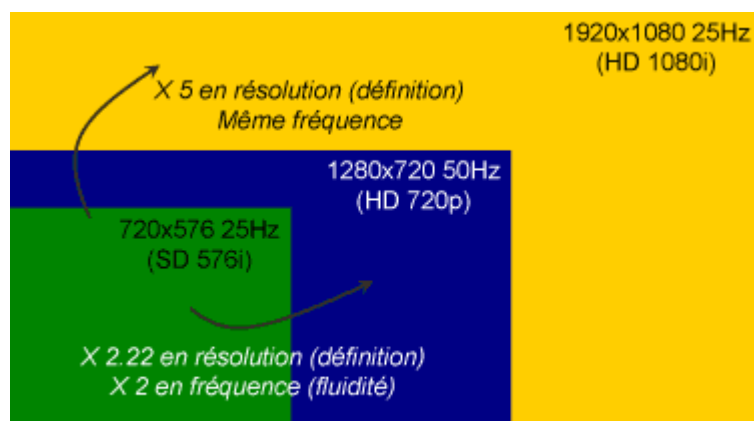
Il représente une évolution de la qualification de définition « standard », SD (Standard Definition) ou SDTV.

Le terme HD s'étend à l'ensemble des technologies audiovisuelles numériques tels que HDV (grand public), HDCam (professionnel), télévision à haute définition (DVB-T, DVB-S, DVB-C, *DSL TV*...) ainsi que les supports Blu-ray Disc, HD DVD, l'enregistrement multimédias sur disque dur ainsi notamment qu'au stockage de données informatiques.

La HD nécessite des capacités techniques bien plus évoluées à cause du volume de données à traiter bien plus important. Par exemple, un film DVD avec une définition de 576 lignes (encodage MPEG-2) a un volume moyen de 7 à 8 Go, alors qu'en HD, le même film a une définition plus que doublée et son volume peut facilement atteindre les 20 à 30 Go MPEG-4.

Il existe deux normes d'images à haute définition :

- La norme SMPTE 274M
 - 1920 X 1080 pixels visibles
 - fréquence de rafraîchissement de 24, 25 ou 30 images par seconde en balayage progressif (1080p), et de 50 ou 60 trames par seconde (1080i).
- La norme SMPTE 296M
 - 1280 X 720 pixels visibles
 - fréquence de rafraîchissement 24, 25 ou 30 images par seconde, uniquement en progressif (720p).



Source : France Télécom

On s'oriente en Europe aujourd'hui vers deux normes labélisées HD Ready : 1080 lignes de 1920 points entrelacé (1080i) et 720 lignes de 1280 points (720p), le nombre de pixels est doublé par rapport au SDTV. le Full HD correspondant à 1080 lignes de 1920 points en progressif.

Les principaux formats d'image utilisables sont les suivants :

- 720p : 1280 X 720 à 50 et 60 Hz en progressif
- 1080i : 1920 X 1080 à 50 et 60 Hz en entrelacé
- 1080p : 1920 X 1080 en progressif

Dans ces normes, on trouve deux catégories de formats d'enregistrement.

Le D6, le D5-HD et le HDCAM SR sont réservés à la production haut de gamme ainsi qu'au mastering film.

Parmi eux, seul le HDCAM SR est un format d'acquisition, les autres n'existent qu'en magnétoscope de studio.

Le D6 est le seul à enregistrer le format HD sans compression.

Dans la deuxième catégorie, on trouve le DVCPRO-HD, le D9-HD et le HDCAM.

Ils visent la production HD courante et existent sous la forme de caméras et de magnétoscopes.

Tous ces systèmes fonctionnent avec des fréquences d'échantillonnage très élevées pour répondre à la lois de Shannon et Nyquist.

La plupart échantillonnent la luminance à 74 MHz et les composantes de chrominance à 37 MHz, si l'on se réfère au 4 :2 :2, nous aurions ici du 22 :11 :11.

Le débit de ce genre de machine peut monter à 1,485 Gbits/sec....

Notons que le HDCAM-SR peut gérer le 22 :22 :22 , que le D9-HD et le DVCPRO-HD travaillent en 14 :7 :7 et que le HDCAM travaille en 17 :6 :6.

Notons aussi qu'en haute définition le nom de 4 :2 :2 a été gardé bien que les fréquence d'échantillonnage soient plus élevées .

- **Algorithme de compression**

Les taux de compression varient de 2,7:1 en HDCAM SR à 25:1 en HDV.

Les types de compression peuvent être intra-image ou long GOP (type MPEG).

- M-JPEG pour D5-HD et HDCAM
- DV pour DVCPRO HD
- MPEG4 ASP pour HDCAM SR
- MPEG4 H264 pour les AVCINTRA (codage intra) et AVCHD (codage inter)
- MPEG2 pour XDCAM HD XDCAM EX et HDV

- **Supports de stockage**

Les plus anciens formats sont basés sur la bande magnétique, les XDCAM sur Blu Ray ou carte flash, les P2 sur carte flash.

1/ Les formats HD non linéaires

- **XDCAM EX** :MPEG2 HD 4:2:0, 25,35Mbits/sec, 1440/1080 GOP=12 8 bits

- **AVCHD** : MPEG4, 4:2:0, 24 Mbits/sec 1440/1080 GOP=12 8bits
- **XDCAM HD** : 2 possibilités : - mpeg2, 18,25,35Mbits/sec 1440/1080 gop=12 8 bits, 4:2:0
 - MPEG2, 50Mbits/sec, 1920/1080 gop=12, 8 bits, 4:2:2
- **P2HD** : 3 possibilités : - DV, 100Mbits/s, 1220/1080, Gop=1 (intra) 10 bits, 4:2:2 (DVCPRO HD)
 - MPEG4, 50Mbits/s, 1440/1080, gop=1, 10 bits 4:2:0 (AVCINTRA50)
 - MPEG4, 100Mbits/s, 1920/1080, gop=1, 10 bits 4:2:2 (AVCINTRA100)

2/Les formats HD sur bande magnétique

	D9-HD	DVCPRO-HD	HDCAM	HDCAM-SR	D5-HD	D6(VooDoo)
fabriquant	JVC	Panasonic	Sony	Sony	Panasonic	Philips, Thomson, Toshiba
bande	½ pouce	¼ pouce	½ pouce	½ pouce	½ pouce	¾ pouce
Débit vidéo(Mbits/sec)	100	100	140	600	235	920
compression	DV 6,7 :1	DV 6,7 :1	MJPEG 4 ,4 :1	MPEG4 2,7 :1	MJPEG 4 :1 en 8bits 5 :1 en 10 bits	Sans compression
Standard supportés	720p 1080i	720p 1080i 24p /25p	1080i 24p /25p	1080i 24p /25p	720p 1080i 24p	1080i 24p
Résolution H	-En 720p Y=816 pixels C=408 pixels -en 1080i Y=1220 pixels C=610 pixels	-En 720p Y=816 pixels C=408 pixels -en 1080i Y=1220 pixels C=610 pixels	Y=1440 pixels C=520 pixels	-En 4 :2 :2 Y=1920 pixels C=960 pixels -en 4 :4 :4 Y=1920 pixels C=1920 pixels	-En 720p Y=1280 pixels C=640 pixels -en 1080i Y=1920 pixels C=960 pixels	Y=1920 pixels C=960 pixels
Fréq d'éch	Y=47MHz C=23MHz	Y=47MHz C=23MHz	Y=57MHz C=20MHz	-en 4 :2 :2 Y=74MHz C=37MHz -en 4 :4 :4 Y=74MHz C=74MHz	Y=74MHz C=37MHz	Y=74MHz C=37MHz
Pistes audio	8	8	8	12	8	12

Le HDV

Le HDV se positionne sur le marché amateur, voir semi-pro.

