

# Réseaux Multimédia

© 2002 Damien Magoni

Toutes les illustrations

© 2001 Pearson Education Limited – Fred Halsall

# Contenu

- Représentation des informations multimédia
  - Numérisation
  - Structure d'un encodeur
  - Structure d'un décodeur
  - Représentation du texte, des images et de l'audio/vidéo
- Compression du texte et des images
- Compression audio et vidéo
- Standards pour les communications multimédia

# Partie 1

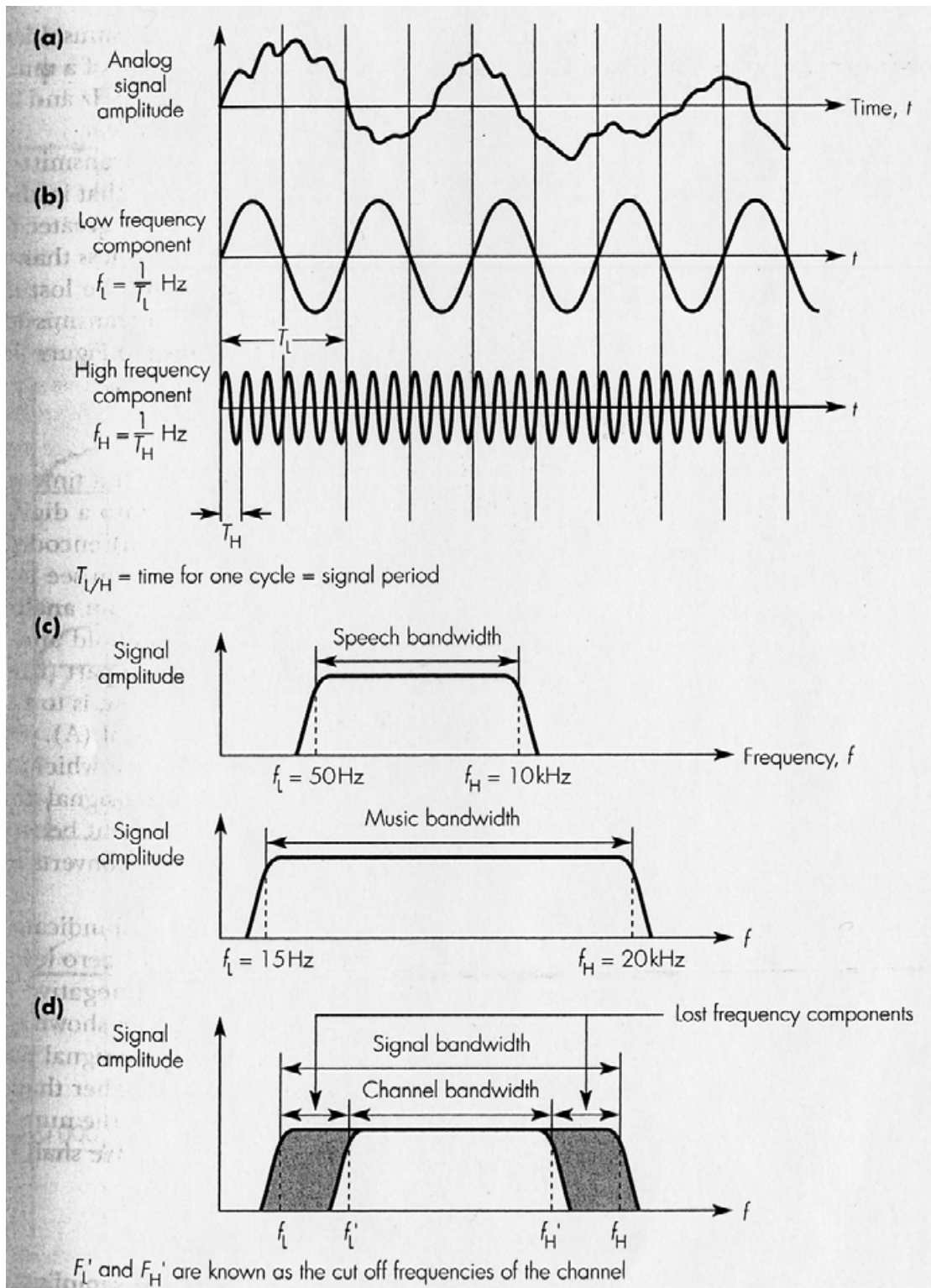
Représentation des  
informations multimédia

# Numérisation

- Tout type d'information est stocké et traité sous forme numérique
- Un signal dont l'amplitude varie dans le temps de manière continue est un signal **analogique**
- La conversion d'un signal analogique en numérique est réalisée par un **encodeur** de signal

# Principe de la numérisation

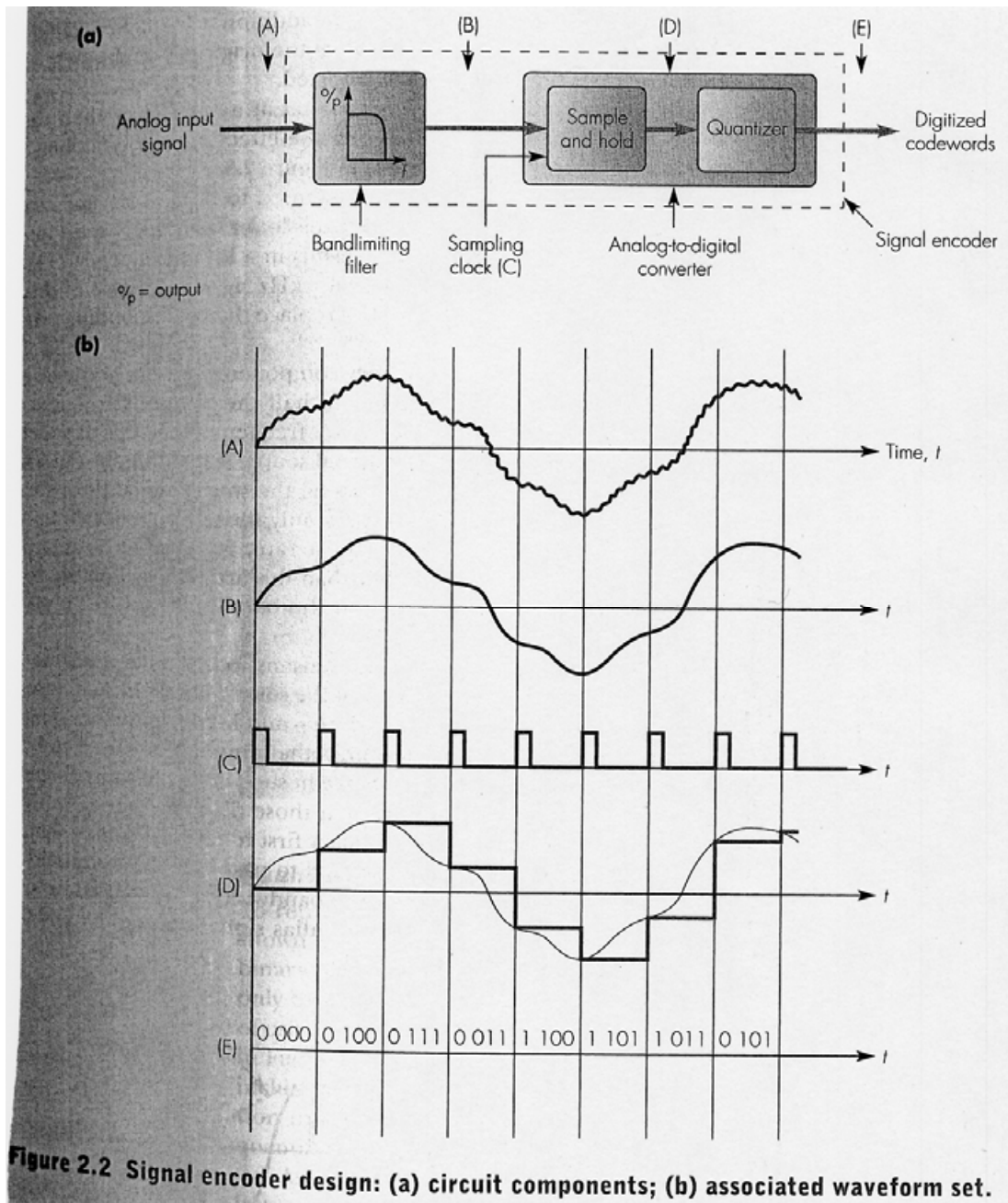
- Tout signal peut être décomposé en une somme de signaux sinusoïdaux de fréquences multiples par une transformation de **Fourier**
- Le spectre des fréquences des composantes sinusoïdales forme la **bande passante** du signal



