

L'histoire de la production vidéo expliquée : Partie 2

Entrelacement

L'image d'un téléviseur est une succession de balayages horizontaux, de gauche à droite, partant du haut, et finissant en bas de l'écran. Au commencement de la télévision, la qualité des éléments phosphorescents du tube est fort médiocre. De ce fait, quand le faisceau balaye le bas de l'écran, le haut a déjà disparu, d'où un phénomène de scintillement, ressenti fortement par l'œil humain pour 25 Hz ou 30 Hz. La solution la plus simple eût été d'accélérer la cadence de balayage, mais ceci imposait également d'augmenter la cadence des images, ce qui était inutile d'un point de vue cinématographique (le mouvement est perçu de la même façon), et fort coûteux en matériel et en bande passante. Une solution plus astucieuse fut de doubler la cadence de balayage, en omettant une ligne sur deux, afin de garder une quantité d'information constante. Ainsi, une première passe affiche toutes les lignes impaires en deux fois moins de temps que pour une image entière et une seconde passe affiche les lignes manquantes paires : c'est ce que l'on appelle l'entrelacement. On obtient bien le même nombre de lignes de balayages pour une image, et on balaye deux fois l'écran pour afficher une seule image. On désigne par le terme « trame » ("field" en anglais) une passe de balayage. Une image est donc constituée de deux trames, puisqu'il faut deux balayages pour définir l'image ("frame" en anglais).

Les caméras, qui fonctionnent comme un « téléviseur inversé », adoptèrent elles aussi cet entrelacement du balayage. Dans la première moitié du temps d'une image, une 1^{re} prise de vue définit toutes les lignes impaires, et une moitié d'image plus tard, une seconde prise de vue définit les lignes paires. Ce qu'il faut bien comprendre ici, c'est que les deux prises de vues sont distantes dans le temps (d'une moitié d'image). Et même si ces deux prises de vue sont complémentaires d'un point de vue spatial (les deux balayages se complètent dans le cadre), ces deux prises de vue n'affichent pas le même contenu ! Si un sujet se déplace dans le champ, il aura une position différente sur chacune des deux trames : on a alors un effet de zig-zag sur chaque frame.

Ce problème est en partie résolu par un dispositif de lames cristallines biréfringentes qui "étaient" les détails en dédoublant les rayons lumineux. Il en découle une perte de définition qui confèrent aux systèmes PAL et SECAM une résolution verticale multipliée par 0,7 (facteur de Kell) et qui n'est plus réellement que de 400 lignes environ.

Du fait de la capture en deux trames de 1/50 de seconde chacune, c'est le temps de pose en vidéo (25i).

Il existe dorénavant de plus en plus d'appareils vidéo capables d'afficher 25, 50 ou 60 images complètes par seconde, l'affichage n'est plus entrelacé, on parle alors de balayage progressif. Parmi les appareils capables d'un tel affichage on trouve : les ordinateurs (leur carte vidéo et leur écran), certains vidéoprojecteurs, les téléviseurs haut de gamme, certaines platines DVD et quelques rares caméscopes. C'est le mode de capture choisi pour les films tournés en HDTV ou en D-cinéma destinés à être transférés et projetés en 35 mm.

Les 25 images progressives (25p) confèrent alors aux caméras un temps de pose de 1/25 de seconde ce qui est trop long en terme de résolution temporelle. On préfère alors limiter le temps d'intégration des trames à 1/50 s (obturateur électronique).

Capture de l'image

Les premières caméras vidéo, fonctionnant sur le même principe que les téléviseurs, analysaient l'image formée par l'objectif à l'aide d'un tube cathodique. Depuis la fin des années 1980, elles sont dotées de capteurs Charge-Coupled Device : CCD ou Dispositif à Transfert de Charges (DTC) en français.

Le transfert de ces charges peut se faire de 3 manières différentes : transfert interligne (capteur IT : Inteline Transfer), transfert trame (capteur FT : Frame Transfer) qui nécessite un obturateur mécanique et est rarement utilisé ou transfert FIT (Frame Interline Transfer).

Au début du XXI^e siècle, les fabricants de capteurs ont décidé d'abandonner cette technologie et construisent désormais des capteurs CMOS (Complementary Metal Oxide Semi-conductor). On trouve cependant encore sur le marché des caméscopes semi-professionnels utilisant la technologie dite "tri-CCD" qui permet d'améliorer notablement le traitement des couleurs.