Caractéristiques du format :

Média	cassettes DV	
Signal vidéo	720/25p	1080/50i
Nbre de pixel	1280x720	1440x1080
Aspect	16/9 ou 4/3	
Compression	MPEG2 MP@H14	
Echantillonnage (luminance)	74,25MHz	55,7MHz
Rapport d'échantillonnage	4 :2 :0	
Quantification	8bits	
Débit	25Mbits	
GOP	6	
Interface	IEEE1394	

Le HDV est construit par plusieurs fabricants et existe sous différentes formes, par ex : mono ou tri ccd, épaulement ou à main, objectifs interchangeables ou fixes...

L'AVCHD

Advanced video codec high definition, développé par Sony et Panasonic en 2006. Il emploie comme support le mini DVD ou une carte compact flash, il est destiné au marché grand public.

Il gère le 1080i et le 720p avec un codec MPEG4 AVC/H264 et un gop de 12 images.

Comme pour le HDV, la conversion sur une station de montage en un format intermédiaire intra image nécessite un temps de conversion assez important, l'autre solution est de passer par un programme compatible avec le mode natif, mais celui ci nécessite une grande puissance de calcul.

L'AVCHD offre une efficacité de compression double du HDV (MPEG4 versus MPEG2).

La structure d'échantillonnage est du 17:8,5:0 (comme un 4:2:0) avec une résolution de 1440/1080.

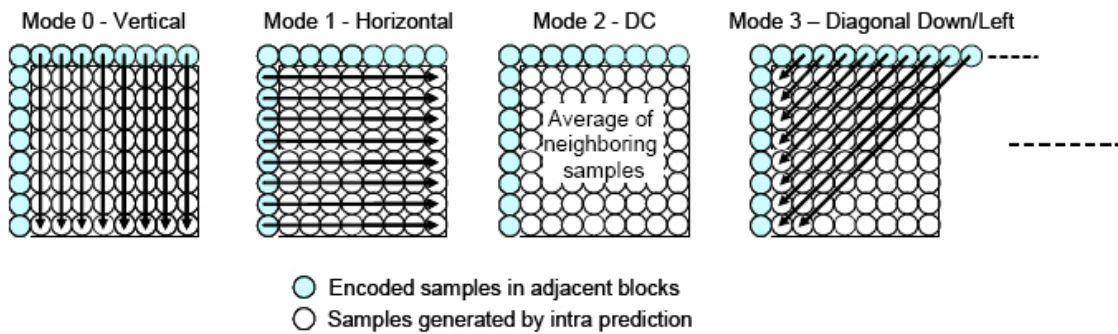
La quantification s'effectue sur 8 bits.

L'AVCINTRA

Compression intra-image propriétaire de Panasonic, utilisé sur les caméras de type P2.

Une caméra P2 peut utiliser deux formats : le DVCPRO HD et le MPEG-4 AVC/H264. Dans ce dernier mode, Panasonic utilise la compression AVCINTRA (totalement différent de l'AVCHD qui utilise un long gop), ici, pas de gop, mais uniquement une prédiction intra-image (ce qui n'est pas non plus une compression JPEG, on parle ici de prédiction).

La prédiction INTRA est une prédiction spatiale intra-image, par exemple, sur le dessin ci-dessous, sur un block de 8x8, différents type de prédiction peuvent être faits (8 en tout), le logiciel intégré dans la P2 choisit la meilleure prédiction possible à partir de l'image.



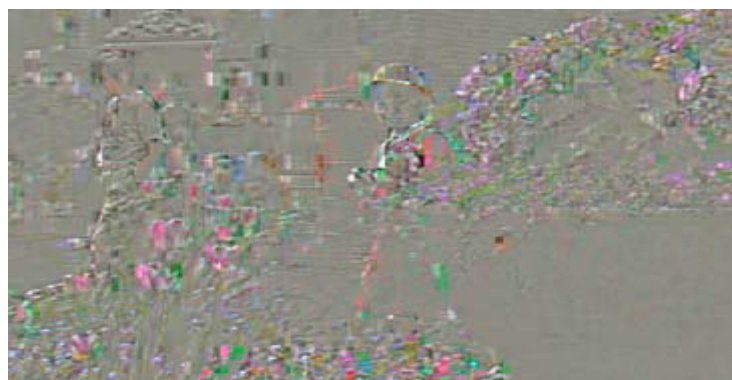
Comme dans toute prédiction (voir MPEG), on compare le résultat du calcul avec l'image réelle, le fichier d'erreur est ajouté au calcul.



Input image



image prédite



erreur résiduelle

Liste actualisée des formats HD

Format	Paramètres	Luma/chroma	Compression	Codec	Débit	Audio	Support	Interfaces
DCI 2K JPEG2000 D-Cinema	4096x2160 (4K) 2048x1080 (2K)	4:4:4 12 bits X Y Z	I frames only 7:1 max	Jpeg2000 Wavelets MXF	250 Mbps VBR	16 canaux PCM 24 bits 48/96 Khz	Disque dur Satellite ADSL - Fibre	Dual HD SDI
Sony HDCAM SR	1920x1080	4:4:4 10 bits 4:2:2 10 bits (1920px en Y - 960px en C)	I frames only 4:2:2 en 2,7:1 4:4:4 en 4,2:1 ou 4:4:4 e 2,1:1	Mpeg-4 SP	440 Mbps 880 Mbps	12 canaux PCM 24 bits 48 Khz	K7	HD SDI
Sony HDCAM	1440x1080	3:1:1 8 bits (1440px en Y - 480 en C)	I frames only 4:1	DCT	140 Mbps	4 canaux PCM 20 bits 48 Khz	K7	HD SDI
Sony XDCAM HD 50	1920x1080	4:2:2 8 bits (1920px en Y - 960 en C) 4:2:2@HL	Long GOP	Mpeg-2 MXF	50 Mbps	4 canaux PCM 20 bits 48 Khz	Blu Ray XDCAM	IEEE1394 HD SDI

Panasonic AVC intra	1920x1080 1440x1080	4:2:2 10 bits (1920px en Y - 960 en C) 4:2:0 10 bits (1440px en Y - 720x540 C)	I frames only	Mpeg- 4 H264 MXF	100 Mbps 50 Mbps	2 canaux PCM 16 bits 48 Khz	P2 memory card	IEEE1394
Sony XDCAM HD	1440x1080	4:2:0 8 bits (1440px en Y - 720x540 C) MP@H1440	Long GOP	Mpeg- 2 MXF	HQ 35 Mbps VBR SP 25 Mbps CBR LP 18 Mbps VBR	2 canaux Mpeg1 L2/PC M 16 bits 48 Khz	Blu Ray XDCAM	IEEE1394 HD SDI
Sony XDCAM EX	HQ 1920x1080 SP 1440x1080	4:2:0 8 bits (1920px en Y - 960x540 C) MP@HL MP@H1440 (1440px en Y - 720x540 C)	Long GOP	Mpeg- 2 MXF	HQ 35 Mbps VBR SP 25 Mbps CBR	2 canaux PCM 16 bits 48 Khz	SxS ExpressCard	IEEE1394 HD SDI
Panasonic DVCP 100 HD	960x720	4:2:2 8 bits (960px en Y - 480px en C)	I frames only 6,7:1	DV x 4	100 Mbps CBR	4 canaux 16 bits 48 Khz	K7 DVCP Pro	IEEE1394