**Figure 2.1 Signal properties: (a) time-varying analog signal; (b) sinusoidal frequency components; (c) signal bandwidth examples; (d) effect of a limited bandwidth transmission channel.**

# Schéma d'un encodeur

- Un encodeur contient deux circuits :
  - Un filtre passe-bas
  - Un convertisseur analogique numérique
- Le filtre passe-bas élimine les hautes fréquences inutiles (A)
- Le signal filtré est échantillonné à intervalles réguliers (D)
- Le signal échantillonné est quantifié (E)





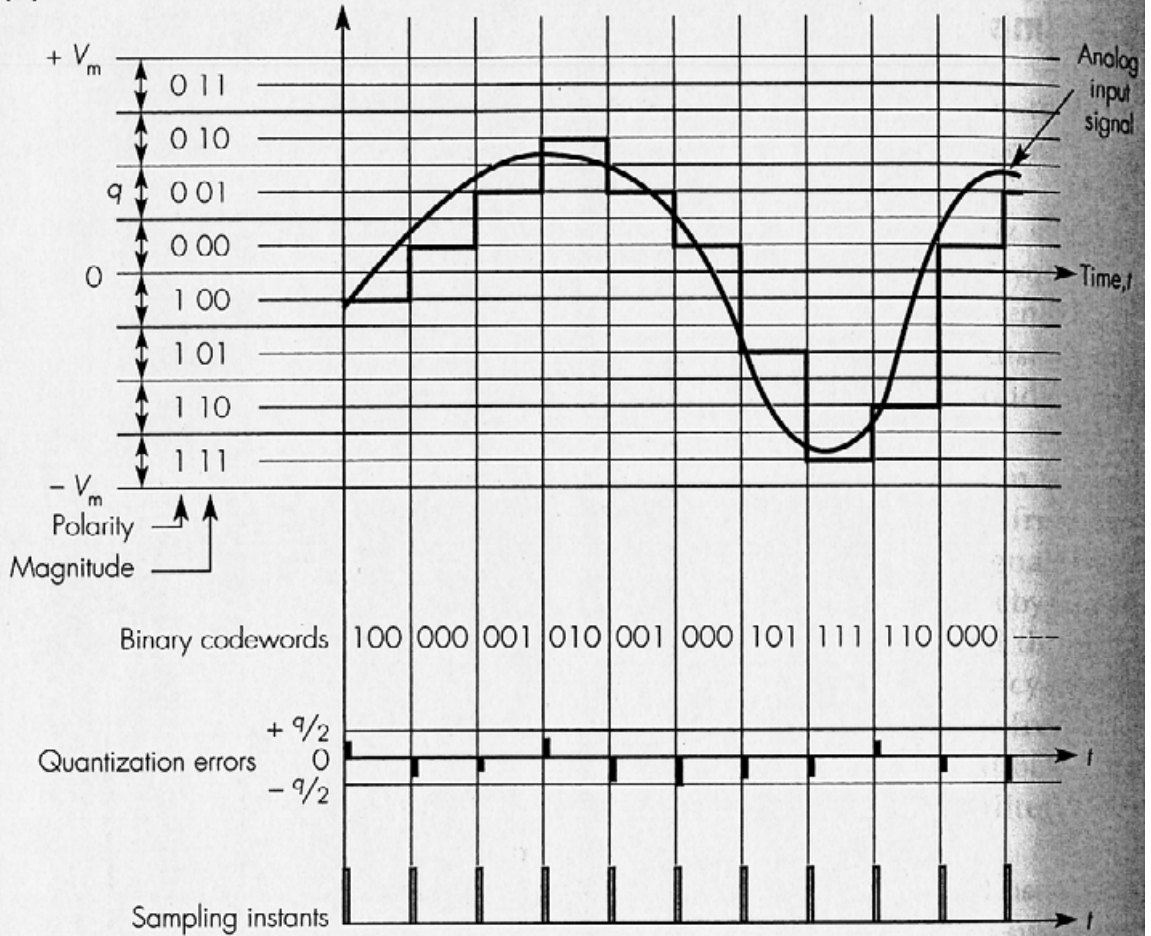
# Fréquence d'échantillonnage

- **Théorème de Nyquist** : pour obtenir une représentation précise du signal il faut l'échantillonner à une fréquence  $\geq$  à 2 x la plus haute fréquence sinusoïdale présente dans le signal
- Cette fréquence dite de Nyquist est mesurée en Hz ou en échantillons par seconde (eps)
- Une fréquence d'échantillonnage inférieure à celle de Nyquist produit des composantes fréquentielles non présentes dans le signal d'origine et le rendent distordu

# Intervalles de quantification

- L'amplitude est représentée par un nombre fini de bits d'où une perte d'information
- Si le découpage est linéaire et utilise  $n$  bits, l'**intervalle de quantification** est donné par :  
$$q = 2V_{\max} \div 2^n$$
- L'**erreur de quantification** est donc au plus égale à  $q \div 2$
- L'erreur de quantification varie entre chaque échantillon d'où le terme de **bruit** de quantification
- Le spectre dynamique du signal est égal à :  
$$D = 20 \log (V_{\max} \div V_{\min}) \text{ dB}$$