Résolution de l'image et fréquence de balayage

Il existe différents formats d'image vidéo, qui dépendent essentiellement de la fréquence de balayage vertical de l'image :

- * 405 lignes 50 Hz (standard anglais abandonné) noir et blanc
- * 525 lignes 60 Hz : résolution 4/3 utile = 711 x 480 (standard américain) couleur NTSC et PAL-N
- * 625 lignes 50 Hz : résolution 4/3 utile = 702 x 576 (standard européen) couleur PAL, SECAM et NTSC-4.43
- * 819 lignes 50 Hz : résolution 4/3 utile = 1024 x 768 (standard français abandonné) noir et blanc

On peut constater à ce point qu'il existe une différence entre le nombre de lignes composant l'image et le nombre de lignes affichées. Ceci représente une différence de 49 lignes en 50 Hz et de 45 lignes en 60 Hz. Ces lignes perdues sont nécessaires, elles représentent le temps nécessaire pour que le faisceau d'électrons balayant le tube cathodique puisse remonter du bas de l'image vers le haut. Ce problème technique n'existe pas avec les panneaux LCD et les dalles plasma, mais il est conservé pour assurer la compatibilité. Les lignes libres sont mises partiellement à profit : on y place les signaux du télétexte, du sous-titrage et aussi le time-code des équipements vidéo professionnels.

Il faut distinguer deux fréquences de balayage de l'image :

- * Le balayage vertical, qui s'effectue de haut en bas et sert à composer l'image. Il s'effectue 50 ou 60 fois par seconde.
- * Le balayage horizontal, qui s'effectue de droite à gauche pour chaque ligne de l'image. La fréquence de balayage horizontal est donc égale à la fréquence verticale multipliée par le nombre de lignes et divisée par deux à cause de l'entrelacement.

$$F_h = \frac{F_v N}{2}$$

Ce qui donne les valeurs suivantes :

- * $F_h(50\text{Hz}) = 50 \times 625 / 2 = 15625 \text{ Hz}$
- * $F_h(60\text{Hz}) = 60 \times 525 / 2 = 15750 \text{ Hz}$

Ce résultat n'est pas du au hasard. Si les fréquences horizontales sont presque les mêmes en 50Hz et en 60 Hz, c'est que cela permet d'utiliser la même circuiterie de balayage horizontal, donc de réaliser des économies.

Couleur

Depuis quelques décennies on connaissait les particularités spectrales de l'œil humain, qui affichaient une très nette préférence pour certaines couleurs. De plus on savait que le spectre chromatique de l'œil peut se décomposer en trois couleurs primaires, qui permettent par mélange de recréer à peu près toutes les autres couleurs du spectre. Le cinéma couleur exploitait ceci en utilisant des émulsions à plusieurs couches, dont chacune était sensible à une couleur donnée.

Les ingénieurs vidéo optèrent pour trois couleurs bien particulières : rouge vert bleu. Ces couleurs sont dites primaires (à ne pas confondre avec les couleurs primaires du domaine de la peinture qui sont le bleu, le rouge et le jaune) car ce sont elles qui, par mélange, vont permettre de recomposer un spectre entier de couleurs.

La prise de vue en couleur s'effectue selon un prisme optique qui répartit la lumière sur trois capteurs, devant lesquels on a respectivement un filtre rouge, vert et bleu. Ainsi, chaque capteur n'enregistre que les informations de lumière concernant sa couleur. Il suffit ensuite d'enregistrer puis restituer les 3 composantes RVB (RGB en anglais) sur un moniteur couleur acceptant les trois entrées RVB. Il faut bien comprendre que l'on obtient trois signaux à la place d'un seul. Il faut non seulement tripler toutes les liaisons câblées entre les différents équipements, mais aussi tripler les pistes d'enregistrement sur un magnétoscope, tripler tous les équipements de production, jusqu'aux équipements de diffusion hertzienne... Le défi était donc de créer un signal unique englobant trois informations différentes, et qui ne devaient pas se mélanger avant le traitement par le poste de réception.

Le défi était aussi de conserver la totale compatibilité avec les postes noir et blanc encore très présents dans les foyers. On travailla donc dans le but de créer un signal vidéo englobant : du rouge, du vert, du bleu, et du noir et blanc dans le même tuyau, sans que ceux-ci se mélangent.

