

5.1 Le signal vidéo analogique

Une image vidéo analogique en couleurs peut être reconstituée par le mélange, à quantité variable, des trois couleurs primaires rouge, vert, bleu. Mais en pratique, le signal vidéo ne se présente pas toujours sous la forme de trois signaux RVB. Il peut en fait revêtir quatre formes différentes.

- RVB : chaque couleur primaire rouge, verte, bleue est véhiculée par trois signaux distincts sur trois liaisons indépendantes (une quatrième voie peut être utilisée pour la synchronisation). La qualité de l'image est optimale, mais aucune compatibilité n'est offerte avec les téléviseurs noir et blanc. De plus, la bande passante requise est triplée par rapport à celle du seul signal noir et blanc, ce qui est totalement prohibitif pour l'enregistrement ou la diffusion. La liaison RVB est essentiellement exploitée en studio pour connecter entre eux certains équipements vidéo, mais également à l'intérieur des outils de traitement de l'image, ainsi qu'en informatique.
- Composantes : les signaux en composantes sont fabriqués à partir d'une recombinaison linéaire des signaux primaires RVB, conduisant à un signal de luminance Y (représentant à lui seul l'image en noir et blanc), et deux signaux dits de différence de couleurs R-Y et B-Y, notés Dr et Db. Les pertes en qualité sont assez minimes et la bande passante requise est beaucoup moins importante qu'en RVB car les signaux Dr et Db, portant uniquement la couleur, sont peu encombrants (§ 5.1.1.1). Cette représentation du signal vidéo en composantes est née avec le format Betacam de Sony au début des années 80. Elle a depuis servi de base à la norme numérique 4:2:2, ainsi qu'à la grande majorité des formats d'enregistrement vidéo.
- Composite : les trois signaux composantes Y, Dr, Db sont combinés en un unique signal vidéo appelé signal composite. Il existe trois standards composites, le PAL, le SECAM et le NTSC, dont les différences portent essentiellement sur la manière dont sont combinés les signaux. Le signal composite

offre une qualité d'image bien inférieure à celle des formats RVB et composantes (surtout dans la représentation des détails), mais il présente l'avantage indéniable d'être transmissible sur une seule voie. Il est utilisé en diffusion hertzienne et a été repris seulement par deux formats numériques, le D2 et le D3.

- Y/C (ou composite séparé) : il s'agit d'un mode de représentation dans lequel la luminance et la chrominance sont codées chacune comme en composite, mais restent séparées l'une de l'autre au niveau de la connectique et des équipements. La qualité de l'image est meilleure qu'en composite (définition supérieure, pas de cross-color,...), les pertes se situant uniquement au niveau de la chrominance. La liaison Y/C est utilisée par les formats grand public S-VHS et Hi8, ainsi que sur certaines cartes de numérisation pour micro-ordinateurs.

Le signal vidéo analogique peut se présenter sous quatre formes – RVB, Y/C, composantes, composite –, dont seules les deux dernières ont servi de base au développement des formats numériques.

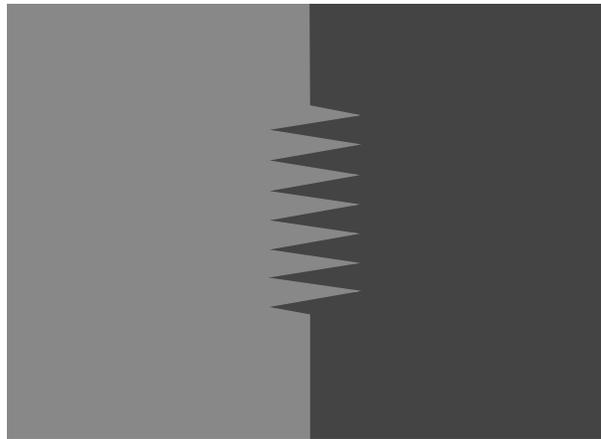
5.1.1. *Le signal composite*

Force est de constater que les normes de codage composite des signaux de télévision en vigueur aujourd'hui reposent sur des spécifications établies il y a plus de 50 ans, en fonction de la technologie et des besoins de l'époque. En effet, les standards PAL, SECAM et NTSC sont tous trois nés de la nécessité d'assurer le passage de la télévision en noir et blanc à la télévision en couleurs de façon totalement compatible. Il était impératif que les nouvelles émissions en couleurs soient restituées correctement par les récepteurs en noir et blanc dont étaient déjà équipés les foyers. Le problème qui s'est alors posé était de savoir comment transmettre les signaux de différence couleurs D_r et D_b dans un canal initialement prévu pour véhiculer un unique signal de luminance, le tout sans perturber la réception monochrome. Deux procédés ont été mis en œuvre pour y parvenir : le premier consiste à réduire la bande passante des signaux de chrominance, et le second à imbriquer leur spectre dans celui de la luminance.

5.1.1.1. Réduction de la bande passante de la chrominance

En 1949, l'américain Bedford montre que l'acuité visuelle de l'œil humain – système optique à une seule lentille non corrigé des aberrations chromatiques – est plus faible pour les détails colorés que pour les détails en noir et blanc. L'une des mires utilisées au cours de ses expériences est représentée sur la figure 5.1. Elle est composée de deux zones qui sont soit noire et blanche, soit de deux couleurs primaires d'égale luminance. Plus l'observateur s'éloigne de cette mire, moins il distingue de façon nette les zigzag de la transition centrale, qui finit par être perçue comme un flou vertical. Par rapport à la mire en noir et blanc prise en référence, l'acuité visuelle pour les mires de couleurs est de 40 % pour la mire rouge/vert, 23 % pour la mire rouge/bleu, et 19 % pour la mire vert/bleu.

Figure 5.1
Un exemple de mire utilisée au cours des expériences de Bedford mettant en évidence la faible acuité visuelle de l'œil humain pour les détails des couleurs. Les deux parties de la mire sont soit noire/blanche, soit rouge/verte, soit rouge/bleue, soit verte/bleue.



Cette constatation, d'apparence mineure, a eu des répercussions directes sur l'établissement de tous les standards de télévision en couleurs dans le monde entier. Elle a en effet permis de réduire la largeur de canal utilisée pour transmettre les composantes de couleur, sans que la qualité subjective de l'image observée ne soit altérée (si le signal de luminance est transmis en pleine bande). La faiblesse du pouvoir de résolution du système visuel humain face aux détails de couleurs a ainsi justifié une réduction

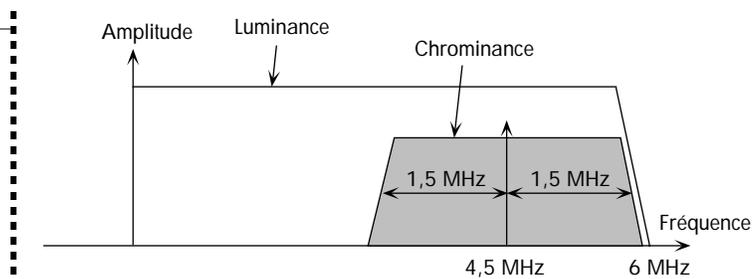
de la bande passante attribuée aux composantes de couleur dans un facteur quatre en diffusion, et dans un facteur deux en production/postproduction (par rapport à la luminance). Dans les systèmes à 625 lignes, la bande passante de la luminance a été fixée à 5,5 MHz pour la diffusion et celle de la chrominance à 1,5 MHz (ces valeurs sont respectivement de 4,2 et 1,4 MHz dans les systèmes à 525 lignes).

L'œil étant moins sensible aux détails colorés qu'aux détails en noir et blanc, la quantité d'information utilisée pour transmettre la chrominance peut être réduite à 25 ou 50 % de celle utilisée pour transmettre la luminance.

5.1.1.2. Imbrication du spectre de la chrominance dans celui de la luminance

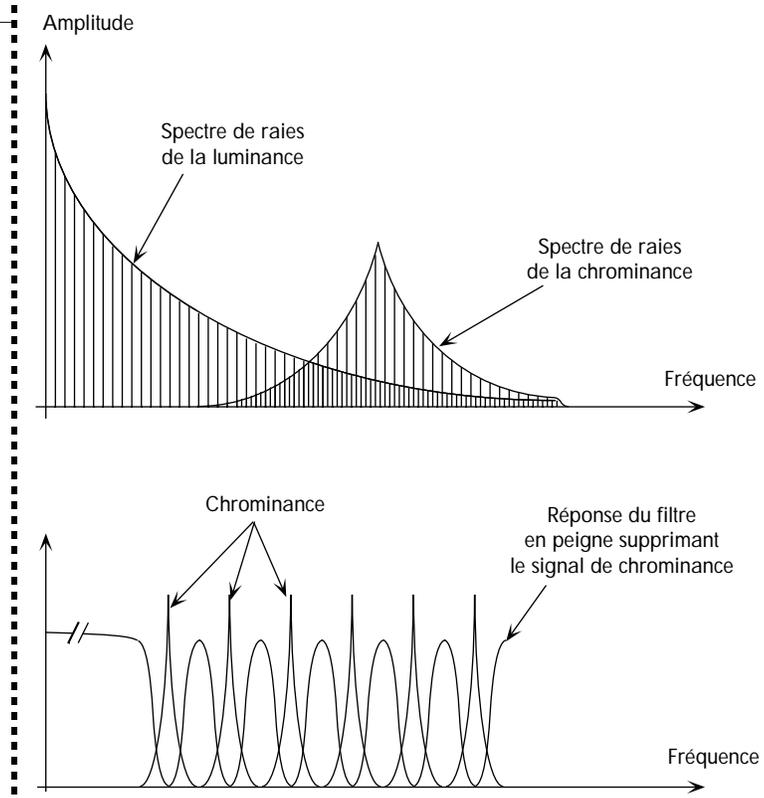
L'analyse séquentielle de l'image de télévision par une succession de lignes se traduit, pour la luminance comme pour la chrominance, par un spectre sous forme de raies ; l'énergie n'est transportée qu'à des fréquences multiples de la fréquence ligne. Cette particularité a permis d'imbriquer les raies du spectre de la chrominance entre celles du spectre de la luminance, et d'inclure ainsi les informations de couleur dans la largeur de canal occupée par la luminance. Dans un signal composite, les signaux de chrominance modulent une sous-porteuse intermédiaire dont la fréquence est choisie dans le haut du spectre de la luminance, partie la moins encombrée car correspondant aux détails fins de l'image.

Figure 5.2
Spectre du signal vidéo composite.



À la réception, la séparation des signaux de luminance et de chrominance devrait théoriquement être réalisée par un filtre « en peigne », comme l'illustre la figure 5.3. Mais elle est, dans la plupart des cas, effectuée par un brutal filtrage passe-bas, plus simple et plus économique, mais qui présente deux inconvénients notoires. D'une part, il réduit de manière significative la bande passante de la luminance. D'autre part, il laisse certaines hautes fréquences de luminance dans la bande spectrale de la chrominance. Un phénomène bien connu de *cross-color* se produit alors quand les composantes de la luminance restées dans la zone de la chrominance sont d'amplitude suffisamment élevées pour être interprétées comme des informations de couleurs. Sur l'image, cela se traduit par l'apparition d'un moiré coloré sur les surfaces garnies de transitions fines.

Figure 5.3
La séparation des signaux de luminance et de chrominance, dont les spectres sont imbriqués, s'effectue de manière idéale par un filtrage « en peigne ».



Le signal unique résultant de cette imbrication des spectres est le signal composite, appelé ainsi parce qu'il est le fruit du mélange des informations de luminance et de chrominance, auxquelles s'ajoutent les signaux de synchronisation. Le mode de combinaison des deux signaux de différence de couleurs Dr et Db entre eux, ainsi que la manière dont ils sont ajoutés au signal de luminance, diffèrent selon les standards de codage NTSC, PAL et SECAM.

Le signal vidéo composite combine, selon l'un des trois standards de codage – NTSC, PAL, SECAM – le signal de luminance, les deux signaux de différence de couleurs, ainsi que les signaux de synchronisation. Le signal composite porte à lui seul toutes les informations nécessaires à la reproduction de l'image vidéo en couleurs. Le spectre du signal de chrominance est inséré dans le haut du spectre du signal de luminance. Un signal de télévision composite en couleurs occupe ainsi le même espace fréquentiel qu'un signal en noir et blanc.

5.1.1.3. Organisation du signal composite

Le signal vidéo est un signal électrique traduisant ligne par ligne les variations de luminosité d'une image en variations de tension proportionnelles. Un noir parfait donne une tension nulle de 0 Volt, tandis que le blanc est représenté par la valeur 0,7 Volt. Le signal vidéo se compose d'une partie dite « utile », contenant les informations visibles de l'image, et d'une partie dite de « suppression », renfermant des données auxiliaires (références de niveau, synchronisation,...).