(a)



(b)

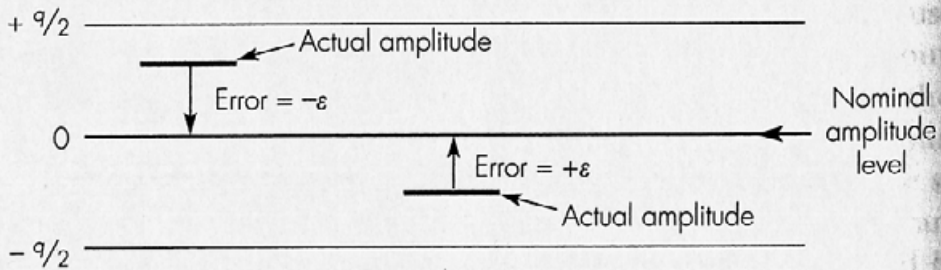
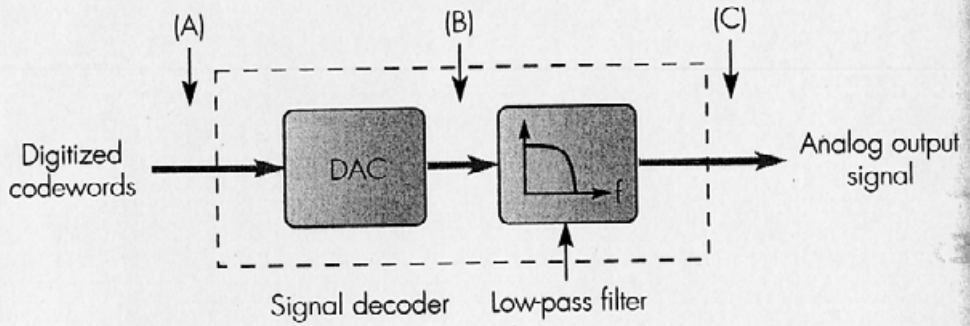


Figure 2.4 Quantization procedure: (a) source of errors; (b) noise polarity.

# Schéma d'un décodeur

- Un décodeur contient deux circuits :
  - Un convertisseur numérique analogique
  - Un filtre passe-bas
- Chaque mot de code est converti en un signal carré analogique (B)
- Cette conversion introduit des hautes fréquences qu'on élimine par un filtre passe-bas dont la fréquence de coupure est la même que celle du FPB de l'encodeur (C)

(a)



(b)

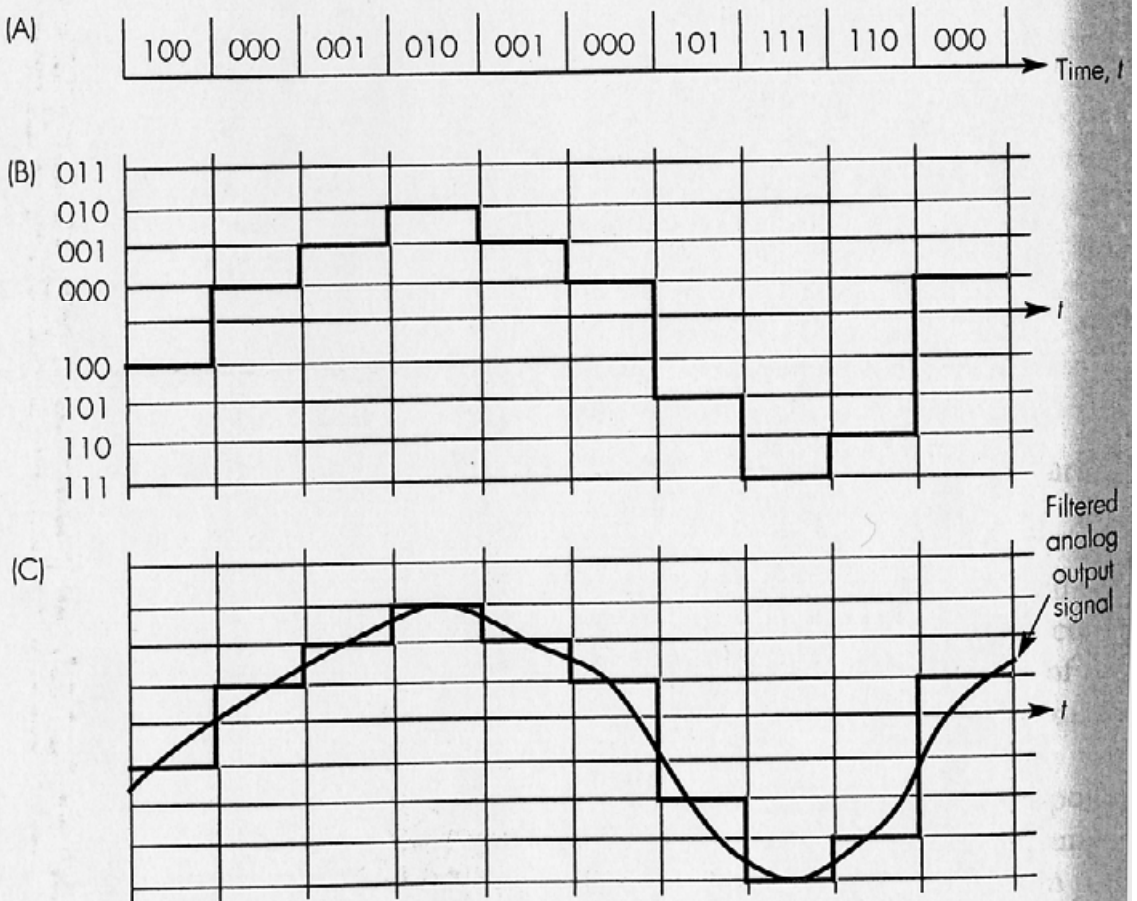


Figure 2.5 Signal decoder design: (a) circuit components; (b) associated waveform set.