HDV	1440x1080	4:2:0 8 bits (1440px en Y - 720x540 C)	Long GOP	Mpeg- 2 TS	25 Mbps CBR	2 canaux Mpeg1 Layer2 16 bits 48 Khz	DV	IEEE1394 HD SDI
AVCHD	720x576 1280x720 1440x1080 1920x1080	4:2:0 8 bits	Long GOP	Mpeg- 4 H264 TS	De 5 à 24 Mbps	2 – 5,1 AC3 16 bits 48 Khz <640 Kbps	Disque dur DVD SD Memory	
Panasonic D5 HD	1920x1080	4:2:2 10 ou 8 bits	I frames only 5:1 en 10 bits 4:1 en 8 bits	DCT	270 Mbps CBR	8 canaux PCM 24 bits 48 Khz	K7	HD SDI
Philips D6 Voodoo	1920x1080	10 bits (Y) 8 bits (C)	I frames only 1:1	DPX	1 Gbps	12 canaux 20 bits 48 Khz	K7	data

Captation HD et cinéma numérique 2K/4K

1. Captation Broadcast HD (1920x1080)

La gamme broadcast HD est généralement caractérisée par 2 paramètres importants vis à vis des gammes inférieures. Ces différences ont des conséquences sur la qualité de l'image ainsi produite:

taille des capteurs 2/3'

format d'enregistrement 4 :2 :2 ou 4:4:4

2. Cinéma numérique (2K / 4K)

Les caméras destinées à la sphère cinéma numérique partagent certaines particularités:

- Un capteur de taille supérieure aux capteurs 2/3' broadcast (souvent une taille équivalente à l'image 35mm)
- Un format couleur 4 :4 :4
- Une gamme d'optique compatible avec la gamme 35mm
- Une cadence d'image à 24 ou 48 images en progressif

Voici une liste non exhaustive de ces caméras :

Panavision GENESIS (et la future EXODUS) Mono CCD taille Super 35mm 12 Mpixels

Panalog sur 10 bits / Monitoring avec LUT REC: HDCAM SR ou SSR1 en Flash memory

4:2:2 (960px en C / 1920px en Y) ou 4:4:4



Sony F23 (2/3')-SRW-9000 (plus compacte) et la **F35** (mono capteur CCD S35) avec un capteur proche de la génésis

3x 2/3' pour F23 et SRW-9000 ou Capteur taille 35mm pour la F35

Traitement 14 bits / S-Log 10 bits ou Hyper Gammas

Monitoring avec LUT

Amélioration du prisme / Gamut / S/N amélioré

REC: HDCAM SR ou SSR1 en Flash memory

4:2:2 (960px en C / 1920px en Y) ou 4:4:4

Ralenti / Ramping



Arri D21

Mono CMOS Super 35mm

Choix de courbes de réponse (type Hyper Gammas) pour la sortie HD

Mode HD 4:2:2 ou 4:4:4 type HDCAM SR (Mode Datas/raw à venir)

ou

DATA: Fichiers images DPX 10bits Log (venom)



Thomson Grass Valley Viper

3 x 2/3' CCD

Venom Flash Pack avec monitoring HD

HDTV single HD SDI 4:2:2

RAW: Film Stream dual link HD SDI 4:4:4

Video Stream



Red One

Super 35mm CMOS 12 Megapixel Mysterium

HD, 2K, 4K (à partir d'un mono capteur 4K...)

REC datas et monitoring HD à fixer (LUT intégrée prochainement en interne)

4096x2304@220 Mbps

RAW compressé en Wavelets Redcode 12 bits sur Compact Flash (fichiers R3D avec QT référence)

ou Red Drive (disque dur)



La plupart des caméras appelées "cinéma numérique" partage une autre philosophie de travail vis à vis du monde habituel de la télévision :

- le format d'enregistrement des valeurs de pixels est déconnecté du protocole standard de diffusion vidéo (noirs à 0V et blancs à +700 mV). Il faut donc "interpréter" (en post prod par exemple) le signal enregistré, avant d'en tirer ("développer" l'image comme en 35mm) une image "visible" sur un dispositif vidéo; cela demande donc un effort d'adaptation sur toute la chaîne numérique pour comprendre un fonctionnement qui peut surprendre.

Deux philosophies extrêmes:

1/ Le format dit RAW (brut), où les données enregistrées proviennent directement du capteur sans aucun traitement, mais rendant la phase de post-production plus lourde: avant de voir une image 'droite', il faut traiter les données comme le font les autres caméras en interne, mais cette fois ci en post prod, depuis les sous-pixels R, G et B de la matrice, dé-matricer (Bayer par exemple) pour reconstruire une image en couleur dans un espace colorimétrique cible, réaliser une réduction de bruit, effectuer une balance des blancs, etc...: c'est la philosophie qui offre la plus grande marge de manœuvre à l'étalonnage; l'image ne peut être ni visionnée ni diffusée telle quelle en sortie de caméra (comme un négatif numérique qu'il faudrait 'révéler' avant de la regarder). Il y a obligatoirement une vie après le tournage...

2/ Le mode 'News' ou 'droit' qui 'clip' l'image dans les hautes lumières et enterre les basses lumières en donnant une image rapide et très contrastée, avec comme seul réglage le diaphragme: C'est la philosophie qui offre le moins de manoeuvre à

l'étalonnage (ce qui est perdu au tournage est perdu pour toujours), mais l'image est prête à être diffusée telle quelle. Il n'y a pas ou peu de vie après le tournage...

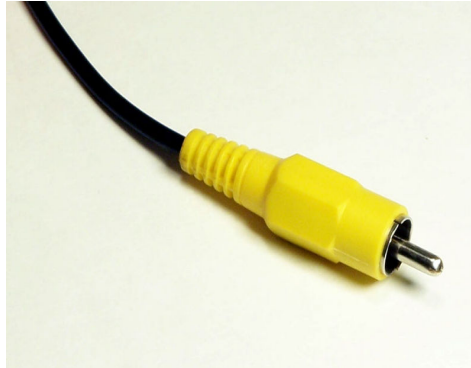
Entre ces deux extrêmes, chaque philosophie trouve sa place, et certaines caméras peuvent fonctionner dans plusieurs modes (linéaire, knee et gamma, courbes en preset, courbes persos, raw)...

Avant l'enregistrement, des courbes de réponse adaptées (users gamma) sont utilisées dans la caméra; après l'enregistrement, des LUT (Look Up tables) 'révèlent' et interprète l'image finale sur un moniteur.

Les interfaces entre les différents appareils

1/connexions analogiques :

a/ Composite



Elle existe dans tous les appareils grand public, c'est en général une fiche RCA jaune. Dans le domaine pro, ce signal aura un connecteur de type BNC (Bayonet Neill-Concelman).



Ce type de signal est essentiellement utilisé pour du monitoring.

b/Composant

Longtemps cantonné dans le professionnel, le composant existe maintenant sur la plupart des lecteurs DVD de salon, signal de meilleure qualité que le composite (pas de sous-porteuse) il permet d'afficher un signal progressif (sur les écran plats).

Nous aurons donc 3 connecteurs (Y R-Y B-Y), parfois notés Y dR dB, en RCA ou BNC.

Ce type de signal existe toujours sur la plupart des magnéto et caméras professionnelles.

2/connexions numériques

a/ IE1394 ou firewire (déjà décrite plus haut)



Deux brochages distincts existent en s400 et s800 : le format à 6 broches permettant l'alimentation des périphériques et le format à 4 broches sans alimentation. Le format à quatre broches est celui des PC portables.