Dans les systèmes de télévision à 625 lignes, chaque image est découpée en deux demi-images ou trames de 312,5 lignes, l'une composée des lignes paires, l'autre des lignes impaires. Sur ces 625 lignes, seules 576 sont dédiées à l'image ; on dit qu'elles sont actives. En effet, les 25 premières lignes de chaque trame ne contiennent pas d'informations visibles. Elles sont utilisées pour transporter soit des signaux de service – synchronisation trame, identification couleur –, soit des informations annexes comme le télétexte. Chaque ligne possède une durée totale de 64 μ s, mais les informations traduisant les variations de l'image – signal vision – ne sont présentes que pendant 52 μ s. Une ligne débute par une impulsion de synchronisation horizontale négative ($-0,3$ V), et

chaque trame commence par un signal de synchronisation vertical. Dans les systèmes à 525 lignes, on compte seulement 487 lignes actives et 19 lignes de suppression sur chaque trame.

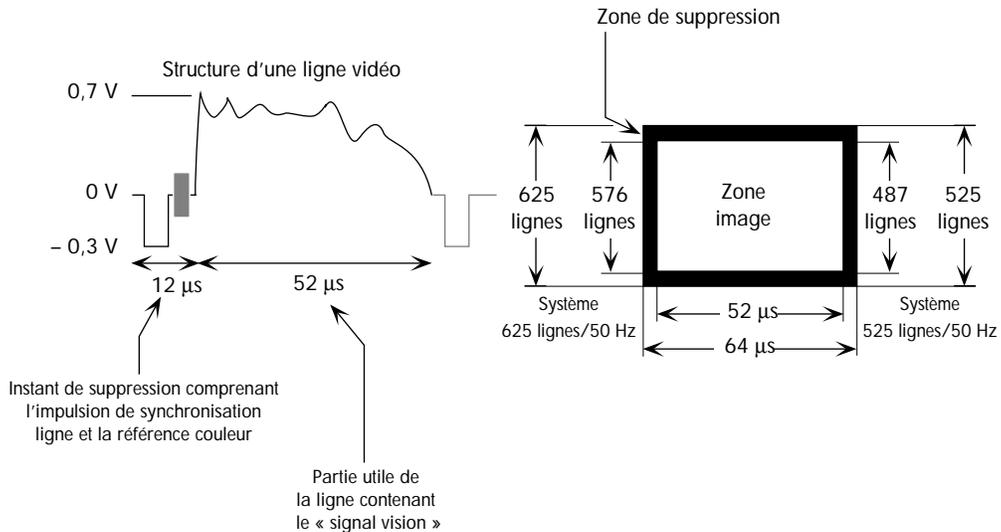


Figure 5.4
Les paramètres du signal vidéo composite.

5.1.1.4. La résolution de l'image vidéo

La résolution horizontale de l'image est directement liée à la largeur de la bande passante utilisée pour transporter ou enregistrer le signal vidéo. Elle est traditionnellement exprimée en « lignes TV » (ou plus simplement en lignes), et fait référence au nombre maximal de lignes verticales blanches et noires pouvant être distinctement perceptibles sur l'écran. Elle se calcule au moyen de l'équation suivante, prenant en compte le ratio 4/3 de l'image :

$$\text{résolution H (lignes TV)} = \frac{2 \times \text{bande passante} \times \text{durée ligne active}}{4/3}$$

Par exemple, la résolution horizontale en PAL/SECAM (bande passante : 5,5 MHz) est de 430 lignes, tandis que celle du VHS (bande passante : 3 MHz) n'est que de 240 lignes.

De manière générale, on retiendra qu'une bande passante de 1 MHz correspond à une résolution horizontale de 80 lignes.

La résolution verticale d'une image affichée avec un balayage entrelacé est égale au nombre de lignes multiplié par le « facteur de Kell », de valeur 0,7. Ce dernier indique le rapport entre la résolution verticale mathématique (576 lignes théoriques) et celle réellement obtenue. Ainsi, en 625/50, la résolution verticale de l'image vidéo est de 403 lignes ($576 \times 0,7$). Avec un balayage progressif, le facteur de Kell est égal à 1 ; la résolution verticale est directement égale au nombre de lignes visibles.

Nous allons à présent passer en revue les trois standards vidéo composites utilisés selon les zones géographiques, qui sont, par ordre de création, le NTSC (525 lignes par image, 60 trames par seconde), le SECAM et le PAL (625 lignes par image, 50 trames par seconde).

5.1.1.5. Le NTSC

Lancé en 1953, le NTSC (*National Television System Committee*) est utilisé principalement par les États-Unis, le Canada, le Mexique et le Japon. Tout comme le PAL mais contrairement au SECAM, le NTSC transmet simultanément les deux composantes de couleurs Dr et Db. Deux porteuses, de même fréquence mais décalées de 90° l'une par rapport à l'autre, sont modulées en amplitude par les signaux de différence de couleurs, puis supprimées. Les signaux résultants sont mélangés pour former un signal unique de chrominance modulé en amplitude (saturation) et en phase (teinte). Ce signal de chrominance est alors ajouté au signal de luminance. Comme deux phases sont possibles en ligne, la périodicité du signal NTSC est de quatre trames, alors que celle du signal en noir et blanc n'est que de deux trames. Pionnier des systèmes de télévision en couleurs et par conséquent tributaire de techniques peu évoluées, le NTSC est sujet à un défaut majeur : la moindre dis-

torsion de phase – due aux circuits du récepteur, à l'antenne, à la propagation, etc. – se traduit par des erreurs de teinte sur l'image. Ainsi, les récepteurs NTSC sont pourvus d'un potentiomètre de réglage de phase pour rattraper d'éventuels défauts. Il faut cependant savoir que les États-Unis vivent actuellement une période de transition vers la diffusion numérique hertzienne standard et à haute définition, au terme de laquelle le NTSC est appelé à disparaître définitivement. Les 1600 diffuseurs américains se sont vu prêter chacun par le gouvernement américain un second canal hertzien de 6 MHz, sur lequel ils sont tenus de diffuser un programme numérique (parallèlement à leur canal NTSC), conformément à un calendrier ayant démarré en 1998 et s'étalant jusqu'en décembre 2006. À l'issue de cette phase transitoire, tous les diffuseurs devront définitivement cesser de diffuser en NTSC et restituer leurs canaux au gouvernement qui les emploiera à d'autres fins.

5.1.1.6. Le PAL

Lancé en 1963, le système de codage PAL (*Phase Alternation Line*) apporte une variation au NTSC, d'apparence mineure, mais qui en supprime la principale faiblesse. Le PAL utilise le même type de modulation que le NTSC, mais avec la particularité d'inverser, à l'émission, la phase de la sous-porteuse relative au signal de différence de couleurs D_r une ligne sur deux, tout en conservant son amplitude (une salve de référence de la phase couleur est insérée sur le palier de suppression ligne). Ainsi, si une rotation de phase accidentelle survient sur une ligne donnée, une rotation de signe opposé se produit sur la ligne suivante. L'œil faisant une synthèse additive des couleurs, il restitue la bonne teinte, puisque les deux déphasages s'annulent mutuellement (on admet que l'information de couleur est quasiment identique sur deux lignes successives). L'alternance de phase sur la composante de différence de couleurs D_r d'une ligne à l'autre engendre une périodicité de quatre trames. Cependant, si l'on considère la fréquence de la sous-porteuse choisie pour le PAL, le nombre de périodes par ligne n'est pas un entier. Au total, la périodicité du signal PAL s'élève en fait à huit trames. En montage, le fait de briser un bloc de quatre images entraîne une saute de couleur provoquée par un déphasage de la sous-porteuse, visible surtout lorsqu'il s'agit de plan sur plan.

5.1.1.7. Le SECAM

Lancé en 1967, le SECAM repose sur la transmission en alternance une ligne sur deux de chacun des deux signaux différence de couleurs Dr et Db (solution par ailleurs récemment reprise par la structure de codage 4:2:0 de la diffusion numérique, du DVD et du format DV). La résolution verticale de la chrominance est ainsi réduite de moitié par rapport à celle de la luminance – ce qui n'est pas le cas des systèmes PAL et NTSC. Rappelons que la résolution horizontale de la chrominance est, de son côté, fortement diminuée en raison de la réduction de sa bande passante. Les inventeurs du SECAM ont estimé superflu de transmettre pour la chrominance une définition meilleure en vertical qu'en horizontal. Un seul signal de différence de couleurs est donc transmis par ligne en modulant en fréquence une sous-porteuse (une sous-porteuse différente est attribuée à chaque signal de différence de couleurs). Cette alternance Dr, Db à la fréquence ligne entraîne une périodicité de quatre trames. Cependant, l'inversion de la phase de la sous-porteuse (une ligne sur trois, une trame sur deux) choisie pour améliorer la compatibilité directe porte cette périodicité à douze trames. À la réception, le circuit de décodage traite, sur chaque ligne, le signal Dr ou Db transmis directement, additionné du signal Db ou Dr de la ligne précédente, conservé en mémoire et considéré comme encore valable pour la ligne actuelle (des signaux d'identification couleur sont transmis à chaque suppression trame et ligne). On s'est cependant assez vite rendu compte que si le SECAM offrait des avantages appréciables pour la transmission (pas de risque d'intermodulation des signaux de chrominance qui ne sont jamais présents simultanément, démodulation simple, etc.), il présentait en contrepartie des inconvénients irréfutables dans le domaine de la production. En effet, le mélange de deux sources SECAM impose de démoduler systématiquement les signaux en composantes Y, Dr, Db, afin de procéder à leur addition, puis de les remoduler en sortie de traitement. C'est pourquoi même dans les pays ayant adopté le SECAM comme norme de diffusion, ce dernier a souvent laissé la place au PAL en studio, qui permet le mélange des sources sans démodulation (si elles sont bien en

phase). C'est ainsi que les standards se sont progressivement divisés entre la production et la diffusion. L'Europe a été partagée entre le PAL et le SECAM, les chaînes de télévision ayant opté pour la diffusion en SECAM étant pour la plupart équipées aussi de matériels PAL.

D'autre part, avec le développement des travaux de postproduction est né le désir de repousser les limites des standards composites, dont la principale faiblesse est la limitation de la bande passante des signaux et l'imbrication de leurs spectres. Car si ce défaut est toléré pour la diffusion, il constitue un sérieux handicap pour les traitements complexes de l'image en régie. Au début des années 1980, l'avenir des normes composites dans les studios commence à s'assombrir.

5.1.2. *Les systèmes en composantes*

En 1983, Sony lance le format d'enregistrement Betacam et, avec lui, le premier magnétoscope capable d'enregistrer séparément les composantes de luminance et de différence de couleurs du signal vidéo. Le format Betacam crée une véritable révolution dans le domaine de l'acquisition en reportage. Il donne naissance aux premiers caméscopes monoblocs intégrant une caméra et un magnétoscope, permettant ainsi la réalisation de reportages vidéo dans d'excellentes conditions de mobilité.

La version améliorée du Betacam, le Betacam SP (1987), connaît un succès énorme dans les studios de production et de postproduction. L'enregistrement du signal vidéo en composantes Betacam SP offre des caractéristiques techniques supérieures au composite et est exempt des problèmes d'intermodulation parasite entre les informations de luminance et de chrominance. Les signaux en composantes analogiques font alors l'objet d'une normalisation. Peu à peu, les autres équipements de studio (mélangeurs, générateurs d'effets, ...) se mettent aux composantes analogiques, qui facilitent grandement les trucages et les incrustations en chromakey. Les caractéristiques originelles de l'image sont en effet conservées tout au long de ses multiples traitements, le codage en composite étant repoussé au

plus près des équipements de diffusion. Cependant, le traitement en composantes séparées, dont les avantages sont sans conteste, reste malgré tout assez limité de par sa nature analogique. Il n'autorise en effet que peu de traitements complexes au travers d'un mélangeur, et pas plus de trois ou quatre copies successives, au-delà desquelles l'image se dégrade. De plus, l'installation d'une régie de production en composantes analogiques n'est pas simple : il faut en effet tripler les chaînes de traitement vidéo au niveau du mélange et tripler également les câbles de liaisons entre équipements, grilles de commutation et distributeurs vidéo. Toutes ces raisons ont conduit à la numérisation des composantes vidéo et à leur codage sous la forme d'un unique signal : le signal 4:2:2.