# Texte

- Trois types de texte sont utilisés pour produire des documents :
  - Texte non formaté (*plain text*) : utilise un jeu de caractère tel que ASCII, EBCDIC, etc.
  - Texte formaté (*rich text*) : ajoute des commandes de formatage de police (italique, gras, etc.) et de paragraphe (e.g. retraits, alignement, etc.), souvent propriétaire tel que RTF, DOC, etc.
  - Hypertexte : ajoute la possibilité de créer des liens qui pointent vers d'autres endroits du texte ou vers d'autres documents, habituellement codé en HTML pour les documents placés sur le Web

# Graphiques

- Deux types de graphiques sont utilisés :
  - Images *bitmap* : elles sont composées d'un ensemble de pixels (*picture elements*), elle proviennent de logiciels de dessin *bitmap*, d'images numérisées par un scanner ou de photos numériques
  - Images vectorielles : elles sont composées d'objets qui sont des courbes mathématiques (dites de Bézier du nom du mathématicien qui les a inventées) ou des droites et segments qui ont une couleur et une taille. Leur principal avantage est de ne pas souffrir du phénomène de pixelisation lorsque on cherche à faire un zoom sur une partie ou la totalité de l'image

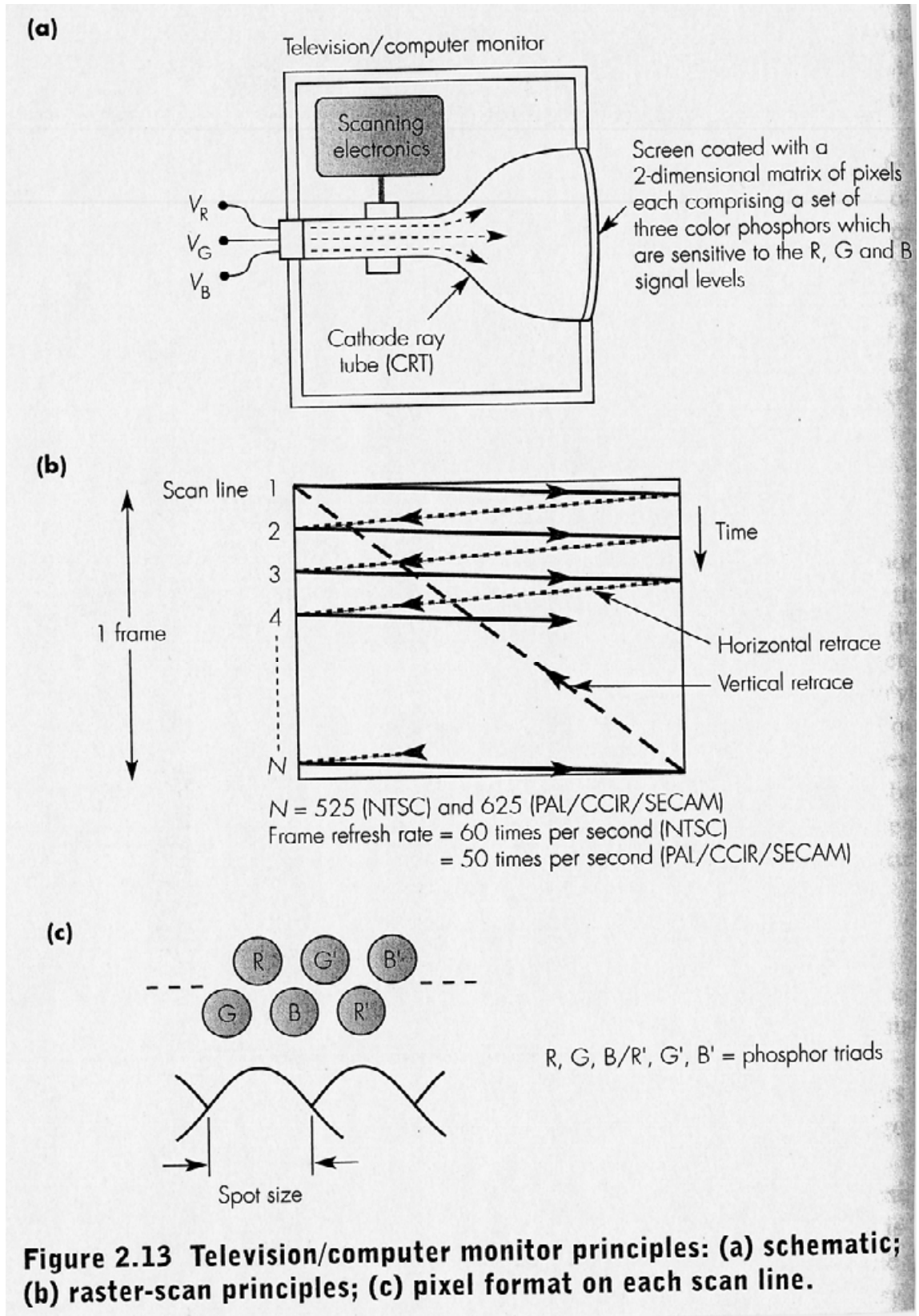
# Couleurs

- La gamme des couleurs peut être obtenue par composition de 3 couleurs primaires
- Mélange additif (rouge, vert et bleu) exprimé en coordonnées RGB :
  - $(0,0,0)$  = noir
  - $(1,1,1)$  = blanc
- Mélange soustractif (cyan, magenta et jaune) exprimé en coordonnées CMY :
  - $(0,0,0)$  = blanc
  - $(1,1,1)$  = noir



# Images sur écran

- Balayage à trames d'en haut à gauche jusqu'en bas à droite : balayage progressif
- Chaque jeu complet de lignes horizontales est une image (*frame*)
- Chaque pixel est un point qui se mélange à ses voisins (un point fait typiquement 0.025 pouces / 0.635 mm)
- Chaque pixel comprend une triade phosphorique (rouge, vert, bleu)
- Le **taux de rafraîchissement** indique la fréquence de retraçage des lignes



# Format d'écran

- Le format (*aspect ratio*) est égal à la largeur de l'image visible à l'écran divisé par sa hauteur
- Écrans :
  - anciens : 4/3
  - nouveaux : 16/9
- Normes :
  - NTSC : USA (480 lignes visibles)
  - PAL : GB (576 lignes visibles)
  - CCIR : Allemagne (idem)
  - SECAM : France (idem)
- Pour conserver un format 4/3 il faut utiliser  $640 \times 480$  sur un moniteur NTSC

# Acquisition numérique d'images

- Utilisation d'un capteur d'image composé d'une grille bidimensionnelle de cellules sensibles à la lumière (photosites)
- Un capteur CCD (*charge coupled device*) convertit l'intensité lumineuse d'un photosite en une charge électrique équivalente
- Pour les capteurs couleurs :
  - Un capteur avec mélange de photosites RGB
  - Trois expositions successives devant les filtres R, G, et B (pour image statique uniquement)
  - Lumière séparée en RGB puis envoyée vers 3 capteurs différents (cher)