Ces connecteurs existent sur la plupart des magnétos et caméras professionnelles.

B/ Le SDI

La Serial Digital Interface ou Interface Numérique Série, est un protocole de transport ou de diffusion des différents formats de vidéo numérique.

Dans une interface série, tout les mots de 8 ou 10 bits sont transmis les uns à la suite des autres dans un seul câble type RG59 avec connecteur BNC.

Elle se décline en 3 versions :

SDI, pour les signaux 4.2.2 non compressés

SDTI, pour les signaux compressés

HD-SDI, pour la haute définition.

Les SDI ne se rencontrent que dans le domaine professionnel.

C/ HDMI

High Definition Multimedia Interface est une interface de connexion audio/video pour les équipement numériques grand public.

Le HDMI gère aussi bien la SD que la HD, avec ou sans compression.

La HDMI encapsule les données selon la technologie TDMS (transmission minimized differential signaling) permettant de transférer de grande quantités de données dans un câble blindé.

Le HDMI comporte 3 canaux TDMS, chacun doté d'une capacité de transfert de 3,4Gbits (total 10,2Gbits).

Le HDMI inclus le HDCP (High Bandwith Digital Content Protection) qui empêche le piratage des flux numériques non compressés.

Longueur max : 15m



Cinéma numérique 2K/4K

Généralités

Le cinéma numérique ne doit pas être confondu avec la télévision à haute définition. En effet il ne dépend pas de l'utilisation de la télévision ou des standards de la TVHD. Le développement du cinéma numérique se fonde sur un standard proposé par les 7 principaux studios américains réunis dans une structure commune appelée Digital Cinema Initiatives (DCI). Ce standard permet de disposer d'équipements qui exploitent les films quelle que soit leur origine, sous réserve que leur préparation, ou encodage, soit effectué en respectant ce standard. Repris par la SMPTE avec le standard SMPTE 428-1, il fait l'objet de normalisations ISO publiée 2008 sous les références ISO 26428-12 et suivantes.

Note sur la SMPTE :

La Society of Motion Picture and Television Engineers ou SMPTE, fondée en 1916, est une association internationale, située aux É.-U., et composée d'ingénieurs. Elle développe des standards vidéos (elle en a déjà plus de 400 à son actif), qui sont utilisés par exemple par la télévision, ou le cinéma numérique. Les standards les plus significatifs à l'actif de la SMTE incluent :
Tous les formats de transmission de cinéma et de télévision, y compris numérique
Les interfaces physiques pour la transmission de signaux de télévision et les données associées (comme par exemple le timecode SMPTE)
Le Material eXchange Format, ou MXF

Le Digital Cinema Initiatives, ou DCI, est un groupement d'experts du cinéma visant à spécifier une architecture pour le cinéma numérique. Cette architecture vise un haut niveau de performance, de contrôle, et de robustesse. Le DCI a été créé en mars 2002, par les studios hollywoodiens The Walt Disney Company, 20th Century Fox, Paramount Pictures, Sony Pictures Entertainment, Universal Studios, Metro-Goldwyn-Mayer et Warner Bros.

Le DCI est connu pour son document « Digital Cinema System Specification » (DCSS, version 1.2), qui sert de référence pour de nombreux systèmes de création de films (DCP mastering) ou d'affichage de films (projection numérique) pour le cinéma numérique. Il spécifie notamment le format conteneur (DCP) des fichiers contenant les films, ainsi que la compression des images de ces derniers, qui s'appuie sur le format JPEG 2000.

Le DCI reconnaît deux technologies conformes à ses spécifications : le DLP Cinema de Texas instruments et le SXRD de Sony3. La plupart des projecteurs numériques commercialisés dans le monde sont basés sur la technologie de Texas Instruments, également nommée 2K pour sa résolution de 2048 pixels par ligne x 1080 pixels par colonne. Après la commercialisation initiale de la série 1, le constructeur livre à partir de 2010 les matrices DLP Cinema série 2, disponible aussi en résolution 4K. Les trois principaux constructeurs de projecteurs numériques, Christie, Barco et NEC, utilisent les puces DLP Cinema de Texas Instruments.

Les signaux audio-vidéo

Vidéo : Les composantes de couleur (YU(R-Y)V(B-Y), pas RGB) sont quantifiées sur 12 bits (valeurs de 0 à 4095), soit 36 bits par pixel (68,7 milliards de combinaisons)

Audio : Échantillonnage à 48 ou 96 kHz, quantification sur 24 bits par

échantillon.

Les spécifications du DCP

Le Digital Cinema Package (DCP) est la copie d'exploitation destinée à être envoyée dans les salles.

La syntaxe pour la description des éléments du DCP (métadonnées) est le XML. La norme de réduction de débit pour les images est la compression JPEG 2000 (compression par ondelettes). Chaque image est compressée indépendamment des autres qui la précèdent ou qui la suivent (images I, ou intra). L'audio n'est pas compressé, le format retenu est le PCM. Le format de fichier pour le conteneur des essences image et audio est le MXF. Il y a un conteneur MXF qui contient les images, un autre pour l'audio. Il peut y avoir d'autres conteneurs MXF, autant que du pistes audio.

Lors de la fabrication de la « copie 0 » au format DCP à partir du DCDM, le « packaging » abouti au cryptage des pistes audio et images. Une clé de lecture principale (dite « clé primaire ») est générée.

Un Digital Cinema Package (DCP) est l'équivalent en cinéma numérique de la copie **de projection, qui en cinéma traditionnel (en argentique) se présente sous forme**

de bobines de film argentique 35 mm.

Un DCP compose un ensemble de fichiers informatiques (images, sons, sous-titres, méta-données...) qui sont destinés à être stockés et joués dans la cabine de projection par un lecteur de DCP, couplé à un projecteur numérique.

La technologie Digital Light Processing (abrégé en DLP, qu'on pourrait traduire par « traitement numérique de la lumière ») est utilisée dans les vidéoprojecteurs.

Chaque pixel correspond à un micro-miroir actionné par un champ électrique : la partie active peut être intégrée dans une puce DMD (Digital Micromirror Device). Chacun de ces miroirs renvoie ou non la lumière de la lampe vers l'écran. Le rapport cyclique de cet état donne la luminosité de chaque pixel variable de 0 à 100 %. En outre, un filtre tricolore RVB en rotation, situé entre la matrice DLP et la lampe, permet de projeter successivement les trois composantes de l'image finale.

Les avantages sont le contraste, l'absence de rémanence, la luminosité, l'absence de pixellisation, le rendu des teintes foncées.

Les inconvénients sont la perte lumineuse et la fatigue oculaire dues au disque coloré, et la perception par certaines personnes de petits « flashes » d'arc-en-ciel pendant la projection. C'est le « rainbow effect ».

La technologie Tri DLP qui consiste à utiliser trois puces

(une pour chaque couleur primaire) permet de supprimer ces problèmes.

Les autres inconvénients sont les prix plus élevés que les LCD, le bruit et la chaleur du système de refroidissement, la durée de vie de la lampe.