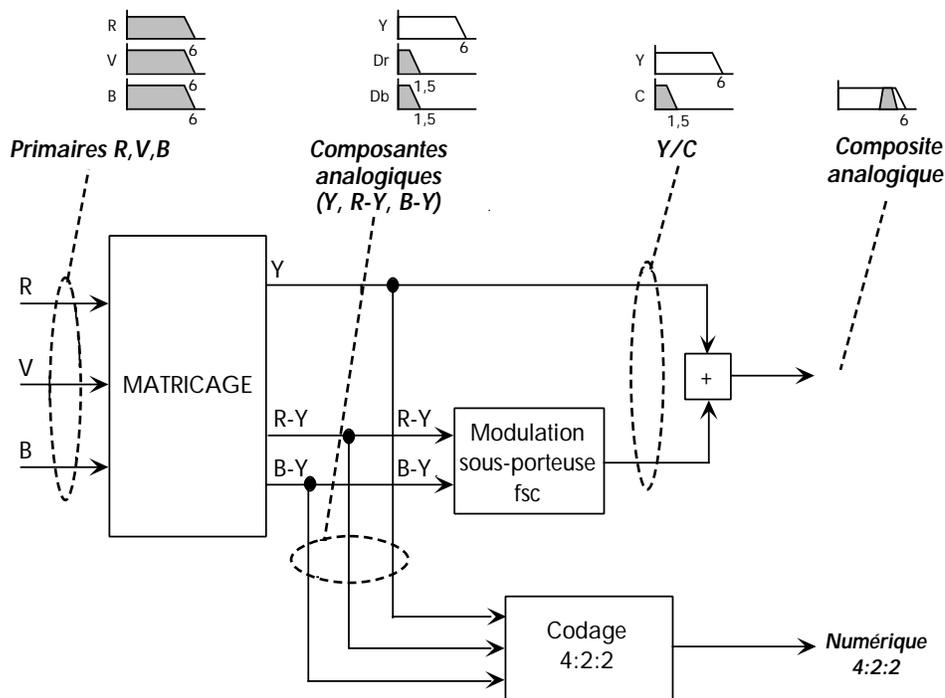


Figure 5.5
Les principaux formats du signal vidéo.

Un système en composantes séparées traite, dans des canaux indépendants, les informations de luminance et de différence de couleurs, supprimant ainsi tous les inconvénients liés au mélange de ces signaux par les systèmes composites. Mais il nécessite de tripler le nombre de câbles nécessaire à transporter un signal vidéo, ce qui alourdit son utilisation en studio.

5.2 Le signal vidéo numérique

5.2.1. *La genèse de la vidéo numérique*

De 1972 à 1982, le numérique est introduit progressivement dans le domaine de la vidéo, donnant naissance à des équipements n'ayant pas d'équivalent en analogique. Des unités de synchronisation – correcteurs de base de temps, synchroniseurs, etc. – de plus en plus évolués, des bibliothèques d'images et des générateurs d'effets font une percée croissante dans le monde broadcast. Industriels et diffuseurs prennent alors conscience du colossal enjeu qu'est en passe de représenter l'avènement de la télévision numérique. Seulement, chaque constructeur propose sa propre solution en matière de numérisation, ce qui fait régner sur le marché une certaine anarchie, et oblige de repasser en analogique dès qu'il s'agit d'interconnecter les appareils.

1982 est une année charnière dans l'histoire de la télévision. Le CCIR (Comité Consultatif International de Radiodiffusion) normalise en effet un format vidéo en composantes numériques compatible à l'échelle mondiale. Cette normalisation se réalise en fait en deux temps. La norme CCIR 601, communément appelée 4:2:2, spécifie les paramètres de codage des signaux de télévision numériques pour une utilisation en studio : signaux à numériser, échantillonnage et quantification. Elle sera ensuite complétée par la norme CCIR 656, qui décrit, quant à elle, les interfaces de liaison entre les différents équipements numériques. Les constructeurs de tous les pays s'efforcent dès lors de développer des équipements dont l'architecture interne est conforme à cette recommandation 601 du CCIR et dont les interfaces d'entrée/sortie répondent à la norme CCIR 656.

À partir de 1986 sont commercialisés les premiers formats de magnétoscopes numériques, concrétisant enfin le rêve de la multigénération sans perte. Les premières régies entièrement numériques de production et de postproduction se développent dès 1987. L'utilisation du numérique fait exploser la créativité en matière d'effets spéciaux, grâce aux étonnantes capacités des générateurs d'effets et stations de travail graphiques, mais aussi grâce aux possibilités de multigénération offertes par les magnétoscopes numériques.

À partir de 1991, le numérique se généralise dans les nouveaux centres de production, ainsi que dans les régies finales des chaînes de télévision. La norme numérique 4:2:2, qui s'accommode aussi bien des ratios d'image 4/3 que 16/9, fera par la suite l'objet de deux évolutions. La première permet une amélioration de la qualité de l'image – passage de 8 à 10 bits de quantification –, tandis que l'autre facilite la connectique en studio – sérialisation du signal – ; de tout cela il sera question plus loin dans ce chapitre.

La norme 4:2:2 est à l'origine d'équipements haut de gamme dont l'excellente qualité satisfait les besoins des utilisateurs les plus exigeants. Aujourd'hui, tout le matériel d'exploitation présent dans une régie existe sous forme numérique et compte déjà plusieurs générations.

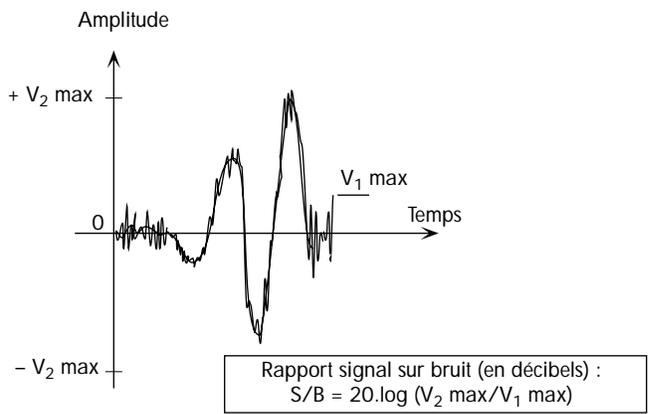
5.2.2. *Signal analogique et signal numérique : quelles différences ?*

5.2.2.1. Le signal analogique

Le signal analogique est celui qui représente le plus naturellement et le plus fidèlement les variations d'un phénomène physique. Ainsi, le microphone transforme en variations de tension les variations de pression acoustique qu'il capte. La caméra fournit, quant à elle, un signal électrique issu de l'analyse séquentielle des variations d'intensité lumineuse recueillies par l'objectif. Dans ces deux cas, une grandeur physique est traduite en un signal analogique dont l'amplitude instantanée porte l'informa-

tion. Une fois généré par les capteurs adéquats, le signal analogique doit être traité avant d'être diffusé et/ou stocké. C'est là que naissent tous les problèmes. Le signal analogique souffre d'un inconvénient majeur : il est très fragile. Son enregistrement, son traitement ou sa transmission lui font subir différents types de dégradations, qui altèrent rapidement sa qualité. Ainsi, la composante de bruit, qui vient inévitablement se superposer au signal utile, est directement liée au nombre et à la qualité des traitements appliqués. Elle est communément quantifiée par le rapport entre l'amplitude maximale du signal d'origine et celle du bruit qui lui est ajouté, appelé rapport signal sur bruit ou S/B (en vidéo, il varie de 40 dB en VHS à plus de 60 dB pour les caméras broadcast). Par ailleurs, la réponse amplitude/fréquence des circuits de traitement n'est jamais parfaitement linéaire ; il s'ensuit une distorsion du signal, donc une modification de l'information véhiculée. Tous ces phénomènes ne font que se multiplier lorsque sont cumulées des copies de bandes successives, copies dont la qualité s'affaiblit très vite avec le nombre de générations.

Figure 5.6
Le rapport signal/bruit d'un signal analogique s'exprime en fonction du rapport de l'amplitude maximale du signal d'origine sur celle du bruit qui lui est superposé.



Le signal analogique est un signal dont la valeur varie avec le temps de manière continue. Il est sujet à de la distorsion lors de son traitement et est souvent dégradé par une composante de bruit venant s'ajouter à lui au cours de son enregistrement ou de sa diffusion.

Le décibel

Le décibel (un dixième du Bel), de symbole dB, exprime, sur la base d'une échelle logarithmique, le rapport entre deux puissances, ou deux grandeurs liées à une puissance. Les puissances acoustiques que perçoit notre oreille, entre le seuil de l'audition (1 microwatt/m²) et celui de la douleur (un mégawatt/m²) peuvent prendre des valeurs ayant un rapport de plus de 10¹². Pour se simplifier la vie et éviter de manipuler des nombres aussi astronomiques, les acousticiens ont eu l'idée de recourir à cette écriture en « puissances de 10 », d'autant que la sensation auditive humaine croît avec le logarithme de l'excitation. De là est né le décibel (plus commode à utiliser que le Bel) défini par l'expression $10 \cdot \log (P2/P1)$, P1 et P2 étant deux puissances. L'échelle des décibels est telle que si l'on double la pression acoustique d'un son, on obtient un niveau sonore supérieur de 3dB.

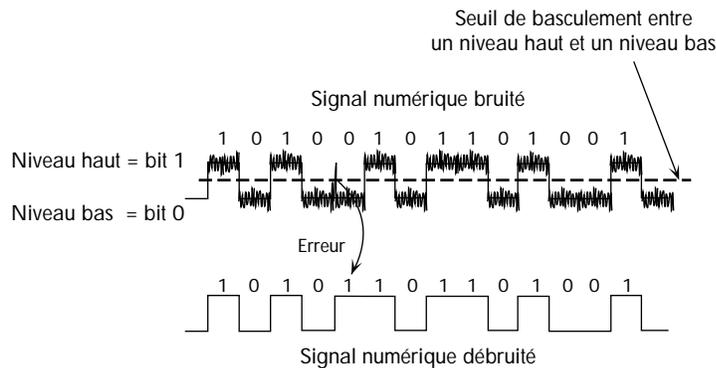
La notation en décibel a par ailleurs été étendue aux rapports de tension électrique. Elle permet alors d'exprimer le gain d'un amplificateur, les pertes de niveau d'une ligne de transmission, ou le rapport signal sur bruit caractérisant la qualité d'un signal. Les rapports de tension en décibel s'expriment cette fois par l'expression $20 \cdot \log V2/V1$ (l'équivalence en décibel d'un rapport en tension est le double de celui en puissance). Un rapport de 2 en tension est représenté par une augmentation de + 6dB. Par exemple, lorsque l'on dit d'un rapport signal sur bruit qu'il a augmenté de 6dB, cela signifie que le niveau du bruit a été réduit de moitié.

5.2.2.2. Le signal numérique

Le signal numérique n'a pas de nature physique. Il se présente sous la forme d'un message composé d'une suite de symboles, et est donc discontinu, le passage d'un symbole à un autre s'effectuant par une transition brutale. Basé sur le langage binaire, le signal numérique est constitué d'une suite de 0 et de 1, traduisant respectivement un niveau bas et un niveau haut du signal électrique. L'avantage fondamental du numérique réside dans le fait qu'il autorise un nombre de traitements complexes très élevé sans que soit affectée l'intégrité de l'information. Les dégradations causées par les perturbations du support de transmission ou d'enregistrement, telles que l'adjonction de bruit, la distorsion, etc., ne touchent que les paramètres analogiques du signal numérique et n'altèrent pas l'information, si elles restent dans certaines limites. En effet, tant que les niveaux électriques hauts et bas correspondant aux deux valeurs binaires se distinguent l'un de l'autre, le message transporté demeure intact. De puissants systèmes de détection et de correction d'erreurs ont en outre été développés pour optimiser la transparence du traitement en transmission ou en enregistrement. Par ailleurs, si le traitement d'un signal analogique s'effectue au moyen de composants qui en modifient les variations, le traitement d'un signal numérique se réalise par des calculs effectués à l'aide

de registres, additionneurs, et multiplicateurs, qui ne font que remplacer un message numérique par un autre. La numérisation du signal confère une plus grande robustesse à l'information et supprime les différents aléas et problèmes de non-linéarité inhérents aux circuits analogiques. En assurant une totale reproductibilité du signal après de multiples traitements et manipulations complexes, le numérique ouvre la voie à un vaste champ d'applications.

Figure 5.7
Les perturbations du support n'affectent pas l'information portée par le signal numérique tant que les niveaux haut (1) et bas (0) sont correctement reconnus.



Le signal numérique est un signal discontinu dans le temps. Il représente la valeur d'une grandeur physique à un instant donné. Le signal numérique est particulièrement stable. Il se prête parfaitement aux traitements les plus complexes et aux copies cumulatives au travers desquels l'information qu'il porte est totalement préservée.

Si les avantages du travail sur un signal numérique sont indéniables, il faut toutefois préciser que les phases de capture du phénomène physique et de restitution sont et resteront analogiques. Les capteurs photosensibles d'une caméra, comme les microphones, généreront toujours un signal analogique, et de l'autre côté de la chaîne, il n'est pas aujourd'hui envisageable de se passer d'écrans, d'enceintes, ou d'écouteurs. Dans une chaîne vidéo, le numérique intervient parfois au niveau de la création de l'image (ordinateurs, palettes graphiques, générateurs de caractères, etc.), et souvent au niveau du traitement ou du stockage d'un signal issu d'un équipement analogique et converti le plus tôt possible. C'est précisément cette étape de conversion qui s'est révélée comme l'une des grandes difficultés techniques à surmonter. Des circuits

capables d'effectuer la conversion de l'analogique vers le numérique existaient depuis longtemps, mais ils ne travaillaient pas à des fréquences suffisamment élevées pour traiter le signal vidéo.