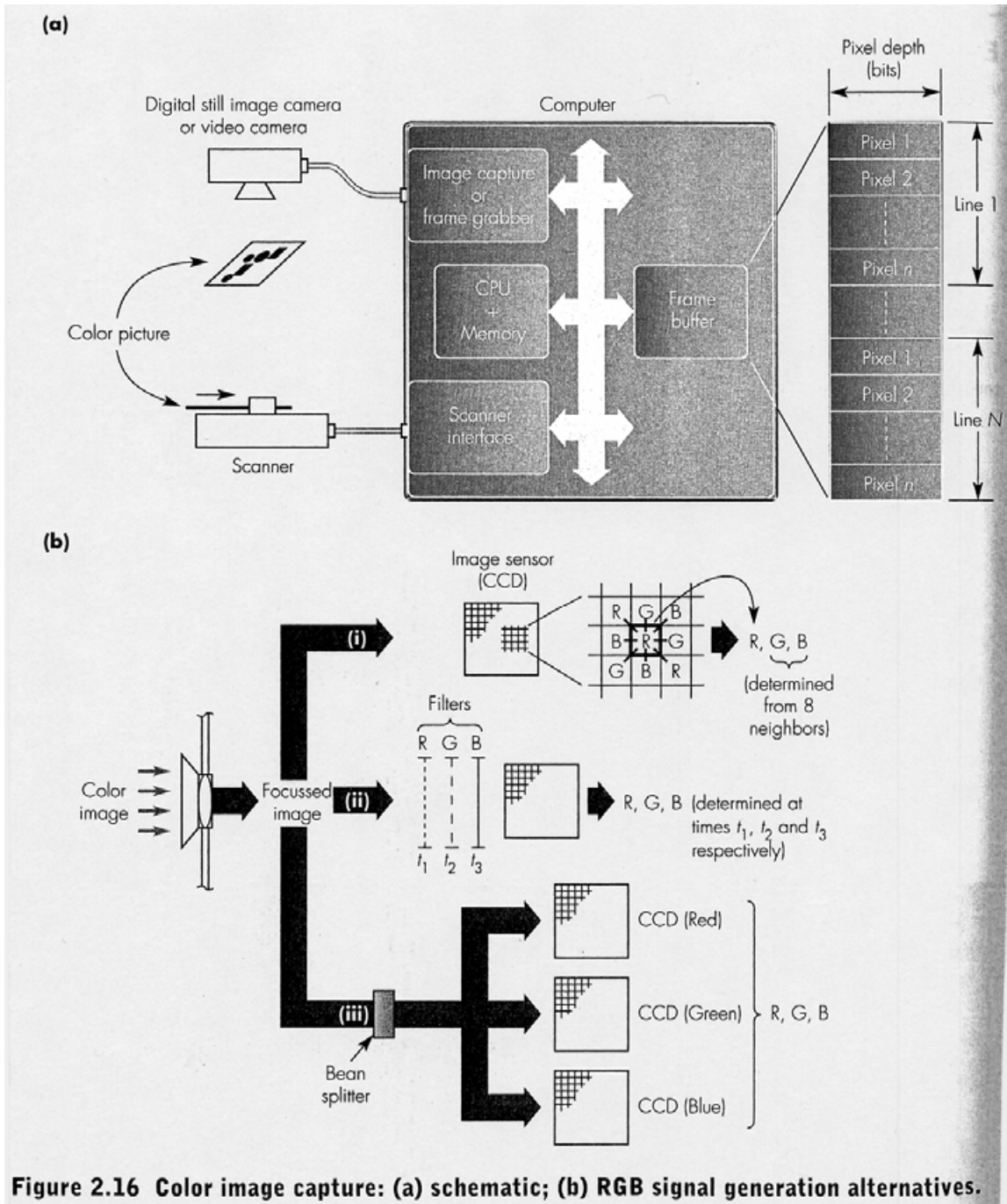


Figure 2.16 Color image capture: (a) schematic; (b) RGB signal generation alternatives.

# Audio

- Deux types de signaux audio :
  - Parole : utilisée dans le téléphone, le vidéo-téléphone, etc.
  - Musique : utilisée dans le CD audio, la télévision, la vidéo
- Bande passante :
  - Parole : 50 à 10k Hz
  - Musique : 15 à 20k Hz
- Nombre de bits nécessaires à la quantification pour avoir une dynamique suffisante :
  - Parole : 12 bits
  - Musique : 16 bits
- La stéréophonie nécessite un débit (*bit rate*) double

# Voie PCM

- Réseau téléphonique public :
  - Bande passante voie : 200 à 3,4k Hz
  - 7/8 bits (US/Europe) par échantillons pour un débit de 56/64 kbps
- La procédure de numérisation est nommée PCM (*pulse code modulation*) et est définie dans la norme de l'ITU-T G.711
- On utilise un compresseur pour rendre les intervalles de quantification non linéaires (l'oreille a une sensibilité logarithmique)
- Deux méthodes de compression-expansion :  $\mu$ -law (USA) et A-law (Europe)