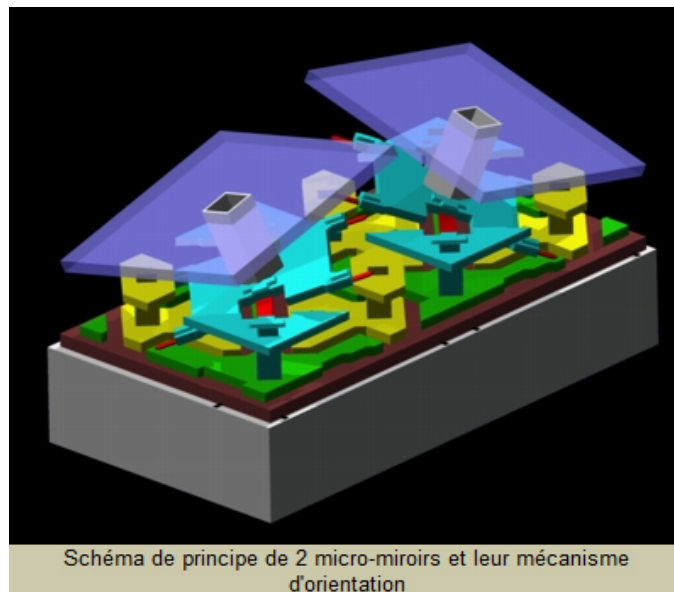
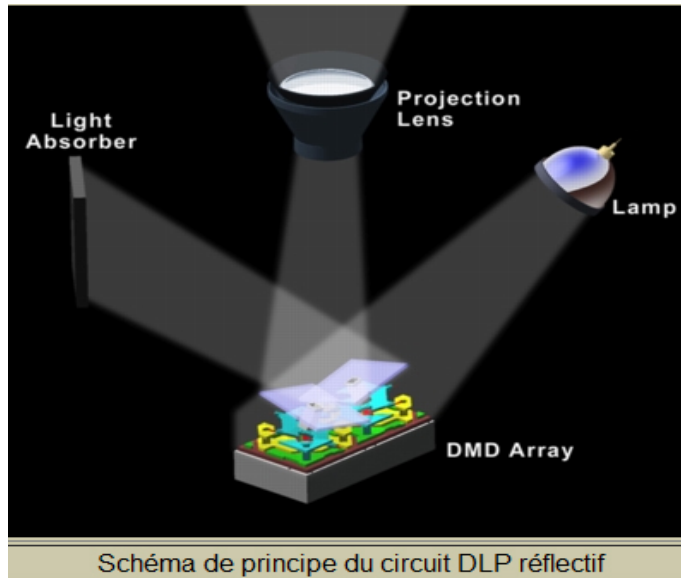
Fonctionnement de la technologie DLP (source Texas Instruments)

Au cœur de chaque système DLP se trouve un semi-conducteur optique, appelé puce DLP, qui fut inventé par le Dr Larry Hornbeck de Texas Instruments en 1987.

Elle est composée d'une matrice rectangulaire qui contient jusqu'à 2 millions de miroirs microscopiques montés sur charnière ; la taille d'un miroir microscopique est inférieure à un cinquième du diamètre d'un cheveu humain.

Lorsqu'une puce DLP est synchronisée avec un signal vidéo ou graphique, une source de lumière et une lentille de projection, ses miroirs peuvent réfléchir une image numérique sur un écran ou toute autre surface.

Les miroirs microscopiques de la puce DLP pivotent en direction de la source lumineuse d'un système de projection DLP (ACTIVÉ) ou s'en éloignent (DÉSACTIVÉ). Cela permet de créer un pixel lumineux ou sombre sur la surface de projection.



Le flux numérique du code de l'image transmis au semi-conducteur active et désactive chaque miroir plusieurs milliers de fois par seconde. Lorsqu'un miroir est plus souvent activé que désactivé, il réfléchit un pixel gris clair. À l'inverse, un miroir qui est plus souvent désactivé qu'activé réfléchit un pixel gris plus sombre.

Ainsi, les miroirs d'un système de projection DLP peuvent réfléchir les pixels en utilisant jusqu'à 1024 niveaux de gris pour convertir le signal transmis à la puce DLP en une image en niveaux de gris extrêmement précise.

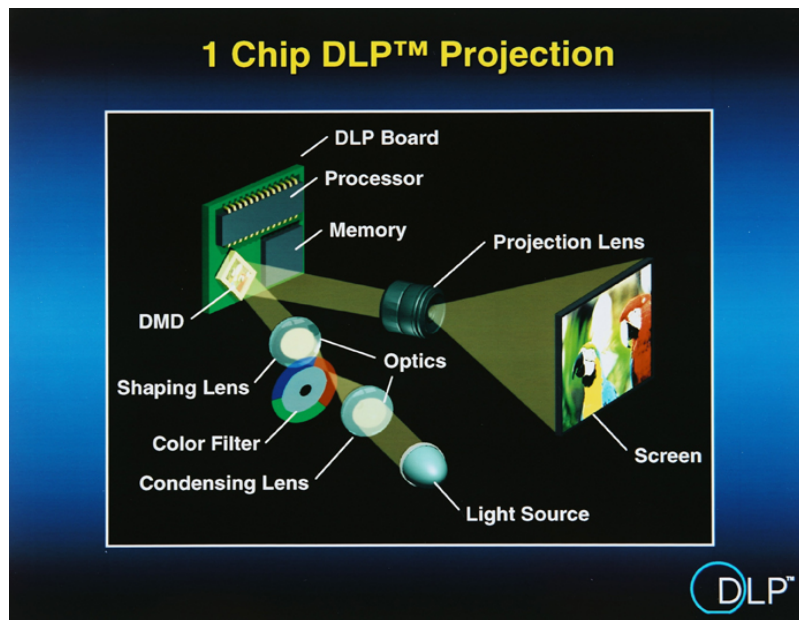
La lumière blanche générée par la lampe dans un système de projection DLP passe à travers un filtre chromatique avant d'atteindre la puce DLP. La lumière est ainsi filtrée en un minimum de rouge, vert et bleu, ce qui permet à un système de projection DLP à une puce de créer au moins 16,7 millions de couleurs.

Certains projecteurs DLP utilisent un système d'éclairage LED, qui remplace la lampe blanche traditionnelle. Par conséquent, la source de lumière émet les couleurs nécessaires, éliminant le filtre chromatique. Certains systèmes DLP reposent sur une architecture à trois puces, notamment les projecteurs destinés aux applications à forte luminosité pour salles de concert ou de cinéma. Ces systèmes sont capables de produire pas moins de 35 trillions de couleurs.

Les états « activé » ou « désactivé » de chaque miroir microscopique sont synchronisés avec les couleurs élémentaires. Par exemple, un miroir devant projeter un pixel violet ne réfléchira que la lumière rouge et bleue sur la surface de projection. Ces couleurs sont ensuite mélangées pour obtenir la nuance voulue de l'image projetée.

SYSTÈME DE PROJECTION DLP À UNE PUCE

De nombreux vidéoprojecteurs et téléviseurs HD dotés de la technologie DLP sont basés sur une configuration à une seule puce comme celle décrite ci-dessus.