5.2.3. *Principe de la conversion analogique/numérique*

On pourrait assimiler le processus de conversion analogique/numérique au passage de la tradition orale à la tradition écrite. Il nous est à tous arrivé d'écouter un discours ou un cours, et de prendre des notes pour conserver une trace du message donné. Cependant, il est difficile d'écouter et d'écrire en même temps, surtout si le débit de l'orateur est élevé. C'est pourquoi il nous faut échantillonner notre attention pour ne saisir et n'écrire en fait que des morceaux de message. Cette notion d'échantillonnage est fondamentale : il faut maintenir un certain temps de concentration pour capter un élément d'information avant de le retranscrire, puis écouter à nouveau, et ainsi de suite. Souvent, une notation par symboles courts et rapides à écrire – sténographie ou sténotypie – est indispensable pour optimiser la saisie de l'information.

La première étape de la conversion analogique/numérique est donc l'échantillonnage : l'amplitude du signal analogique est prélevée ponctuellement à des instants réguliers et suffisamment rapprochés. Les échantillons de tension ainsi récoltés décrivent la forme du signal point par point. Il faut ensuite remplacer leur valeur par un nombre entier de longueur fixe, codé en base 2 : c'est la seconde étape de la numérisation du signal, appelée quantification. Cette base 2 est très facile à manipuler à l'aide d'un signal électrique à deux états, puisque un « 1 » peut correspondre à une tension positive de 5 V, par exemple, et un « 0 », à une tension nulle. Chaque échantillon saisi est pesé – comme on pèse une denrée alimentaire à l'aide de poids pour déterminer son prix – afin que soit repéré, dans une table d'équivalence, le nombre binaire qui lui est le plus proche. Imaginons une balance à deux plateaux : sur un plateau, on place l'échantillon à peser et sur l'autre, des poids de plus en plus petits pour tenter d'approcher

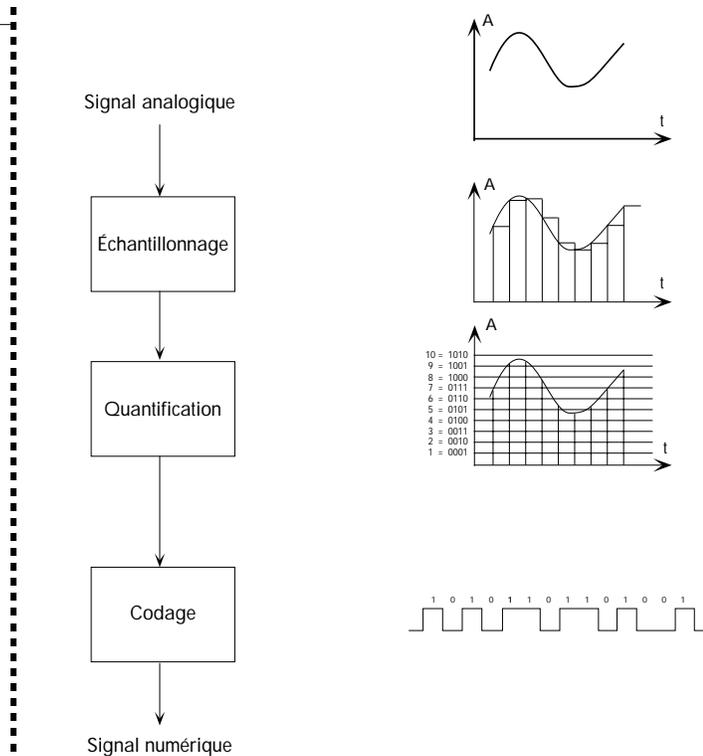
l'équilibre – sans l'obtenir réellement, sauf en cas de grande chance. La précision de la pesée dépend donc de la valeur du plus petit poids disponible. Le nombre binaire attribué à chaque échantillon est en effet forcément arrondi à ce plus petit poids. Cette approximation est source d'une erreur de quantification, ou bruit de quantification. L'opération de quantification prend un certain temps, dont dépendra au final la vitesse de conversion du système. Cette notion de vitesse de conversion est importante ; c'est par ailleurs le point qui a posé le plus de problèmes en vidéo, le débit d'information étant extrêmement important comparé à celui de l'audio, par exemple. Puis vient l'étape du codage, au cours de laquelle le flux de données est mis en forme en vue de son stockage ou de sa transmission. Par ailleurs, pour conférer aux informations une bonne immunité face aux perturbations amenées par le support d'enregistrement ou de transmission, il est nécessaire de leur adjoindre un certain nombre de données supplémentaires qui permettront, lors du décodage, de détecter et corriger les erreurs introduites (décodage d'un 1 pour un 0 et inversement). Ces données redondantes ajoutées aux données utiles ne font cependant qu'accroître la quantité déjà énorme d'informations à débiter par unité de temps, donc la bande passante nécessaire à transporter le signal numérique.

Le langage numérique : les bits et les octets

Le bit (contraction de *binary digit*) est la plus petite unité d'un système numérique ; sa valeur est « 1 » ou « 0 ». Un bit isolé peut représenter deux niveaux, ou états (*on/off*, blanc/noir,...), deux bits peuvent en représenter quatre (par exemple deux niveaux de gris entre le blanc et le noir), et plus généralement n bits peuvent en représenter 2^n . Ainsi, le codage d'une image vidéo sur 8 bits permet de distinguer 256 niveaux de gris, ou valeurs de couleurs, alors qu'un codage sur 10 bits permet d'en obtenir 1024. Un « mot » de 8 bits forme un octet (*byte* en anglais). Un mot peut cependant être composé de 10, 16, 24 ou 32 bits. Notons que chaque bit supplémentaire double le nombre de valeurs discrètes qu'il est possible d'obtenir. Dans un mot de n bits, le bit situé le plus à gauche est le plus significatif (*MSB : Most Significant Bit*), tandis que celui situé le plus à droite est le bit le moins significatif (*LSB = Least Significant Bit*). Dans l'exemple ci-dessous (« 10010111 » = 151), le premier 1 (à gauche) est le bit le plus précieux car il représente à lui seul la valeur 128, tandis que le dernier 1 (à droite) ne représente que la valeur 1.

<i>MSB</i>	1	0	0	1	0	1	1	1	<i>LSB</i>
	$\times 128$	$\times 64$	$\times 32$	$\times 16$	$\times 8$	$\times 4$	$\times 2$	$\times 1$	
	128	0	0	16	0	4	2	1	
	+ + + + + + + + + = 151								

Figure 5.8
Les principales phases
de la conversion
analogique/numérique.



Voilà donc exposé en quelques lignes le principe de la conversion analogique/numérique : échantillonnage, quantification, codage et adjonction de codes de correction d'erreurs en sont les étapes principales. Nous allons les détailler dans ce qui suit, en nous limitant au cas spécifique du signal vidéo.

5.2.4. Les signaux vidéo à numériser

Pour supprimer les différents systèmes de couleurs liés aux codages composites et favoriser l'échange international des programmes, il a été décidé de numériser des données communes à tous les pays. C'est donc logiquement la numérisation des composantes Y, Dr, Db qui a été retenue. Selon la convention adoptée, les composantes de couleur Dr, Db numérisées deviennent Cr, Cb.

5.2.5. L'échantillonnage

Figure 5.9
La succession d'échantillons d'amplitudes prélevés sur le signal analogique à des instants réguliers et suffisamment rapprochés permet de décrire ce signal point par point.

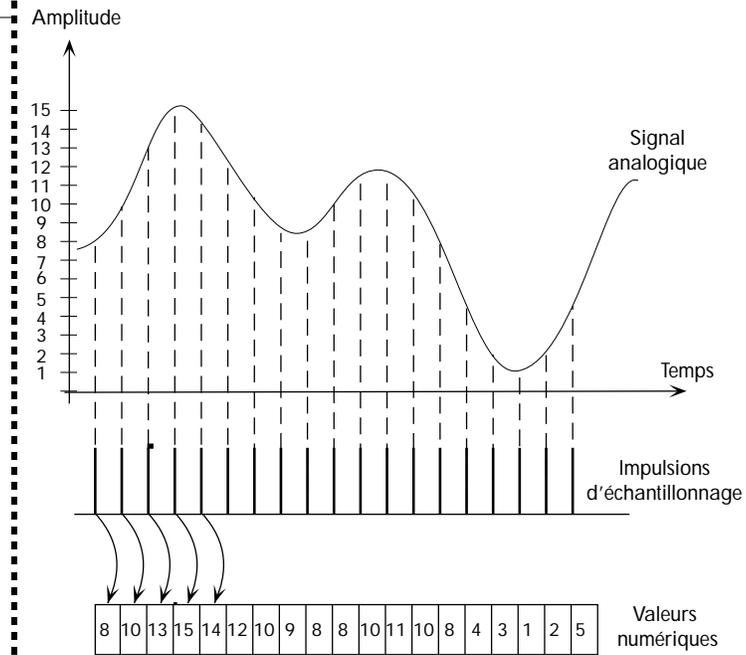
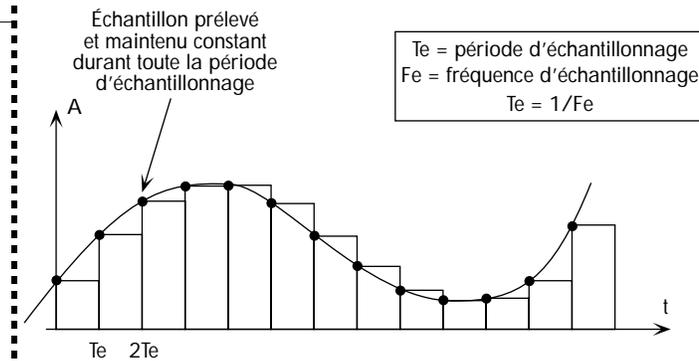


Figure 5.10
Principe de l'échantillonnage d'un signal. Un échantillon représente une amplitude instantanée discrète du signal.



5.2.5.1. La loi de Shannon et Nyquist

Échantillonner un signal variant de façon continue consiste, nous l'avons vu, à le découper en tranches temporelles et prélever

ponctuellement ses valeurs à des instants réguliers. On pourrait *a priori* penser que la succession d'échantillons ainsi générée est déficitaire en information par rapport au signal d'origine. En fait, si l'on respecte un ensemble de conditions concernant la bande passante du signal et l'espace maximal entre les échantillons, l'information portée par ces derniers sera une représentation intégrale du signal analogique. Tout doit être fait pour que le rythme de découpe soit suffisamment rapide pour être à l'affût de la variation la plus brève du signal d'origine. Car si la distance entre deux échantillons est supérieure à la plus rapide des variations du signal analogique, cette dernière passera inaperçue et sera définitivement ignorée. Il faut bien comprendre qu'entre deux échantillons, on ignore tout du signal d'origine. Autrement dit, pour échantillonner un signal en préservant toute son information, il est nécessaire de connaître au préalable la fréquence la plus élevée à laquelle il est susceptible de varier. On admet que cette fréquence maximale est de 6 MHz pour la vidéo et 20 kHz pour l'audio. Une loi mathématique, établie par Shannon et Nyquist, permet de déterminer la fréquence d'échantillonnage minimale à choisir pour ne pas manquer la plus petite (en termes de durée) des informations à saisir : « Un signal dont le spectre est limité à la fréquence F_{\max} est entièrement déterminé par la suite complète de ses échantillons prélevés à des intervalles de temps réguliers de valeur $T = 1/2.F_{\max}$. » Autrement dit, le signal échantillonné sera la représentation exacte du signal original si la fréquence d'échantillonnage F_e est au moins supérieure à deux fois la fréquence maximale du signal : $F_e \geq 2.F_{\max}$. Si ce critère n'est pas vérifié, les composantes spectrales répétitives du signal échantillonné ne sont pas assez espacées et se chevauchent, comme l'illustre la figure 5.13. La zone d'interférence, appelée zone de repliement ou d'aliasing, donne naissance à des fréquences aberrantes dans la bande utile. C'est pourquoi un filtre passe-bas, dit anti-repliement, ou anti-aliasing, à pente très abrupte, est impérativement requis en amont du processus de conversion. Son rôle est de rejeter toutes les fréquences indésirables du signal d'entrée qui sont supérieures à la moitié de la fréquence d'échantillonnage choisie. Il peut en effet exister dans le signal électrique des composantes à variation très rapide

(bruit, parasites, etc.), dont le contenu fréquentiel est plus élevé que la limite théorique. Ce filtre anti-aliasing doit être caractérisé par un gain constant dans la bande utile et par un affaiblissement rapide aux fréquences supérieures à la demi-fréquence d'échantillonnage.

Figure 5.11
Si la fréquence d'échantillonnage est trop faible, les variations rapides du signal original ne sont pas représentées par la suite d'échantillons prélevés.