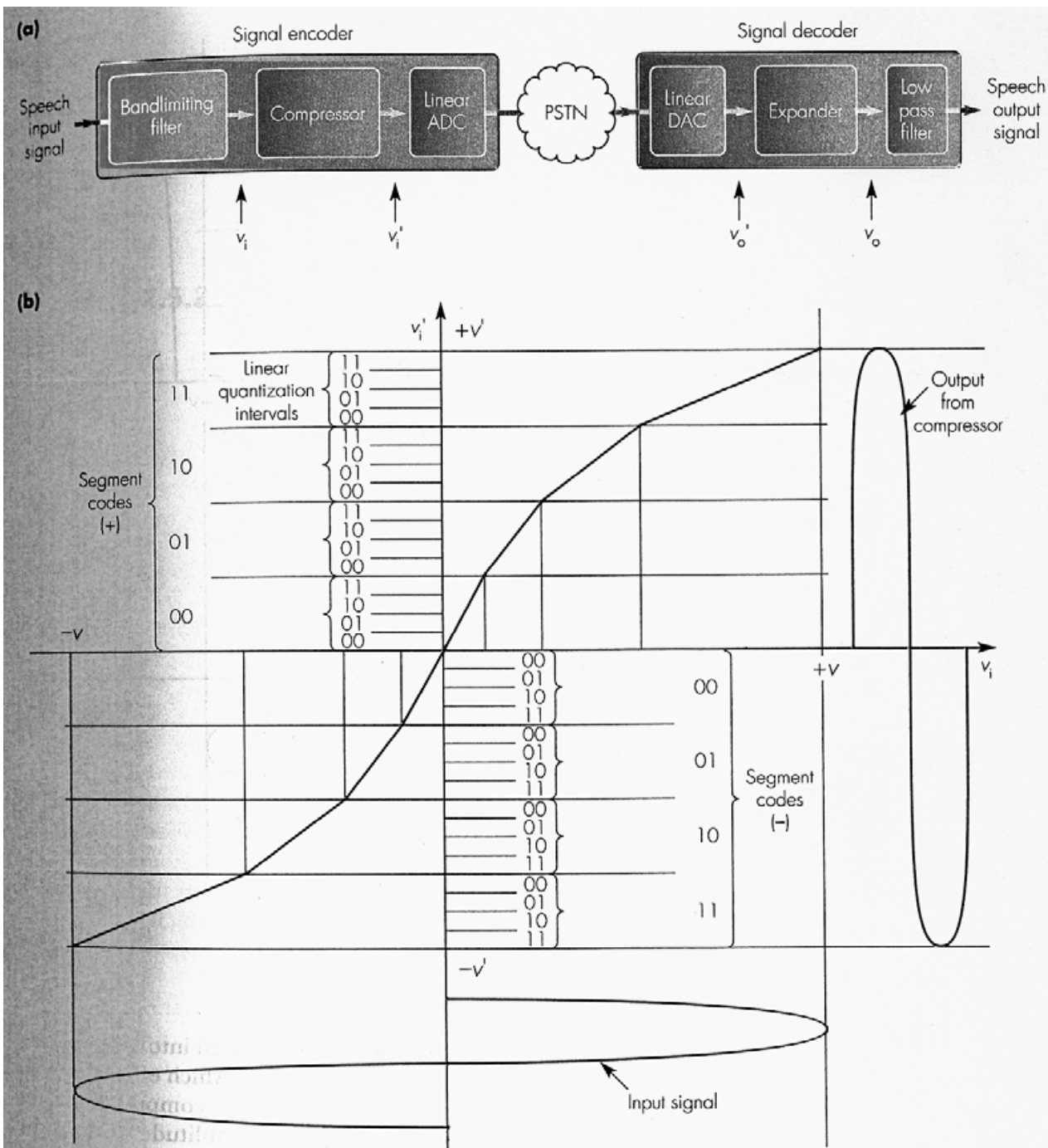
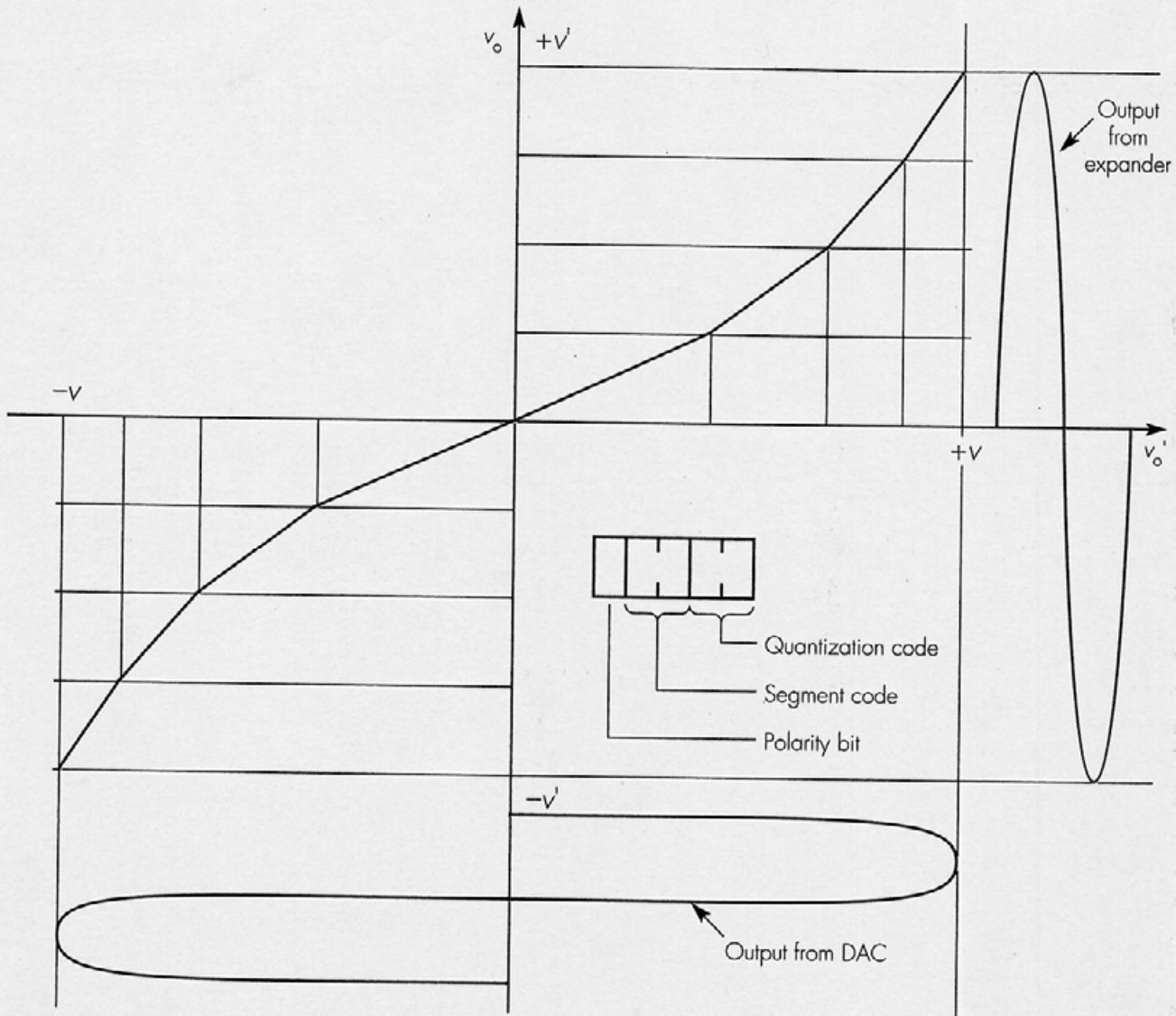


Figure 2.17 PCM principles: (a) signal encoding and decoding schematic; (b) compressor characteristic; (c) expander characteristic.



(c)



Note that in the G.711 standard a 3-bit segment code and 4-bit quantization code are used.

**Figure 2.17 Continued**

# Musique de qualité CD

- Standard CD-DA (*digital audio*)
- Fréquence de Nyquist : 44,1 kHz
- Quantification linéaire sur 16 bits
- Débit stéréo : 1,411 Mbps