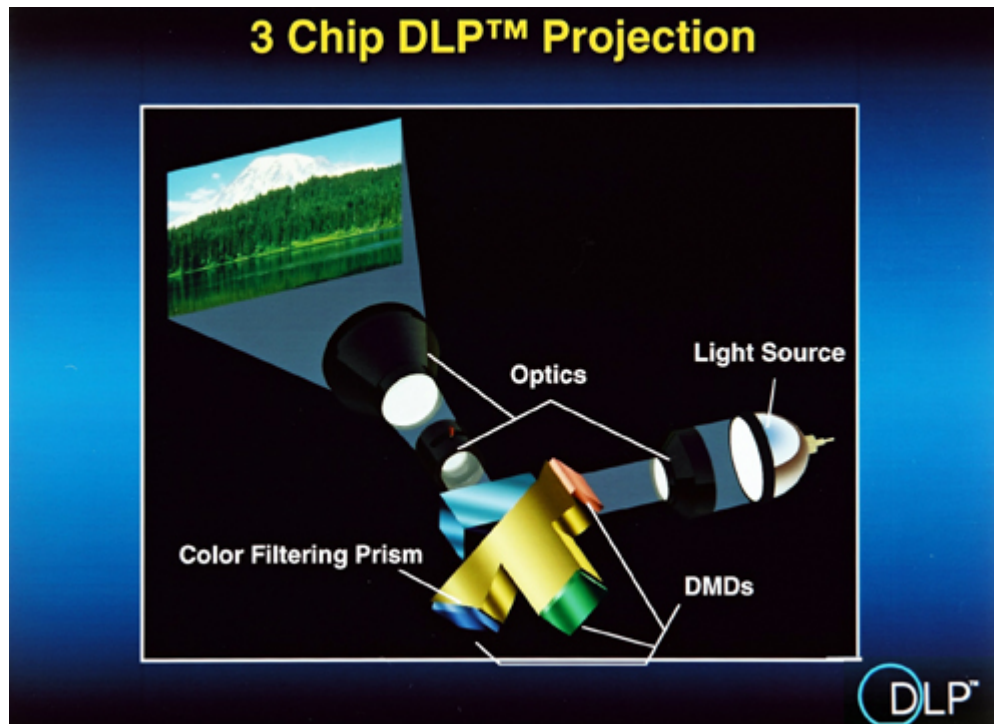


La lumière blanche passe à travers le filtre d'une roue chromatique, provoquant la projection séquentielle de lumière rouge, verte et bleue, ainsi que d'autres couleurs primaires telles que le jaune, le cyan et le magenta, sur la surface de la puce DLP. La commutation des miroirs et la durée proportionnelle de leur « activation » ou « désactivation » sont synchronisées en fonction de la lumière qui leur est envoyée. Les couleurs séquentielles se mélangent ensuite pour former l'image en couleur que vous voyez à l'écran.

SYSTÈME DE PROJECTION DLP À TROIS PUCES

Les projecteurs dotés de la technologie DLP destinés à des applications à forte luminosité, notamment la projection cinématographique ou sur grand écran, reposent sur une configuration à trois puces pour reproduire des images, qu'elles soient animées ou fixes.

Dans un système à 3 puces, la lumière blanche générée par la lampe passe à travers un prisme qui la sépare en rouge, vert et bleu. Chaque puce DLP est identifiée à l'une de ces trois couleurs. La lumière de couleur que les miroirs microscopiques réfléchissent est ensuite amalgamée et passe à travers la lentille de projection pour créer une image.



JPEG2000

jpeg 200 est un standard de compression qui donne une réduction de débit plus importante que le jpeg traditionnel basé sur la DCT. ce format permet en outre d'extraire à partir d'un seul original des images de différentes résolution en fonction des applications visées. Jpeg 2000 supporte le codage vidéo sur 8/10 ou 12 bits et offre une capacité de compression de 20 à 40% plus efficace que le jpeg. Son principal inconvénient est de nécessiter beaucoup plus de ressources aussi bien à l'encodage qu'à la décompression.

La plus grosse différence par rapport au jpeg est de faire appel à une transformée par ondelette plutôt qu'à une dct.

Un filtrage par ondelette consiste à réduire à chaque passage la résolution de l'image d'un facteur 2 et à conserver en mémoire les info de détail retirés. Ensuite l'image est compressée résolution par résolution. le décodage est ensuite directement effectué sur l'image la plus adaptée à l'utilisation cible.

Les informations sont ensuite ordonnées dans le flux binaire de manière croissante en résolution, l'image apparait donc de plus en plus détaillée suivant sa décompression.

les ondelettes visent à séparer les basses fréquences (les aplats) des hautes fréquence (détails contours) qui sont contenues dans une image.

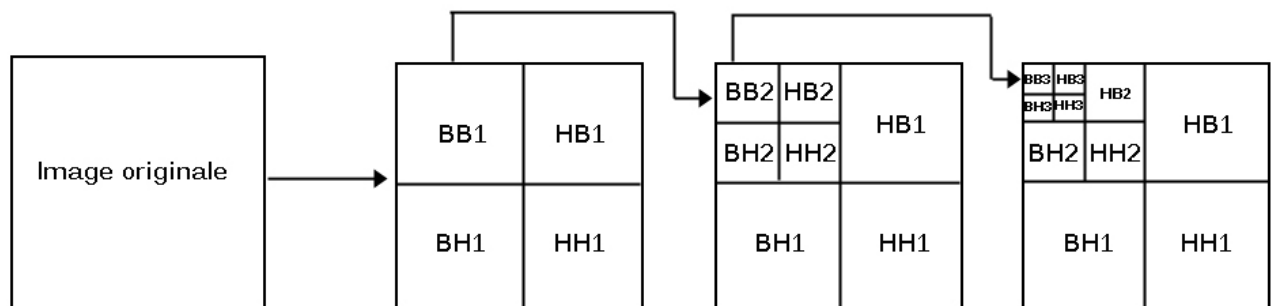
L'objectif est donc le même qu'en DCT, mais la manière est différente. Ici, l'image n'est pas divisée en blocs de pixel comme avec la dct, mais elle est décomposée en sous bandes,

c'est à dire une série d'image de résolutions différentes.

La première étape est un sous échantillonnage d'ordre 2, le résultat donne 4 sous espaces, une sous image dotée d'une résolution réduite de moitié et trois sous espaces renfermant

respectivement les différences en haute fréquence entre la sous image et l'image originale, dans les trois directions, horizontale, verticale, diagonale.

Ensuite un autre sous échantillonnage d'ordre 2 est appliqué à la sous image précédemment obtenue, ce qui donne 4 nouveaux sous espaces, et ainsi de suite.