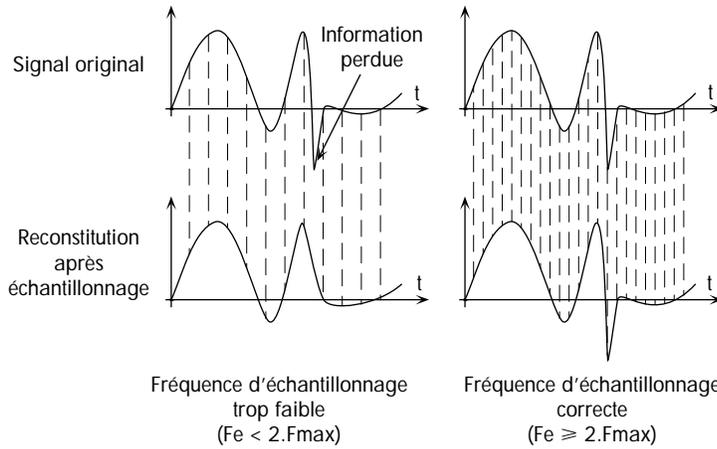
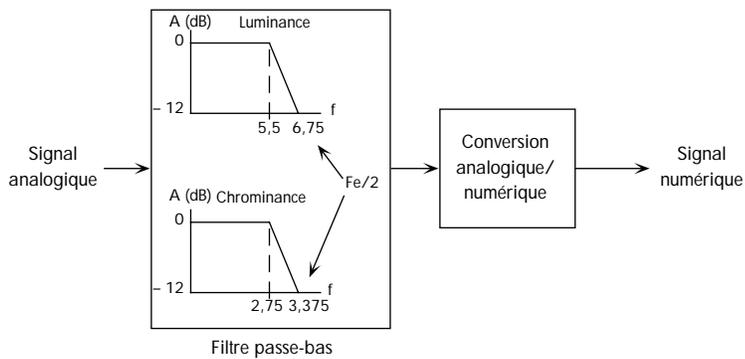


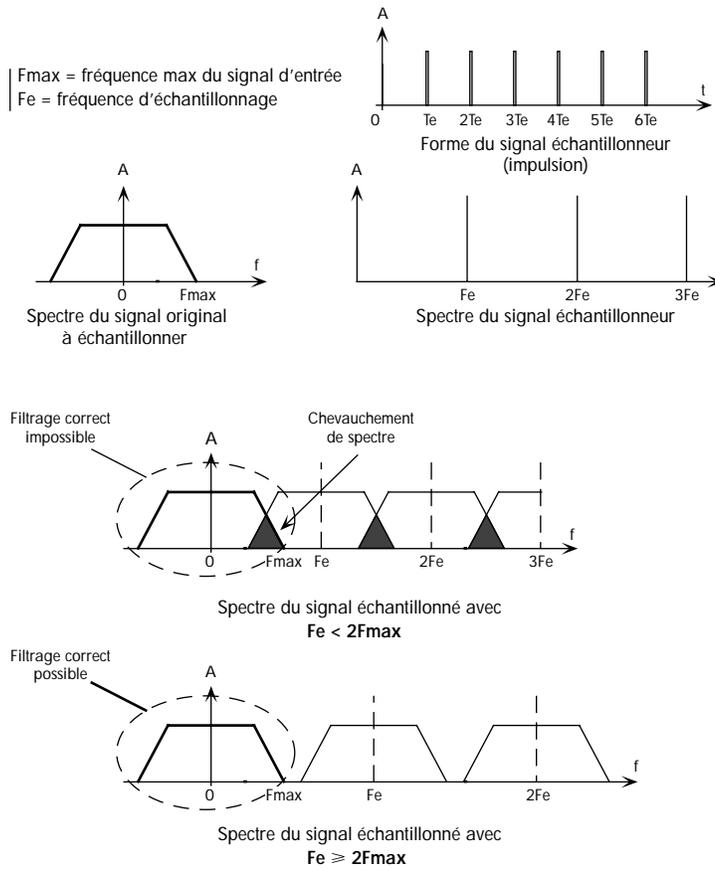
Figure 5.12
Le filtrage anti-repliement a pour rôle de supprimer, avant l'échantillonnage, les fréquences du signal vidéo supérieures à la moitié de la fréquence d'échantillonnage.



Précisons par ailleurs que le temps de calcul de l'échantillon est très faible, mais qu'il n'est pas nul. Il est donc indispensable que l'amplitude de l'échantillon à mesurer soit maintenue constante pendant le temps de conversion.

Figure 5.13

Le spectre du signal échantillonné est composé d'une succession de translations du spectre original à des fréquences multiples de la fréquence d'échantillonnage (F_e , $2F_e$, $3F_e$...). Pour que les répétitions spectrales soient disjointes, il faut que $F_e \geq 2F_{max}$. Le signal initial peut alors être reconstitué par un filtrage passe-bas supprimant, sur le signal échantillonné, toutes les fréquences supérieures à F_{max} .



Le processus d'échantillonnage consiste à prélever ponctuellement et régulièrement des mesures d'amplitudes du signal analogique. Conséquence de cette opération : le spectre du signal est dupliqué autour des multiples de la fréquence d'échantillonnage. C'est pourquoi la fréquence d'échantillonnage doit être au moins supérieure au double de la fréquence maximale du signal à convertir, afin d'éviter tout chevauchement de spectre. Si cette condition est vérifiée, l'échantillonnage n'a alors aucun rôle perturbateur.

5.2.5.2. Les fréquences d'échantillonnage du signal vidéo

• **La luminance**

Les fréquences d'échantillonnage des composantes du signal vidéo ont été choisies communes aux systèmes à 625 et 525 lignes, de manière à s'affranchir définitivement des problèmes d'incompatibilité entre les différentes zones géographiques du globe. L'information de luminance dispose d'une bande passante de 6 MHz. Celle-ci est nominalement plate jusqu'à 5,5 MHz, avec un affaiblissement d'au moins 12 dB à 6,75 MHz. La fréquence d'échantillonnage du signal vidéo doit donc, pour respecter le critère de Shannon-Nyquist, être au moins égale à 12 MHz. La compatibilité à l'échelle mondiale impose par ailleurs de choisir une valeur qui soit un multiple commun des fréquences lignes des systèmes à 625 et 525 lignes, soit respectivement 15 625 Hz et 15 734,25 Hz. Plusieurs essais ont été effectués avec des fréquences allant de 12 MHz à 14,3 MHz. Les paramètres étudiés ont principalement été la qualité de l'image avant et après traitement, le rapport qualité/coût, la capacité potentielle des magnétoscopes numériques, ainsi que la réduction du débit binaire. C'est ainsi qu'a été adoptée, comme norme mondiale, une fréquence d'échantillonnage du signal de luminance égale à :

$$F_e(Y) = 13,5 \text{ MHz.}$$

Cette valeur est égale à 864 fois la fréquence ligne des systèmes à 625 lignes et à 858 fois celle des systèmes à 525 lignes. Dans les deux cas, une ligne active numérique renferme 720 échantillons de luminance.

• **La chrominance**

Les signaux de différence de couleurs ont une bande passante nominalement plate jusqu'à 2,75 MHz, avec un affaiblissement d'au moins 12 dB à 3,375 MHz. Ils sont échantillonnés à une fréquence deux fois plus faible que le signal de luminance :

$$F_e(\text{Cr}) = F_e(\text{Cb}) = 6,75 \text{ MHz.}$$

Cette valeur est égale à 432 fois la fréquence ligne des systèmes à 625 lignes et à 429 fois celle des systèmes à 525 lignes.

Sur une ligne, il y a donc deux fois moins d'échantillons de chrominance que de luminance, soit 360 par ligne active. Compte tenu des filtres réalisables, les bandes passantes équivalentes sont de 5,75 MHz pour Y et de 2,75 MHz pour Cr, Cb ; elles sont plus élevées que celles fixées des standards analogiques composites et composantes (surtout en chrominance).

5.2.6. *La norme 4:2:2*

La norme de codage numérique destinée aux applications de studio a été baptisée « 4:2:2 ». Ces nombres représentent le rapport entre les fréquences d'échantillonnage du signal de luminance (le « 4 ») et des deux signaux de différence de couleur (les « 2 »). Le choix du « 4 » remonte à l'origine des travaux sur la numérisation du signal vidéo, quand la première solution envisagée reposait sur la numérisation du signal composite. La fréquence d'échantillonnage de la luminance avait été choisie égale à 4 fois la fréquence de la sous-porteuse du signal PAL ou NTSC, soit respectivement 17,7 MHz ou 14,3 MHz. Lorsque cette solution a été abandonnée en faveur d'un standard en composantes unique à l'échelle mondiale, le chiffre « 4 » a été conservé, bien que n'ayant plus de véritable signification. Selon la convention adoptée, ce chiffre représente la fréquence d'échantillonnage de la luminance, soit 13,5 MHz. Les chiffres suivants indiquent, quant à eux, la fraction de cette fréquence qui est attribuée aux signaux de différence de couleur, soit ici la moitié : 6,75 MHz (on peut dire que le « 1 » présente la fréquence unitaire de 3,75 MHz).

Cependant, on s'est assez rapidement rendu compte qu'il était peu probable que l'ensemble des applications actuelles et futures de la télévision s'accommodent d'un seul niveau de codage numérique, à savoir le 4:2:2. C'est pourquoi d'autres paramètres de codage ont été également définis, formant des normes « légères » d'une part, et des normes « haute qualité » d'autre part. Notons que les modifications apportées ne concernent que le traitement de la chrominance, la luminance restant intacte. Les normes légères sont caractérisées par un sous-échantillonnage des composantes de couleurs en horizontal (4:1:1) ou en vertical (4:2:0).

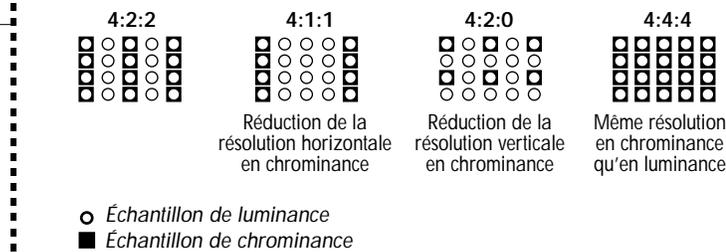
Elles sont aujourd'hui largement employées dans le domaine de la diffusion et de l'enregistrement numérique. Les normes haute qualité accordent au contraire à ces signaux de couleurs une bande passante aussi élevée que celle du signal de luminance (4:4:4), et peuvent également inclure un signal de découpe, alors représenté par un quatrième chiffre (4:2:2:4, 4:4:4:4). Elles sont utilisées dans certains équipements de postproduction haut de gamme effectuant des traitements complexes, comme des incrustations en chromakey ou des corrections colorimétriques précises.

La plus faible sensibilité de l'œil humain à la couleur qu'à la luminosité autorise de prélever deux fois moins d'échantillons de chrominance que d'échantillons de luminance. Dans la norme définie pour les applications de studio, le rapport entre les fréquences d'échantillonnage de Y, R-Y, B-Y est égal à celui des nombres 4:2:2. La bande passante du signal de luminance est de 5,75 MHz ; celle des signaux de différence de couleur est de 2,75 MHz.

Le 4:2:2 et ses déclinaisons

- 4:2:2** La luminance est échantillonnée à 13,5 MHz, les composantes de couleurs sont échantillonnées à 6,75 MHz. Le 4:2:2 est le niveau de codage principal de studio, utilisé dans l'ensemble des équipements numériques, ainsi que par les formats d'enregistrement haut de gamme.
 - 4:2:2:4** C'est un codage 4:2:2 incluant un signal de découpe échantillonné à 13,5 MHz.
 - 4:1:1** La luminance est échantillonnée à 13,5 MHz, les composantes de couleurs sont échantillonnées à une fréquence quatre fois plus faible, soit 3,375 MHz. Le 4:1:1 est notamment utilisé par le format d'enregistrement DVCPRO25.
 - 4:2:0** La luminance est échantillonné à 13,5 MHz, les composantes de couleurs sont échantillonnées à 6,75 MHz, mais alternativement une ligne sur deux. Le 4:2:0 est utilisé par les systèmes de diffusion numérique, le DVD, ainsi que par le DV et le DVCAM (en 625/50).
 - 4:4:4** La luminance et les composantes de couleurs sont échantillonnées à la même fréquence de 13,5 MHz, sur toutes les lignes. Le 4:4:4 notamment est utilisé par les stations informatiques de retouche et de compositing.
 - 4:4:4:4** C'est un codage 4:4:4 incluant un signal de découpe échantillonné à 13,5 MHz.
-

Figure 5.14. Sous-échantillonnage de la chrominance.
 « 4 » = 13,5 MHz
 « 2 » = 6,75 MHz
 « 1 » = 3,375 MHz



5.2.7. Choix de la structure d'échantillonnage

On appelle structure d'échantillonnage la répartition des échantillons dans le temps et l'espace. Rappelons qu'avec l'apparition de la couleur, le signal vidéo a perdu la périodicité de 2 trames qu'il avait en noir et blanc (4 trames en NTSC, 4 ou 8 trames en PAL, 4 ou 12 trames en SECAM). Pour le numérique, on a souhaité éliminer ce problème de gestion des cycles et revenir à une périodicité de 2 trames pour la structure d'échantillonnage, ce qui sous-entend que les images soient de structure identique. Trois types de structures ont alors été envisagées : orthogonale, quinconce ligne et quinconce trame.