Pour commencer, il était impensable d'avoir une caméra noir et blanc ET une caméra couleur. Il fallait donc fabriquer du noir et blanc à partir des trois composantes RVB. Se basant sur les sensibilités de l'œil aux différentes couleurs, on prit 59 % de vert, 30 % de rouge, et 11 % de bleu qu'on mélangea copieusement. On venait d'inventer un nouveau terme : la luminance (Y). Les télé noir et blanc pourraient donc voir en noir et blanc des images issues de caméra couleur. Comment maintenant rajouter à ce Y les informations de couleurs nous permettant de retrouver notre RVB original ? Puisqu'on avait déjà la lumière de notre image (le Y), il fallait « colorier » ce noir et blanc avec des informations de couleurs qui ne contenaient elles, aucune valeur de lumière, mais uniquement des indications de teinte et de saturation.

Une fois d'accord pour ce noir et blanc colorisé, il fallut trouver l'astuce qui permettrait de transmettre la lumière (Y) et la chroma (que nous appellerons C pour faire simple). Des procédés électroniques aux noms aussi effrayants que « modulation d'amplitude en quadrature de phase, à sous-porteuse supprimée » virent le jour. Ces solutions se devaient à la fois de mixer 2 signaux de manière à pouvoir les discriminer à la réception, mais aussi de n'avoir aucune interférence visible dans le spectre du signal noir et blanc.

Ces solutions furent trouvées et appliquées. Ainsi sont nés le NTSC (National Television System Committee) aux États-Unis, le SECAM (SÉquentiel Couleur À Mémoire) en France, et le PAL (Phase Alternate Line) en Allemagne. La technique employée pour transformer du RVB en signal couleur compatible noir et blanc s'appelle le codage. Le NTSC, le SECAM et le PAL sont trois types de codages différents, et bien entendu, incompatibles entre eux. Passer d'un type de codage à un autre s'appelle transcodage.

Aucune des trois solutions n'est néanmoins transparente, loin s'en faut. Un signal codé souffre d'artefacts plus ou moins visibles selon le codage.

Un signal vidéo codé de la sorte est dit signal composite, car il contient plusieurs sources de nature différente. Les standards vidéo utilisant le composite vont de l'U-MATIC / U-MATIC SP au VHS en passant par le 8mm ou Video 8, le Betamax, le VCR ou encore le V2000. Au vu des dégradations causées par le codage, il devenait urgent de s'en absoudre en production.

Au début des années 80, SONY mit au point un format vidéo à composantes séparées, constitué de plusieurs signaux distincts, véhiculés par des câbles distincts : le Betacam / Betacam SP. Pour rester compatible N&B, on évita soigneusement le RVB, et on choisit naturellement un format comportant le fameux Y (signal N&B), plus des informations de chrominance véhiculées par 2 signaux : U & V (appelés aussi Cr et Cb). Pour ceux qui n'auraient pas encore décroché, le $U = R - Y$, le $V = B - Y$, où $Y = 0,30R + 0,59V + 0,11B$ (les coefficients étant différents selon le codage utilisé). Cette transformation de RVB en YUV s'appelle Matriçage. Contrairement au codage, le matriçage est une opération très simple, qui ne génère pas de dégradation, tout en offrant l'avantage de la compatibilité Y.

Quelques années plus tard, on vit apparaître un format grand public dit S-Video ou Y/C, où la luminance Y et la chrominance C (codée en NTSC, PAL ou SECAM) étaient séparées (S-VHS, Hi-8, Super-Betamax). Ce format est de qualité meilleure qu'un format composite, puisque la chrominance n'empiète plus sur la bande de fréquences de la luminance, ce qui pouvait amener à des artefacts colorés sur des détails fins. La résolution horizontale de ces formats pouvait donc être quasiment doublée (400 points/ligne au lieu de 240-250).

La vidéo numérique

Introduction

L'histoire du numérique dans la vidéo commence véritablement de 1972 à 1982. À l'origine équipements de synchronisation, les appareils se sophistiquèrent avant d'entrer dans le milieu professionnel. Dès lors, les industriels prirent conscience de l'avènement de ce nouveau phénomène et présentèrent des normes en matière de numérisation. Une certaine anarchie numérique régna alors sur le marché ce qui força la main au CCIR (Comité consultatif international de radiodiffusion) à normaliser un format vidéo en composantes numériques compatible dans le monde entier : cette norme c'est le 4:2:2, ou CCIR 601. Elle spécifie les paramètres de codage de signaux à numériser (échantillonnage, quantification...) Dès lors les innovations ne cessèrent de s'enchaîner pour permettre aujourd'hui, à la vidéo numérique, de se généraliser dans les centres de production, chaînes TV et régie de post-production pour assister le montage vidéo.