# Musique assistée par ordinateur

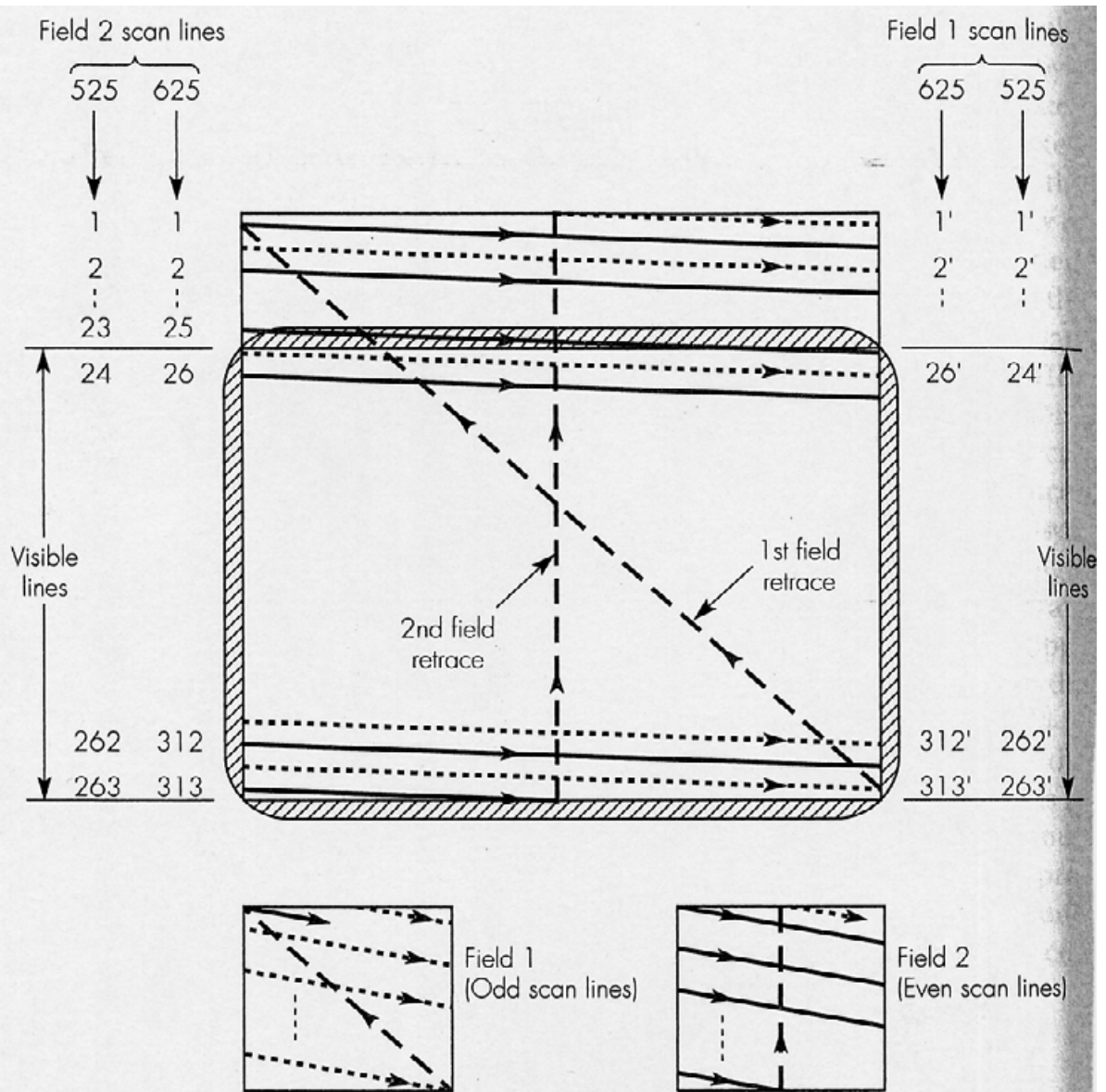
- Standard MIDI (*Musical Instrument Digital Interface*)
- Un fichier MIDI ne stocke pas les sons mais des évènements :  
note jouée, vélocité, durée, etc.
- Trois éléments :
  - Un contrôleur MIDI qui produit des évènements ou messages MIDI
  - Un séquenceur MIDI qui enregistre et / ou joue des évènements MIDI
  - Un générateur de sons qui produit des sons correspondant aux évènements MIDI reçus
- On peut enregistrer au format PCM à la sortie du générateur de sons

# Vidéo

- Plusieurs types de signaux vidéo :
  - Télévision, enregistrements VCR/DVD
  - Vidéo-téléphonie et visioconférence
  - Clips vidéo

# Entrelacement

- 50 images par seconde sont nécessaires pour éviter le scintillement (*flicker*) mais 25 suffisent pour produire un mouvement fluide
- Chaque image est transmise en deux moitiés nommées champs, le premier contient les lignes impaires et le deuxième les lignes paires
- Les deux champs sont intégrés ensemble dans le récepteur en utilisant un balayage entrelacé



525-line systems : 262.5 each field, 240 visible  
 625-line systems : 312.5 each field, 288 visible

**Figure 2.19 Interlaced scanning principles.**

# Signaux couleurs

- Les postes couleurs devaient être compatibles avec les diffusions noir et blanc
- Propriétés d'une couleur :
  - Brillance (*brightness*) : mesure l'énergie stimulant l'œil et variant sur une échelle de gris de noir à blanc et ne dépend pas de la couleur de la source
  - Nuance (*hue*) : représente la couleur effective selon sa longueur d'onde
  - Saturation : mesure la force de la couleur, une couleur pastel est moins saturée que son original
- Luminance = brillance
- Chrominance = nuance et saturation

# Codage des couleurs en télévision

- La luminance vaut :
$$Y_s = 0,299R_s + 0,587G_s + B_s$$
- Elle est utilisée par les postes noir et blanc
- La coloration est codée par les chrominances bleu ( $C_b$ ) et rouge  $C_r$  et valent :
$$C_b = B_s - Y_s \text{ et } C_r = R_s - Y_s$$
- Ces trois signaux sont combinés pour diffusion et forment alors le signal vidéo composite
- Pour éviter une luminance trop élevée en noir et blanc, on réduit les chrominances



# Codage et bande passante des couleurs

- PAL :

$$Y = 0,299 R + 0,587 G + B$$

$$U = 0,493 (B - Y)$$

$$V = 0,877 (R - Y)$$

- NTSC :

$$Y = 0,299 R + 0,587 G + B$$

$$I = 0,74 (R - Y) - 0,27 (B - Y)$$

$$Q = 0,48 (R - Y) + 0,41 (B - Y)$$

- Les chrominances sont placées en bout du spectre de la luminance (elles sont déphasées de  $90^\circ$  et modulées en amplitude et en phase)
- La figure illustre le spectre d'un signal télé couleur en bande de base