- Dans la structure orthogonale, la phase horloge de l'échantillonnage est identique pour chaque ligne et chaque trame. Les échantillons sont situés aux mêmes emplacements d'une ligne à l'autre et d'une trame à l'autre.
- Dans la structure quinconce ligne, la phase horloge de l'échantillonnage est décalée d'une demi-période à chaque ligne. Les échantillons sont décalés en quinconce d'une ligne à l'autre. La périodicité de 2 trames est maintenue grâce à un déphasage à la fréquence image.
- Dans la structure quinconce trame, chaque trame est orthogonale, mais la phase horloge de l'échantillonnage est décalée d'une demi-période à chaque trame. Les échantillons sont alignés d'une ligne par rapport à l'autre sur une trame, mais ils sont décalés en quinconce entre deux trames.

Figure 5.15
Les différentes structures d'échantillonnage fixes possibles (périodicité 2 trames).

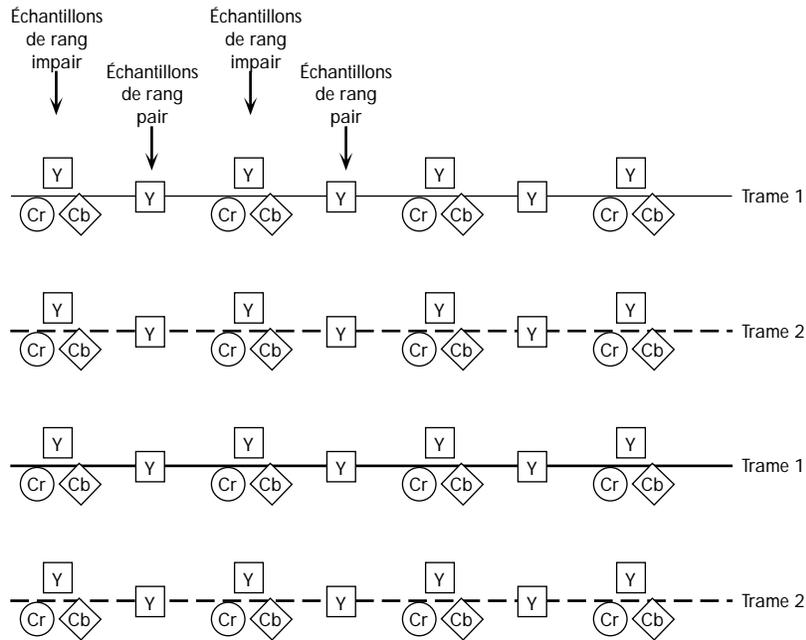
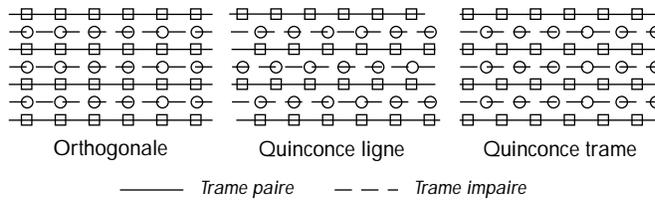


Figure 5.16

La structure orthogonale. Sur une image 4:2:2, les échantillons de différence de couleurs Cr et Cb coïncident avec les emplacements impairs des échantillons de luminance sur chaque ligne.

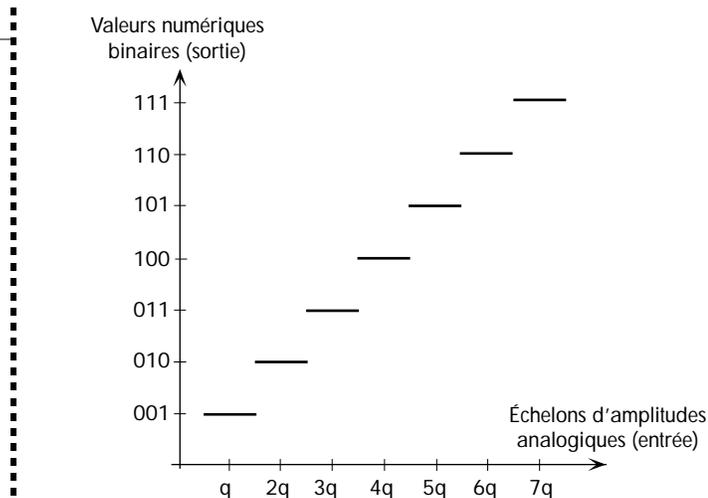
À la suite de nombreux essais subjectifs, c'est finalement la structure orthogonale (au demeurant la plus simple des trois) qui a été choisie (fig. 5.16). Nous retiendrons qu'elle se répète identique à elle-même à chaque ligne, chaque trame et chaque image. Dans la norme 4:2:2, les échantillons de différence de couleurs Cr et Cb coïncident avec les emplacements impairs des échantillons de luminance Y sur chaque ligne ; les emplacements pairs ne sont occupés que par les échantillons de luminance.

5.2.8. La quantification

5.2.8.1. Principe de base

La quantification est, après l'échantillonnage, la deuxième grande étape du processus de conversion analogique/numérique. Elle a pour but de faire correspondre à chaque amplitude discrète mesurée un nombre entier exprimé en base 2, dans laquelle n chiffres permettent de coder $N = 2^n$ valeurs distinctes. Au cours de cette étape de quantification, le signal analogique, qui peut prendre une infinité de valeurs, est converti en un signal constitué d'un nombre fini (N) de valeurs numériques codées sur n bits. On comprend alors que des erreurs d'arrondi par défaut ou par excès se produiront inévitablement. Car à plusieurs valeurs proches, mais cependant différentes, du signal analogique, correspondra une seule et même valeur numérique binaire. La précision du signal converti sera donc directement liée au nombre de valeurs disponibles pour coder les échantillons.

Figure 5.17
Principe de la quantification avec 3 bits – cas peu réaliste. Toutes les amplitudes analogiques inscrites à l'intérieur d'un échelon sont converties en une seule et même valeur numérique, celle du centre de l'échelon.



La figure 5.17 permet de bien comprendre cette phase de quantification et l'erreur qu'elle apporte. En fonction de la plage d'amplitudes à quantifier, on définit une échelle constituée d'un nombre fini d'intervalles « q » appelés « pas de quantification » ou « échelons de quantification », ou encore « quantums ».

À chaque échelon $q, 2q, 3q, \dots$ est associée une valeur numérique. À un instant t , l'amplitude du signal se trouvant à l'intérieur d'un échelon est remplacée par la valeur de cet échelon ; la valeur exacte de l'amplitude n'est pas prise en considération. Il est évident que plus les échelons sont petits, plus ils sont nombreux sur une plage donnée, et plus la précision du signal quantifié est grande.

Quoi qu'il en soit, l'erreur de quantification est toujours inférieure au quantum q , qui est en fait la plus petite valeur mesurable par le convertisseur ; pour reprendre l'analogie avec la pesée au moyen d'une balance, q représente le plus petit poids disponible. Elle représente une erreur de 1 unité sur le nombre total de quanta. Ainsi, un système travaillant sur 3 bits est caractérisé par une erreur maximale de $1/8$, soit un taux de 12,5 % ; un système à 8 bits donne une erreur maximale de $1/256$, soit 0,39 % ; tandis que le taux d'erreur d'un système à 10 bits est de 0,09 % ($1/1024$). De manière générale, plus le nombre de bits utilisé pour la quantification est élevé, plus la précision du système de conversion est grande, mais plus le temps de traitement est long. L'expression mathématique du quantum en fonction de l'amplitude maximale du signal d'entrée ($V_{\max} - V_{\min}$) et du nombre de bits de quantification n est la suivante :

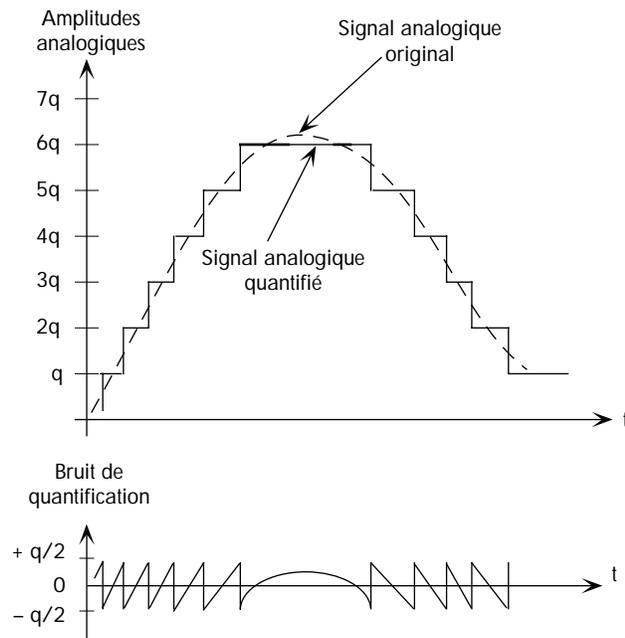
$$q = \frac{V_{\max} - V_{\min}}{2^n}$$

En pratique, il a été décidé que le signal de sortie change de valeur quand le signal d'entrée passe par l'une des valeurs $(2n + 1).q/2$, ce qui limite l'erreur à $\pm q/2$. La succession d'erreurs de quantification se traduit par la création d'un signal aléatoire, dont l'amplitude maximale est un demi-pas de quantification – il varie entre $\pm q/2$. Ce signal, appelé bruit de quantification, se superpose au signal utile. Plus le nombre de bits est élevé, plus le bruit de quantification est faible, donc plus le rapport signal sur bruit est élevé. En télévision, le rapport S/B d'un signal numérique est approché par la relation :

$$S/B \text{ (dB)} = 6 N + 10 \quad (N = \text{nombre de bits})$$

On retiendra que chaque bit supplémentaire améliore le rapport S/B de 6 dB.

Figure 5.18
La courbe représentant le signal quantifié a l'allure de marches d'escaliers. La succession des erreurs engendre un signal appelé bruit de quantification.



Au cours de l'étape de quantification, l'amplitude de chaque échantillon analogique prélevé sur le signal d'origine est convertie en un nombre binaire. Un nombre infini d'informations est donc remplacé par un nombre limité de valeurs disponibles. La précision de la quantification dépend de ce nombre de valeurs disponibles, c'est-à-dire du nombre de bits utilisés pour le codage des valeurs numériques : n bits donnent 2^n valeurs possibles.

5.2.8.2. La quantification du signal vidéo

À l'origine, la norme 4:2:2 spécifiait une quantification sur 8 bits des composantes du signal vidéo. Une quantification sur 8 bits permet de disposer de 256 niveaux numériques, dont 220 utiles pour représenter les niveaux sur une échelle de gris – avec une marge de réserve en dessous du noir et au-dessus du blanc. Or, si ce nombre de niveaux convenait pour la diffusion et le reportage,

Dans la pratique, il faut cependant tenir compte du comportement des supports d'enregistrement et des voies de transmission, qui présentent parfois des défaillances. Des erreurs peuvent alors être introduites dans le flux de données numériques et modifier le contenu du message.

Sur un magnétoscope, les principales sources d'erreurs sont les irrégularités du mécanisme d'entraînement (*gigue*, ou *jitter* en anglais), les fluctuations du niveau du signal, mais aussi la poussière, un défaut d'enduit magnétique, un léger froissement de la bande, etc.

En transmission, les dégradations pouvant affecter le signal sont essentiellement dues au bruit et aux distorsions : écho, évanouissement, diaphonie, intermodulation parasite, etc.

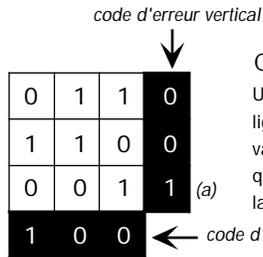
Une erreur en numérique, c'est soit une inversion de valeur binaire dans un mot, soit une absence fugitive de données. Une erreur est d'autant plus gênante qu'elle affecte les bits de poids fort, c'est-à-dire les plus significatifs.

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

La figure 5.20 donne un exemple extrêmement simplifié de correction d'erreurs, basé sur la loi de parité. Dans la réalité, les choses sont évidemment beaucoup plus complexes.

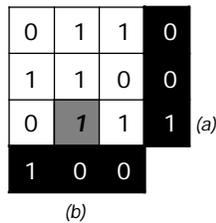
0	1	1
1	1	0
0	0	1

Données originales



Codage : ajout d'un code de correction d'erreur

Un bit supplémentaire, appelé bit de parité, est ajouté à la fin de chaque ligne et de chaque colonne du tableau de données. Ce bit prend une valeur qui rend pair le nombre total de « 1 » présent dans la ligne ou la colonne qu'il complète. Par exemple, sur la ligne (a), un « 1 » est ajouté pour que la somme des chiffres sur cette ligne soit égale à 2.