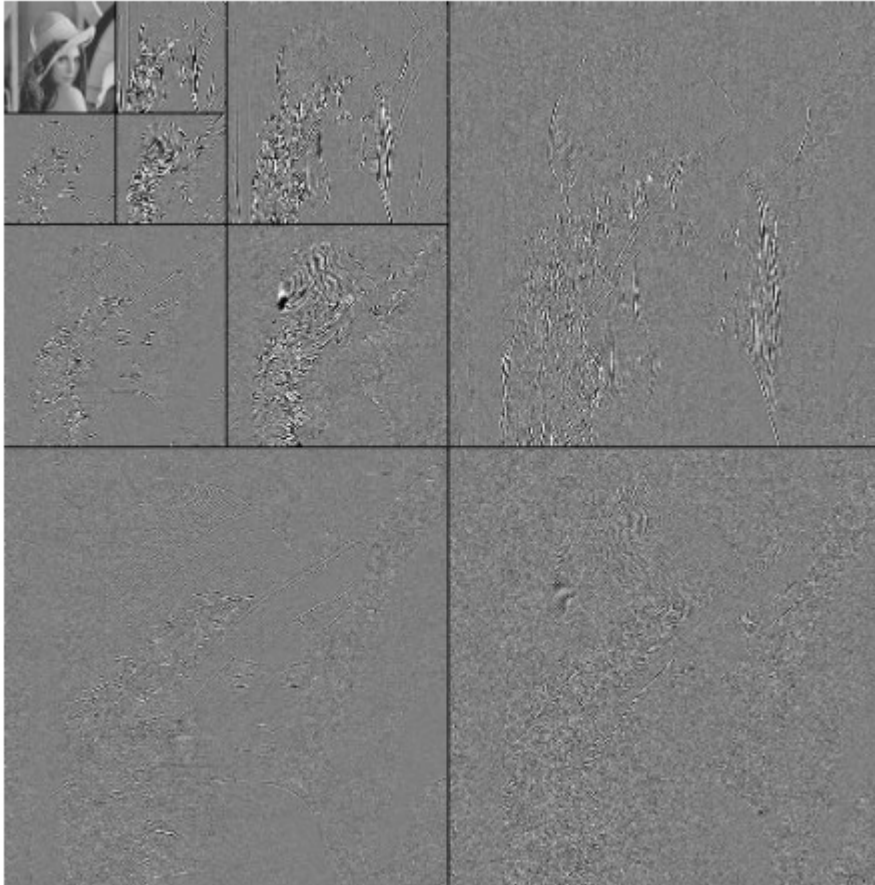


BB=sous bande horizontale et verticale

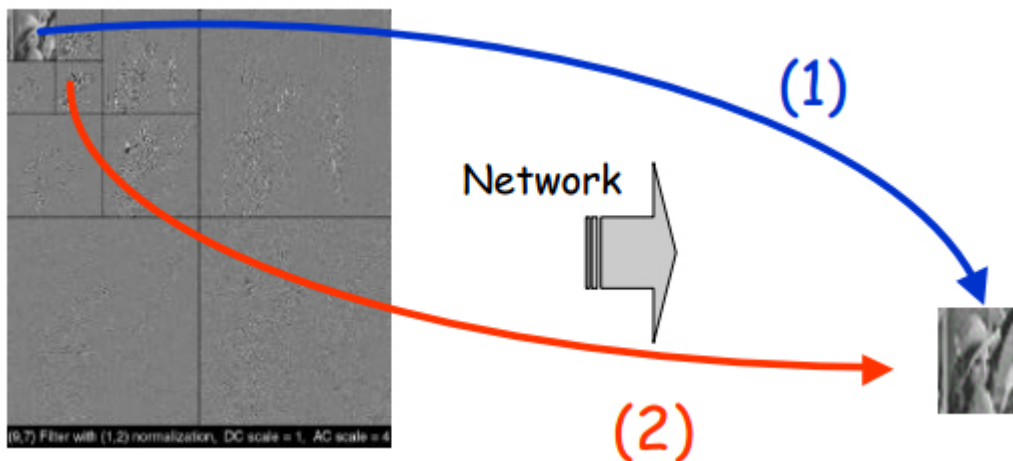
HB= haute fréquence horizontale, basse fréquence verticale

BH= basse fréquence horizontale, haute fréquence verticale

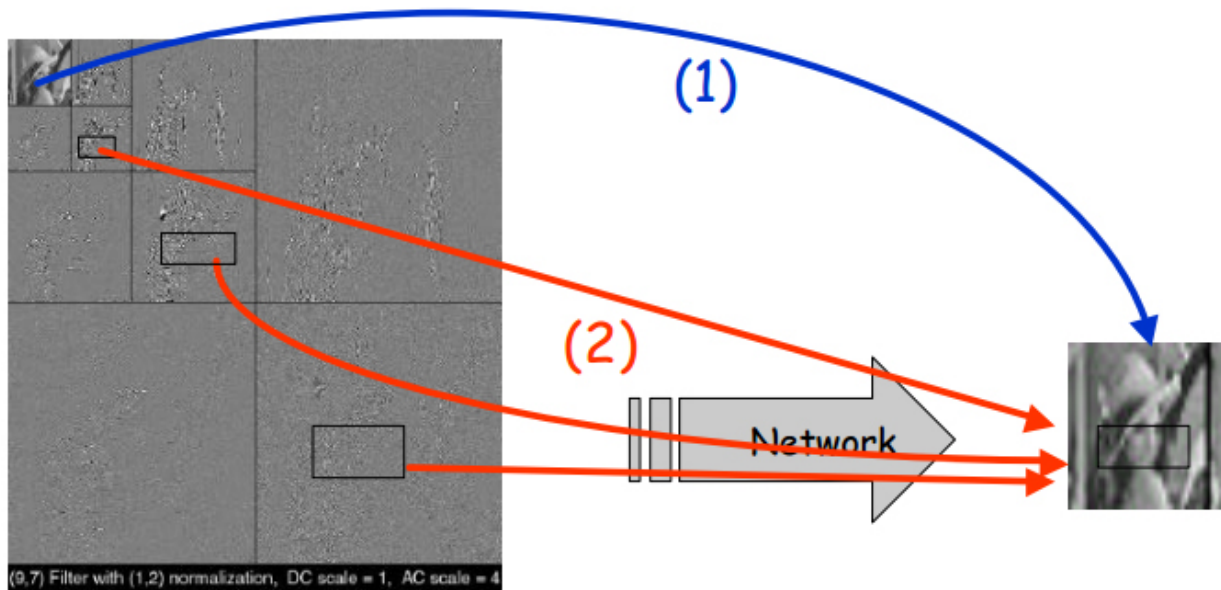
HH= haute fréquence verticale et horizontale.



En pratique si l'on parle d'une image HD, 1920x1080, la transformée en ondelette générera une image 960x540 puis une 480x270 et ainsi de suite... Toutes ces images de plus en plus petites feront partie du flux de données et seront accessibles au décodage (consultation basse résolution sur un « petit » ordinateur).



Lorsque l'on accède à une image JPEG2000 par un réseau, on peut extraire une image basse résolution (1), si l'on veut une meilleure résolution, des informations peuvent être extraites du serveur pour améliorer la qualité (2).



JPEG2000 supporte aussi l'extraction sélective de données pour une zone de l'image particulière.

La quantification s'effectue sur le même principe qu'en JPEG, un seuil est défini pour chaque niveau de détail, seuil en dessous duquel les coefficients sont abandonnés.

Pour résumer :

Le D-Cinema DCI, qui répond à un cahier des charges précis (DCI) et à une norme (SMPTE, AFNOR, etc...): Jpeg2000, 2K, Espace couleur XYZ, 4:4:4, 12 Bits... etc, et qui ne concerne **QUE** le processus industriel de distribution des œuvres cinématographiques, **QUEL QUE SOIT** le matériel utilisé en captation (35mm, HD, etc...). On parle alors d'une norme qui permet l'interopérabilité des matériels, quelque soit le pays et l'œuvre distribuée, comme l'était le format 35mm Edison.

Pour fabriquer une œuvre destinée à une exploitation dans une salle de cinéma numérique labellisée DCI, on peut donc utiliser un large choix de matériel (HD, 2K, 4K...) en captation, sans que cette captation ne soit encadrée par une norme quelconque.

Introduction aux philosophies de captation numérique (source Le repaire)

On peut proposer 6 grandes familles philosophiques de caméras HD:

La philosophie « News », avec une image linéaire fabriquée au diaph (toutes les cams utilisées pour fabriquer une image news, ainsi que Z1, etc...).

La philosophie « RAW » (à l'opposé du mode News) qui déplace une grande partie des choix techniques en post-production (Viper, Red one...).

La philosophie « Gamma et Knee » qui permet de 'sculpter' l'image de son choix manuellement lors du tournage (HDW-750 et les copines..., HVX200, EX1, Z7, etc...). On peut bien entendu réaliser des profils sur des "scenes files" sur memorstick par exemple pour les recharger selon les besoins, selon les cams.

La philosophie « Hyper Gamma », des courbes toutes prêtes qui permettent d'exploiter au mieux la dynamique de la caméra dans la plupart des conditions d'éclairage (HDW F900 et les copines...): HG 1 et 2 pour un signal sur 100%, HG 2 et 4 pour un signal sur

109% pour augmenter la dynamique si la chaîne post prod peut interpréter ces 9% supplémentaires (HG1= moins de lumière et HG2=fort contraste extérieur).