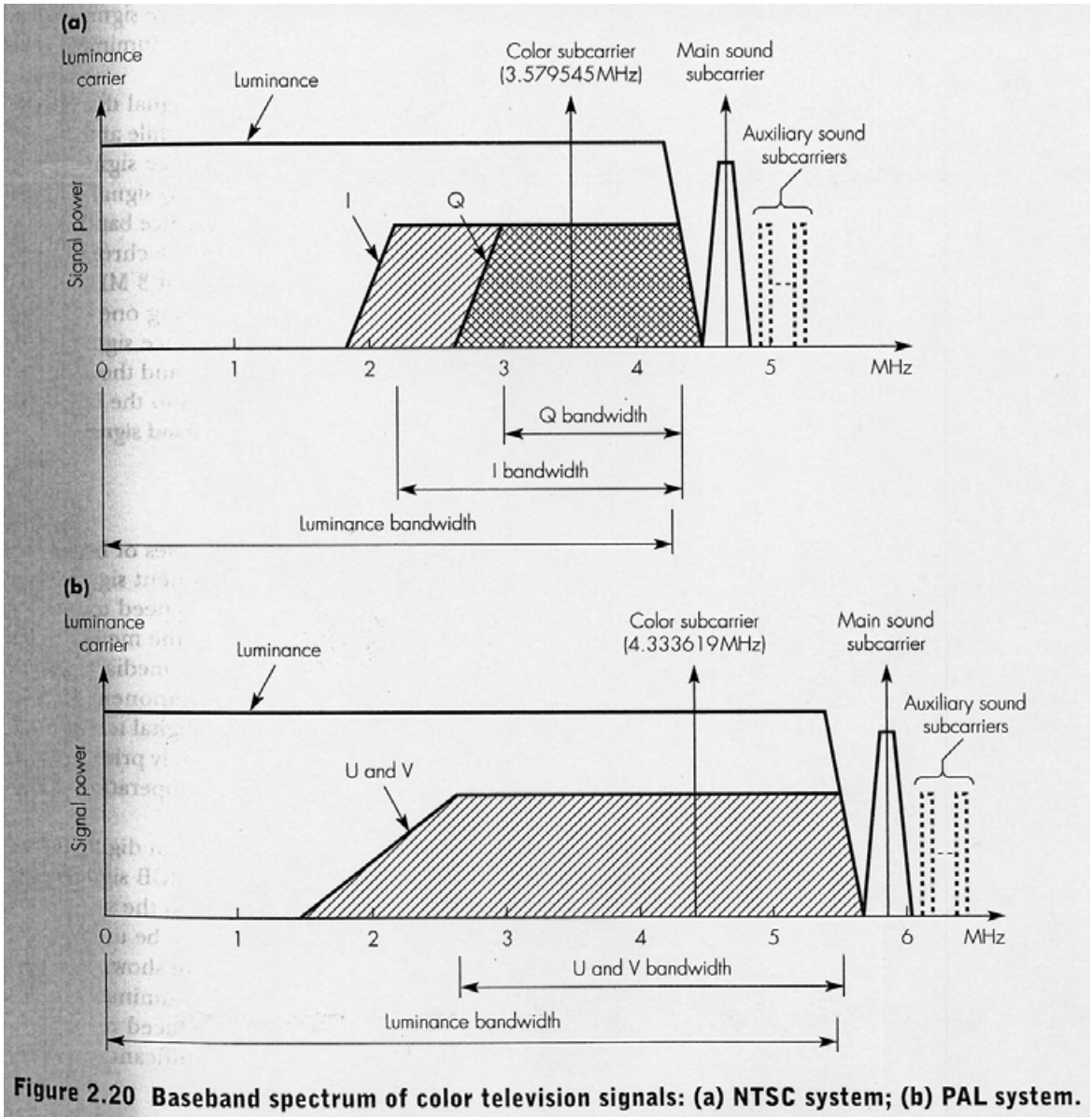


Figure 2.20 Baseband spectrum of color television signals: (a) NTSC system; (b) PAL system.

Caractéristiques	NTSC (National Television Standard Committee)	SECAM (SEquentiel Couleur A Mémoire)	PAL (Phase Alternating Line)
Date de mise en service	1950	1960	1960
Fréquence trame image	60 et 30 Hz	50 et 25 Hz	50 et 25 Hz
Nombre de lignes	525	625	625
Largeur du canal	6 MHz	8 MHz	8 MHz
Largeur de bande vidéo	4.2 MHz	6 MHz	5 MHz
Séparation image/son	4.5 MHz	6.5 MHz	5.5 MHz
Bande latérale atténuée	0.75 MHz	1.25 MHz	0.75 MHz
Modulation de l'image	négative	positive	négative
Modulation du son	FM	AM	FM
Localisation	Amérique du Nord Une partie de l'Amérique du Sud Quelques pays d'Asie (32 pays en tout)	France Pays de l'Est Afrique (42 pays en tout)	Allemagne Brésil Argentine... (63 pays en tout)

# Plan des porteuses

- Bande HF (fréquences de 1,8 kHz à 29 700 kHz)
- Bande VHF (fréquences de 50 MHz à 225 MHz)
- Bande UHF (fréquence de 430 MHz à 2450MHz)
- Bande SHF (fréquences de 3300 MHz à 24,25 GHz)
- Bande EHF (fréquences de 47 GHz à 400 GHz)

# Vidéo numérique

- Signaux numériques RGB séparés ayant la même résolution (i.e. taux d'échantillonnage et bits/échantillons)
- Requiert le triple de bande passante
- L'œil est plus sensible à la luminance, donc on utilise une résolution moindre pour les chrominances pour gagner de la bande passante

# Format 4:2:2

- Format numérique utilisé par les studios de télévision défini par la recommandation CCIR-601 :
  - 6MHz pour la luminance et 3 MHz pour les chrominances
  - Taux d'échantillonnage par ligne de 13,5 MHz (pour avoir un nombre entier d'échantillons par ligne en 525 et 625 lignes) pour la luminance et 6,75 MHz pour les chrominances
- En pratique, 720 échantillons par ligne pour la luminance et 360 échantillons par ligne pour les chrominances
- 8 bits par échantillon, résolution verticale de 480 / 576 lignes
- Échantillonnage orthogonal (points fixes):
  - 525 lignes :  $Y = 720 \times 480$ ,  $C_b = C_r = 360 \times 480$
  - 625 lignes :  $Y = 720 \times 576$ ,  $C_b = C_r = 360 \times 576$

# Format 4:2:0

- Utilisé pour les applications de diffusion vidéo numérique
- Balayage entrelacé et absence d'échantillons de chrominance dans les lignes alternatives
- Échantillonnage orthogonal (points fixes) :
  - 525 lignes / rafraîchissement 60Hz :  
 $Y = 720 \times 480, C_b = C_r = 360 \times 480$
  - 625 lignes / rafraîchissement 50Hz :  
 $Y = 720 \times 576, C_b = C_r = 360 \times 576$

# Autres formats

- HDTV : 16/9, 1920 × 1152, 4:2:2/0
- Source Intermediate Format (SIF) aussi appelé 4:1:1 :
  - 525 lignes / rafraîchissement 30Hz :  
 $Y = 360 \times 240, C_b = C_r = 180 \times 120$
  - 625 lignes / rafraîchissement 25Hz :  
 $Y = 360 \times 288, C_b = C_r = 180 \times 144$
  - Qualité VCR
- Common IF (CIF) :
  - Rafraîchissement 30 Hz :  
 $Y = 360 \times 288, C_b = C_r = 180 \times 144$
  - Utilisation en vidéoconférence
- Quarter IF (QCIF) :
  - Rafraîchissement 15 / 7,5 Hz :  
 $Y = 180 \times 144, C_b = C_r = 90 \times 72$
  - Utilisation en vidéo-téléphonie