Décodage : vérification de la loi de parité et détection des erreurs éventuelles

Le nombre de « 1 » sur chaque ligne et sur chaque colonne est compté afin de vérifier la loi de parité établie lors du codage. Si, par exemple, sur une ligne, un nombre impair de « 1 » est trouvé, une recherche par colonne permet d'identifier le bit erroné. Dans l'exemple ci-contre, on compte un nombre impair de « 1 » sur la ligne (a) ainsi que sur la colonne (b). La valeur du bit se trouvant à leur intersection est erronée.

0	1	1
1	1	0
0	0	1

Correction de l'erreur

Le bit erroné est remplacé par sa valeur correcte.

.....

Figure 5.20
Principe de correction d'erreur par adjonction d'un code de parité.

5.2.9.1. L'entrelacement des données

La plupart des erreurs qui perturbent un signal numérique affectent généralement plusieurs dizaines de bits consécutifs. C'est pourquoi a été inventé le principe du brassage, ou entrelacement, des données. Celui-ci consiste à éloigner les unes des autres les informations à l'origine consécutives, dans le but de briser et disperser tout paquet d'erreurs. Ce brassage est pseudo-aléatoire, sa

formule étant connue par le codeur comme par le décodeur (chargé, lui, de tout remettre en ordre). À la lecture ou à la réception du signal, le désentrelacement a pour effet de répartir les petites erreurs spatialement sur l'image. Ces erreurs isolées sont plus facilement détectables et corrigibles, en tous cas beaucoup moins perceptibles, voire pas du tout.

5.2.9.2. La dissimulation

Il faut bien comprendre qu'un système de correction d'erreurs ne se contente pas d'effectuer une approximation de la valeur erronée, mais qu'il restitue intégralement la donnée d'origine. Cependant, quand les capacités de correction sont insuffisantes, on fait appel à un processus de dissimulation. On distingue plusieurs degrés de dissimulation d'erreurs détectées mais non corrigées. Parmi elles, la duplication de données adjacentes – qui donne un résultat assez médiocre en vidéo – et le calcul de la valeur moyenne entre échantillons proches : c'est le meilleur compromis en vidéo.

Avant la transmission ou l'enregistrement d'un signal numérique, on prévient tout risque d'erreurs d'une part en brassant les données, et d'autre part en leur ajoutant des données de contrôle. Ces dernières sont utilisées par le décodeur pour vérifier la validité des données reçues et, le cas échéant, détecter et corriger – ou du moins dissimuler – les erreurs.

5.2.10. *Le codage de canal*

Le codage de canal a pour but de moduler le flux de données numériques pour l'adapter aux caractéristiques du canal de transport ou d'enregistrement. Par exemple, les longues suites de 1 ou de 0 du message numérique créent des composantes continues, impossibles à relire une fois enregistrées. Il faut donc les rompre par un codage particulier, dont l'unique but est de modifier la forme du signal, sans bien sûr toucher à son contenu. Il existe plusieurs codes, chacun ayant ses avantages et ses inconvénients qui les rendent plus appropriés à tel ou tel type d'application : réduction de la composante continue, mais aussi recouvrement de l'horloge, distribution spectrale, etc. Quelques-uns d'entre eux sont décrits ci-après et représentés sur la figure 5.21 (p. 220).

5.2.10.1. NRZ (Non Retour à Zéro)

C'est le plus simple de tous. Une donnée binaire « 1 » engendre un niveau haut du signal et une donnée binaire « 0 » un niveau bas. Ce code est caractérisé par une importante composante continue, ainsi que par une absence du signal d'horloge, qu'il faut donc régénérer à la réception.

5.2.10.2. S-NRZ (*Scrambled NRZ* = NRZ embrouillé)

Il s'agit d'une variante du code NRZ dans laquelle le signal est mélangé – somme modulo 2 – avec une séquence binaire pseudo-aléatoire. Cela a pour effet d'une part de générer des transitions fréquentes en brisant les longues suites de symboles identiques, et d'autre part de réduire la valeur de la composante continue, ce qui est essentiel pour le transformateur du tambour de têtes d'un magnétoscope.

5.2.10.3. NRZI (Non Retour à Zéro Inversé)

Un « 1 » détermine une transition au milieu de la demi-période d'horloge, un « 0 » n'a aucun effet. Ce code, qui présente l'avantage d'être insensible aux inversions de polarité, est caractérisé par une faible composante continue. Il est notamment utilisé dans les liaisons série 4:2:2.

5.2.10.4. Biphase Mark

Un « 0 » provoque une transition et un maintien du niveau pendant toute la période d'horloge, tandis qu'un « 1 » entraîne une transition et un changement de niveau à la moitié de la demi-période d'horloge. Ce code présente une composante continue nulle et contient tous les fronts d'horloge – il est dit auto-synchroniseur. Il est employé par le code temporel longitudinal LTC des magnétoscopes.

5.2.10.5. Miller, ou MFM (*Modified Frequency Modulation*)

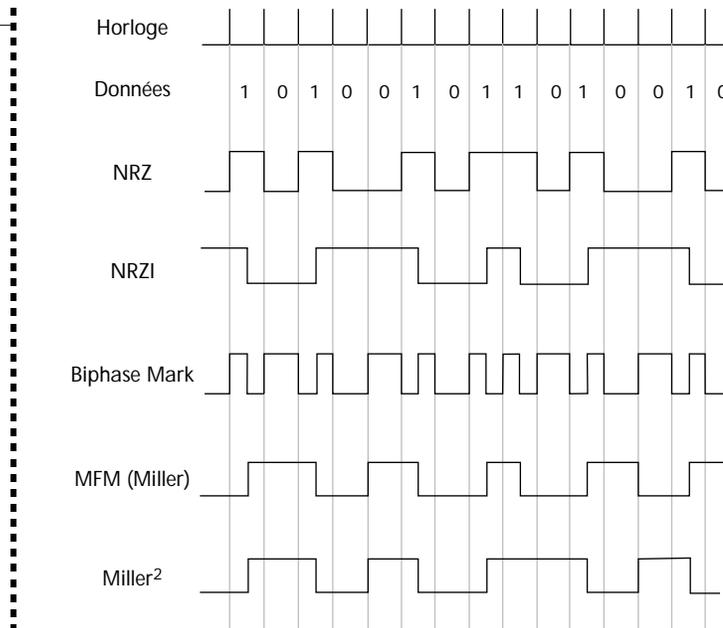
Un « 1 » donne une transition au milieu de la demi-période d'horloge, un « 0 » isolé ne donne aucune transition, mais une transition se produit entre deux « 0 » consécutifs. Ce code, inventé par

A. Miller, de la société Ampex, peut comporter une composante continue. Quant à l'horloge, elle est facilement extractible, le signal présentant au moins une transition tous les deux bits.

5.2.10.6. Miller carré (Miller²)

Il possède les mêmes caractéristiques que le code MFM, auxquelles s'ajoute la règle suivante : la dernière transition d'une suite paire de « 1 » est omise. Ce code, qui présente une composante continue très faible, est utilisé dans certains magnétoscopes numériques.

Figure 5.21
Les principaux codages de canal.



Il existe d'autres codes plus complexes et plus puissants, parmi lesquels on peut citer le très sophistiqué code 8/14 (*Eight to Fourteen Modulation* ou EFM), utilisé par le Compact Disc et par certains magnétoscopes. Son principe repose sur la conversion des mots de 8 bits de données en mots de 14 bits – dont 3 de redondance –, dans le but de réduire la composante continue ainsi que la gamme de fréquences en haut et en bas du spectre.

Le codage de canal a pour but de structurer le flux numérique, afin de lui conférer des propriétés spectrales optimisées pour son enregistrement ou sa transmission.

5.2.11. La ligne vidéo numérique

Les lignes actives analogiques des systèmes à 625 et 525 lignes sont de durées légèrement différentes.

La ligne active numérique doit logiquement contenir un nombre suffisant d'échantillons pour couvrir la plus longue des deux, c'est-à-dire celle des systèmes à 525/60 lignes, qui en requiert 710. Il a finalement été choisi 720 échantillons pour le signal de luminance et 360 échantillons pour chaque signal de différence de couleurs. Une ligne active numérique renferme donc un total de 1 440 échantillons. La référence des temps pour l'opération de conversion analogique-numérique est donnée par le front avant des impulsions de synchronisation ligne à mi-amplitude ; c'est donc à cet instant qu'apparaît le premier échantillon.

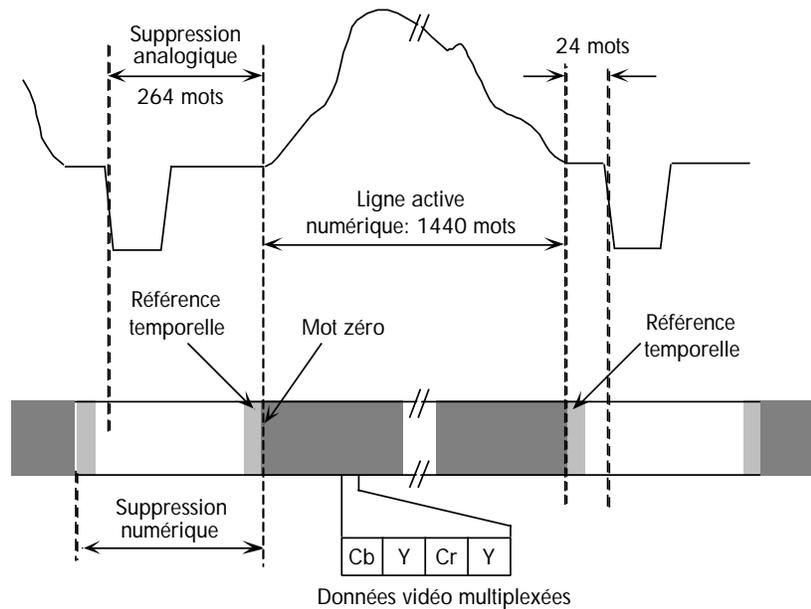


Figure 5.22
Relation temporelle entre la ligne vidéo analogique et la ligne vidéo numérique.

5.2.12. La trame vidéo numérique

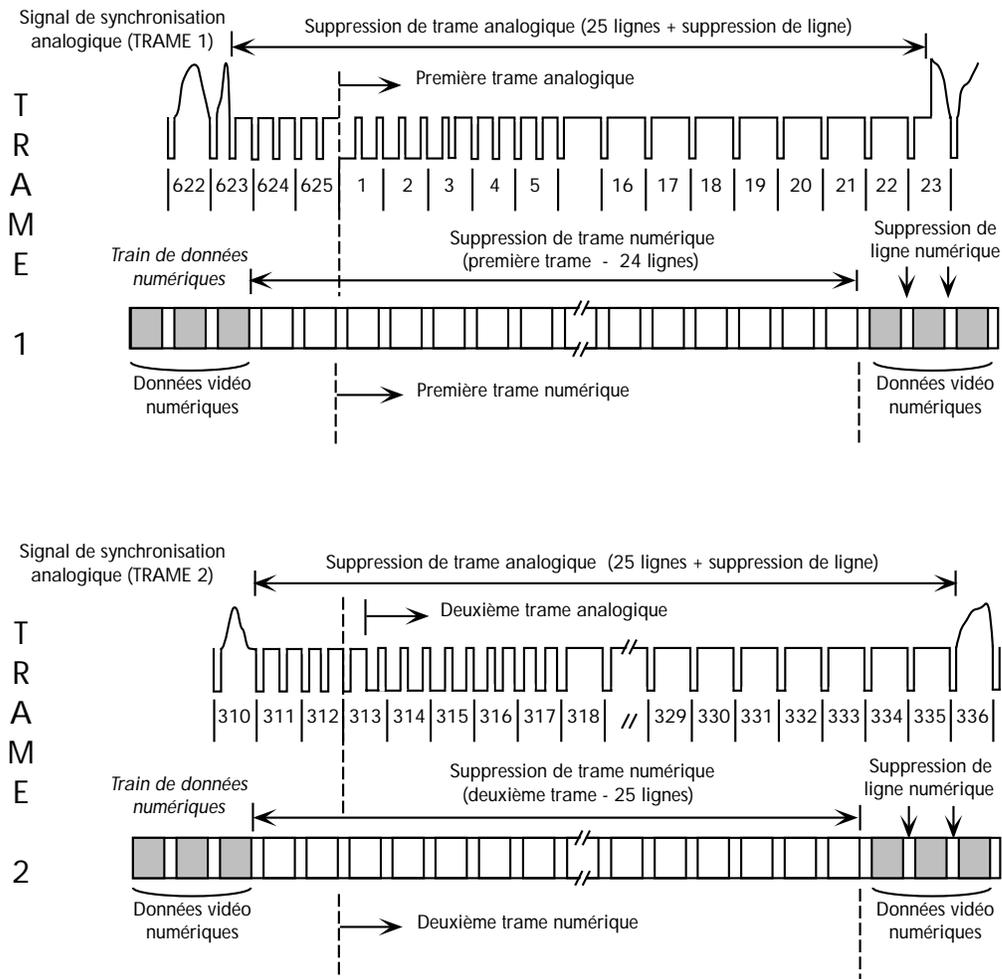


Figure 5.23
Relations entre les trames analogiques et les trames numériques.