La philosophie « Courbes de réponse sur mesures », des courbes fabriquées pour des conditions particulières et des caméras qui acceptent l'injection de courbes (F900 par exemple).

Des courbes LOG à étalonner (Genesis (Panalog), F23 et F35 (S-LOG), HPX3700 P-LOG, etc...). C'est ce qu'on appelle du simili RAW, la totalité de la dynamique du capteur mais dans un workflow vidéo, sur 10 bits en LOG (l'image utile de 0 à 50% du signal linéaire, les hautes lumières de 50 à 100%), ou les nuances se retrouvent dans le bas de la courbe (le blanc standard à 100% se retrouve à 70%, et les visages passent de 70% à 40%).

Deux philosophies extrêmes:

Le format dit RAW (brut), où les données enregistrées proviennent directement du capteur sans aucun traitement, mais rendant la phase de post-production plus lourde: avant de voir une image 'droite', il faut traiter les données comme le font les autres caméras en interne, mais cette fois ci en post prod, depuis les sous-pixels R, G et B de la matrice, dé-matricer (Bayer par exemple) pour reconstruire une image en couleur dans un espace colorimétrique cible, réaliser une réduction de bruit, effectuer une balance des blancs, etc...: c'est la philosophie qui offre la plus grande marge de manœuvre à l'étalonnage; l'image ne peut être ni visionnée ni diffusée telle quelle en sortie de caméra (comme un négatif numérique qu'il faudrait 'révéler' avant de la regarder). Il y a obligatoirement une vie après le tournage...

Le mode 'News' ou 'droit' qui 'clip' l'image dans les hautes lumières et enterre les basses lumières en donnant une image rapide et très contrastée, avec comme seul réglage le diaphragme: C'est la philosophie qui offre le moins de manœuvre à l'étalonnage (ce qui est perdu au tournage est perdu pour toujours), mais l'image est prête à être diffusée telle quelle. Il n'y a pas ou peu de vie après le tournage...

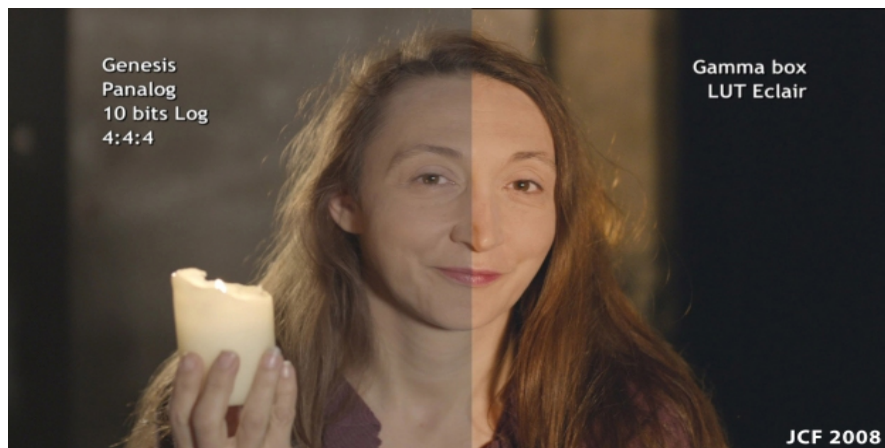
Entre ces deux extrêmes, chaque philosophie trouve sa place, et certaines caméras peuvent fonctionner dans plusieurs modes (linéaire, knee et gamma, courbes en preset, courbes persos, raw)...

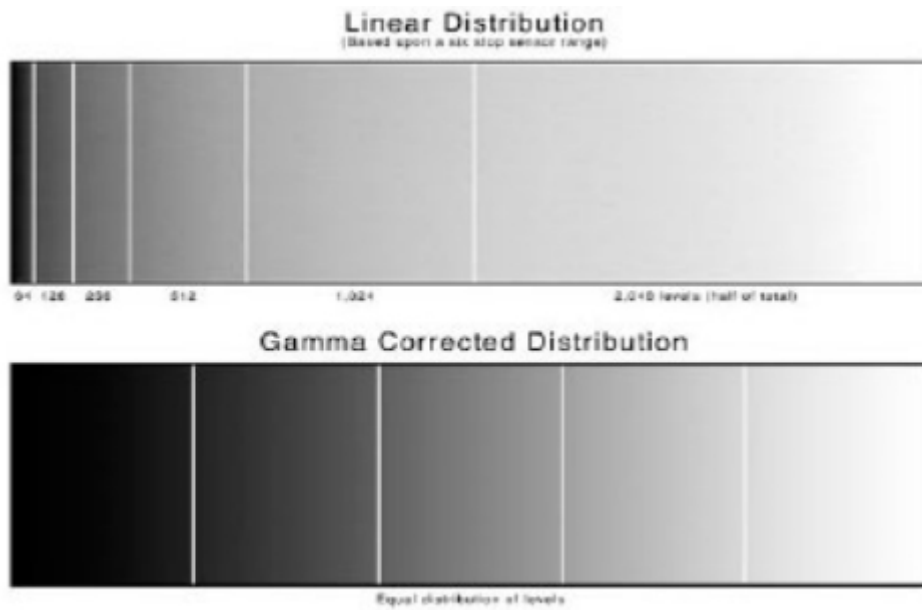
Avant l'enregistrement, des courbes de réponse adaptées (users gamma) sont utilisées dans la caméra; après l'enregistrement, des LUT (Look Up tables) 'révèlent' et interprète l'image finale sur un moniteur.



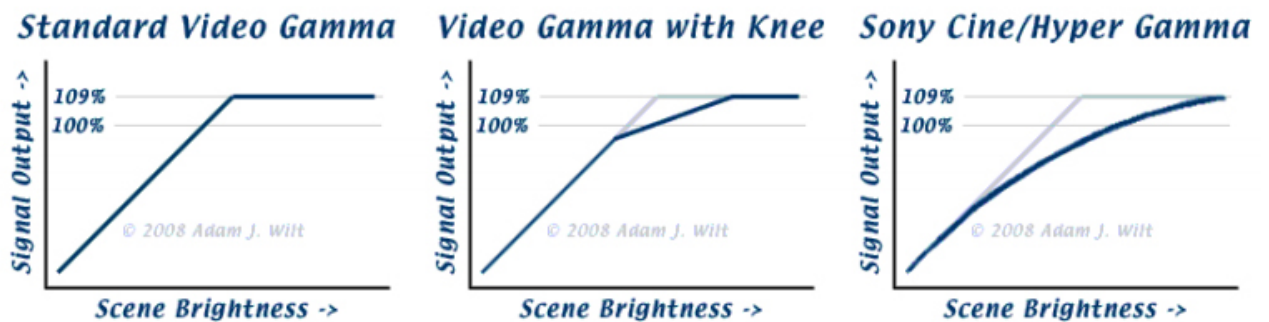


image révélée: application d'une courbe de Gamma standard HD ITU REC 709 (image corrigée sur Red Alert – gratuit)





Cette image illustre bien la recherche que l'on a effectuée afin de palier à la dynamique du gamma linéaire : Les noirs ont été amplifiés et les blancs compressés



Les premières caméras HD proposaient des image très « Vidéo », les constructeurs ont ensuite développé des courbe de réponse spécifiques, corrigeant les transitions brutales sur les hautes lumières, aboutissant à une image plus proche du rendu argentiques

En résumé, les différents workflows possibles :