La figure 5.23 montre les relations entre les trames numériques et les trames analogiques, ainsi que la position de l'intervalle de suppression de trame numérique pour les systèmes à 625 lignes. Pour éviter d'avoir à créer des demi-lignes numériques, les

débuts et fins de suppression trame coïncident avec les débuts et fins de suppression ligne. En 625/50, chaque trame active renferme ainsi un nombre entier de lignes complètes, et une image est constituée de 576 lignes utiles. La suppression trame numérique s'étend sur 24 lignes dans la première trame et sur 25 lignes dans la deuxième trame. Trois lignes de suppression sur chaque trame sont réservées au transport de données auxiliaires, qui bénéficient ainsi d'un débit de 1,7 Mbits/s.

Tableau 5.1.
Les principaux paramètres de la norme 4:2:2.

	625/50	525/60
Signaux à coder	Y, (R-Y), (B-Y) corrigés en gamma	
Fréquences d'échantillonnage Luminance Différence de couleurs	13,5 MHz 6,75 MHz	
Bandes passantes équivalentes Luminance Différence de couleurs	5,75 MHz 2,75 MHz	
Nombre d'échantillons par ligne complète Luminance Différence de couleurs	864 432	858 429
Nombre d'échantillons utiles par ligne Luminance Y Différence de couleurs	720 360	
Structure d'échantillonnage	Orthogonale. Les échantillons de chrominance coïncident avec les échantillons impairs de luminance	
Quantification (8 bits)	256 niveaux dont : 220 utiles pour Y, 225 pour Cr, Cb	
Quantification (10 bits)	1 024 niveaux dont : 880 utiles pour Y, 900 pour Cr, Cb	

Tableau 5.2.
Les débits du signal vidéo.

	Signal utile 720 × 576		Signal complet 864 × 625	
	8 bits	10 bits	8 bits	10 bits
Signal 4:2:2	166 Mbits/s	207 Mbits/s	216 Mbits/s	270 Mbits/s
Signal 4:2:0 ou 4:1:1	124 Mbits/s		162 Mbits/s	

5.2.13. *Les interfaces numériques parallèle et série*

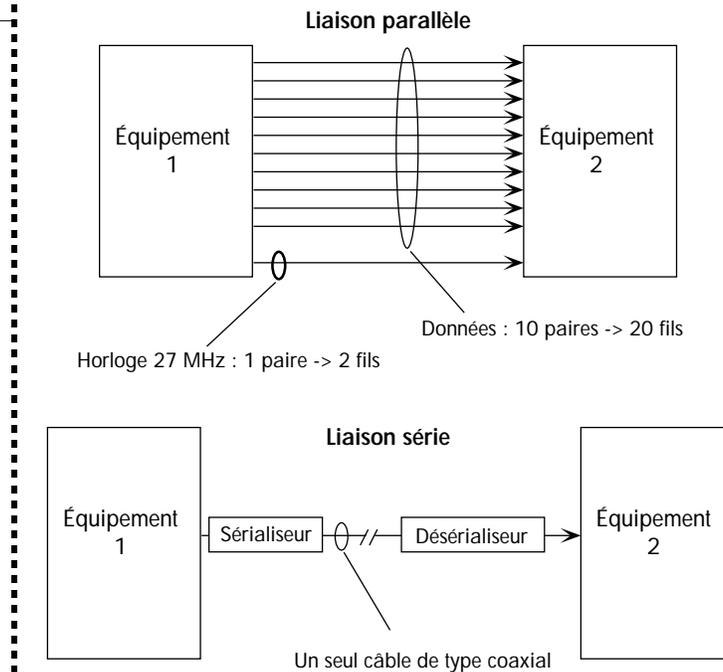
On désigne par le terme interface l'ensemble des paramètres relatifs à l'interconnexion entre deux équipements ; ces paramètres concernent le type, le nombre et la fonction des circuits de liaison, ainsi que la forme des signaux échangés. Le signal numérique 4:2:2 peut être transporté par une interface parallèle ou une interface série.

Dans une interface parallèle, tous les 8 ou 10 bits d'un mot de données sont transportés simultanément sur des supports distincts ; c'est le mode de transmission le plus naturel. Une liaison parallèle véhicule en même temps 8 ou 10 trains numériques composés des mots binaires Cb, Y, Cr, Y, Cb, Y, Cr, etc., et rythmés à la même fréquence d'horloge de 27 MHz (13,5 + 6,75 + 6,75). Ces trains de bits sont transmis sur des paires symétriques (avec deux fils par donnée), selon une polarité définie. L'architecture interne de la plupart des équipements numériques fonctionne sous une forme parallèle qui se contente de circuits relativement lents et simples, donc bon marché. Cependant, lorsque le signal doit circuler entre différents équipements, l'utilisation d'une liaison parallèle présente plusieurs inconvénients. Le câble à multi-conducteur requis est onéreux et peu maniable ; le connecteur normalisé de type 25 broches est loin d'être idéal pour la réalisation d'un panneau de brassage ou d'une grille de commutation. De plus, la longueur maximale du câble ne peut excéder quelques dizaines de mètres – au-delà, les temps de propagation des signaux peuvent varier –, ce qui peut rendre très lourde la connectique d'un studio.

Dans une interface série, tous les 8 ou 10 bits d'un mot de données et tous les mots successifs sont transmis les uns à la suite des autres sur un seul et unique support. La sérialisation du signal numérique 4:2:2 présente ainsi l'avantage considérable de permettre son transport sur un seul câble coaxial traditionnel (le même que celui utilisé dans les installations analogiques) pouvant atteindre 300 m de long. Le problème de la distribution des signaux et du raccordement des équipements dans le studio est ainsi résolu, puisqu'il devient presque aussi simple qu'en composite, avec un câble par source. L'électronique de l'interface série

est cependant plus complexe que celle de l'interface parallèle, tant au niveau de l'émetteur (sérialisation) qu'à celui du récepteur (désérialisation). Sony a développé un circuit intégré assurant toutes les fonctions nécessaires à la conversion d'un signal parallèle (8 ou 10 bits) en un signal numérique série à 270 Mbits/s. Ce circuit permet de transmettre sur un câble unique le signal vidéo numérique composantes 4:2:2 (525 ou 625 lignes), avec quatre pistes audio numériques dites *embedded* et un code temporel insérés durant les intervalles de suppression. En bout de chaîne, un autre circuit effectue les opérations inverses pour redonner au signal sa forme parallèle. L'interface numérique série aura été un élément clé dans le succès rencontré par le numérique. Elle est normalisée sous l'appellation *Serial Digital Interface (SDI)*. Cependant, le débit total du signal délivré est beaucoup trop élevé pour être transmis par les réseaux de diffusion traditionnels. Seules sont donc concernées ici les applications de production et de postproduction broadcast. Pour le reste, il faut faire appel aux techniques de réduction de débit décrites dans le chapitre 6.

Figure 5.24
Liaisons parallèle et série
10 bits entre deux équipements. Le sérialiseur et le désérialiseur
sont des circuits intégrés.



Le transport d'un signal numérique 10 bits en parallèle nécessite 10 paires symétriques à 27 Mbits/s, additionnées d'une paire pour l'horloge.
Si ce signal est mis sous forme série, un simple câble coaxial permet de le véhiculer, avec un débit élevé à 270 Mbits/s.

5.2.13.1. Sériation du signal 4:2:2

Comme un unique support est utilisé pour transporter sous forme série tous les bits du signal les uns à la suite des autres, le débit, et par conséquent la bande passante requise, est multiplié par le nombre de bits de quantification.

Prenons l'exemple d'un signal parallèle sur 10 bits échantillonné à la fréquence de 1 MHz : 1 bit est produit toutes les 1 μ s. Si ce signal doit être transmis sous forme série, ce n'est plus 1 mais 10 bits qui doivent être produits en 1 μ s ; le débit du signal sérialisé passe donc de 1 à 10 MHz.

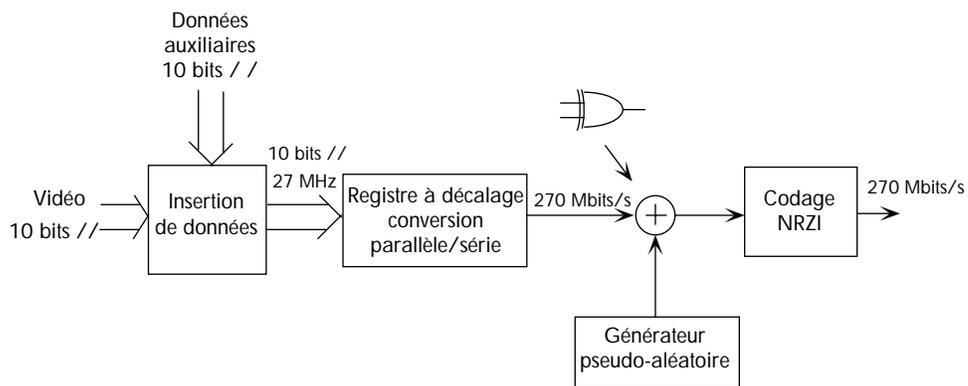


Figure 5.25
Principe de la sérialisation du signal numérique.

Pour en revenir à la vidéo, les 10 fils à 27 Mbits/s sont remplacés par un fil à 270 Mbits/s. Avec cette augmentation du débit se pose le problème de la synchronisation : le décodeur qui reçoit le flot de données doit être capable d'identifier chaque bit qui lui parvient afin de reconstituer correctement le signal vidéo parallèle. Comme

aucun signal de référence temporelle externe n'est transmis – le but de la sérialisation est de minimiser le nombre de connexions –, l'information d'horloge doit être extraite du signal numérique lui-même, ce qui impose une certaine mise en forme côté émission. Un codage de canal auto-référencé est pour cela utilisé : il doit permettre de recouvrer l'horloge dans le décodeur, mais également de minimiser la composante continue – pour éviter toute distorsion pouvant affecter la forme du signal –, ainsi que d'homogénéiser la distribution spectrale du signal. Pour satisfaire à ces exigences, les informations binaires, codées en NRZ, sont d'abord entrelacées à l'aide d'une séquence bien définie, fournie par un générateur pseudo-aléatoire. Le signal est ensuite soumis à un codage NRZI à l'issue duquel il renferme un nombre de transitions suffisant pour que le décodeur puisse en extraire les impulsions d'horloge.

5.2.13.2. Désérialisation du signal 4:2:2

À l'autre extrémité de la liaison, il faut désérialiser le signal pour qu'il puisse attaquer les équipements en parallèle. La désérialisation se décompose en plusieurs étapes.

Une correction automatique de câble – égalisation – est d'abord appliquée sur la totalité de la réponse en fréquences pour corriger les pertes causées par la longueur du câble. L'horloge à 270 MHz est ensuite régénérée grâce à la détection des nombreux fronts de transitions du codage NRZI. Le signal NRZI est alors converti en un signal NRZ, et les données sont désentrelacées pour retrouver leur ordre chronologique. Un registre à décalage synchronisé à 270 MHz convertit alors les données série sous forme parallèle. Puis les informations auxiliaires – audio, code temporel – sont séparées du signal vidéo parallèle et transmises séparément vers les circuits adéquats.

Précisons enfin que lorsqu'un signal sur 10 bits arrive en entrée d'un équipement travaillant sur seulement 8 bits, les deux derniers bits sont mis à zéro, l'erreur commise étant dissimulée par un arrondi soit par défaut, soit par excès.

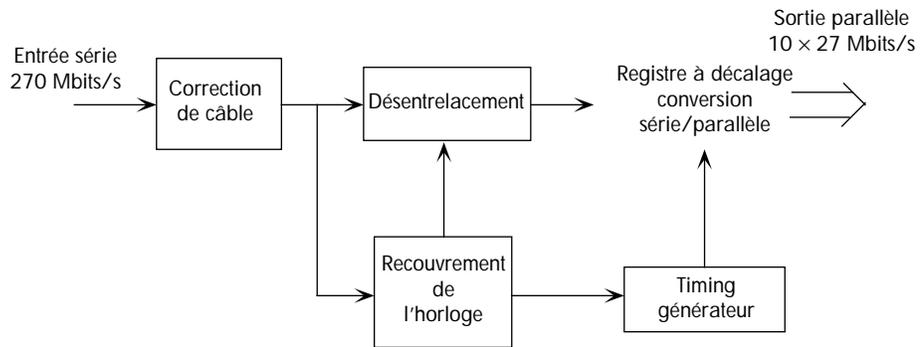


Figure 5.26
Principe de la désérialisation du signal numérique.