

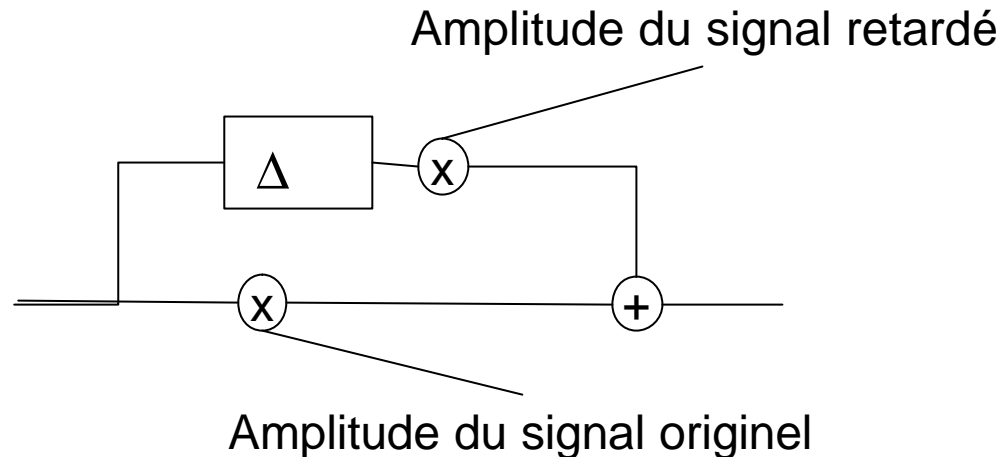
Transformation et codage du son numérique (2)

S. Natkin

Mai 2003

Les effets de retards

Schéma d'une ligne à retard simple

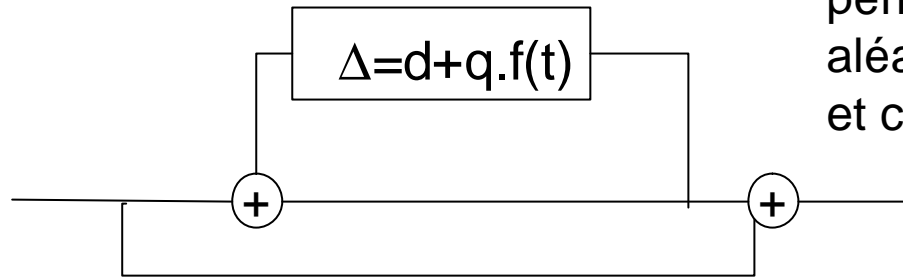


- Court < 10 ms modification du spectre (un ou deux échantillons passe bas FIR, plus long filtre en peigne)
- Moyen > 10 ms et < 50 ms effet d'ambiance, rehausse un signal mince
- Long > 50 ms effet d'écho

Utiliser aussi pour donner un effet de localisation dans l'espace



Flanger



$f(t)$ est une fonction périodique ou une fonction aléatoire de moyenne nulle et comprise entre -1 et +1



Rolling Stone

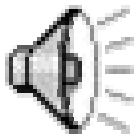
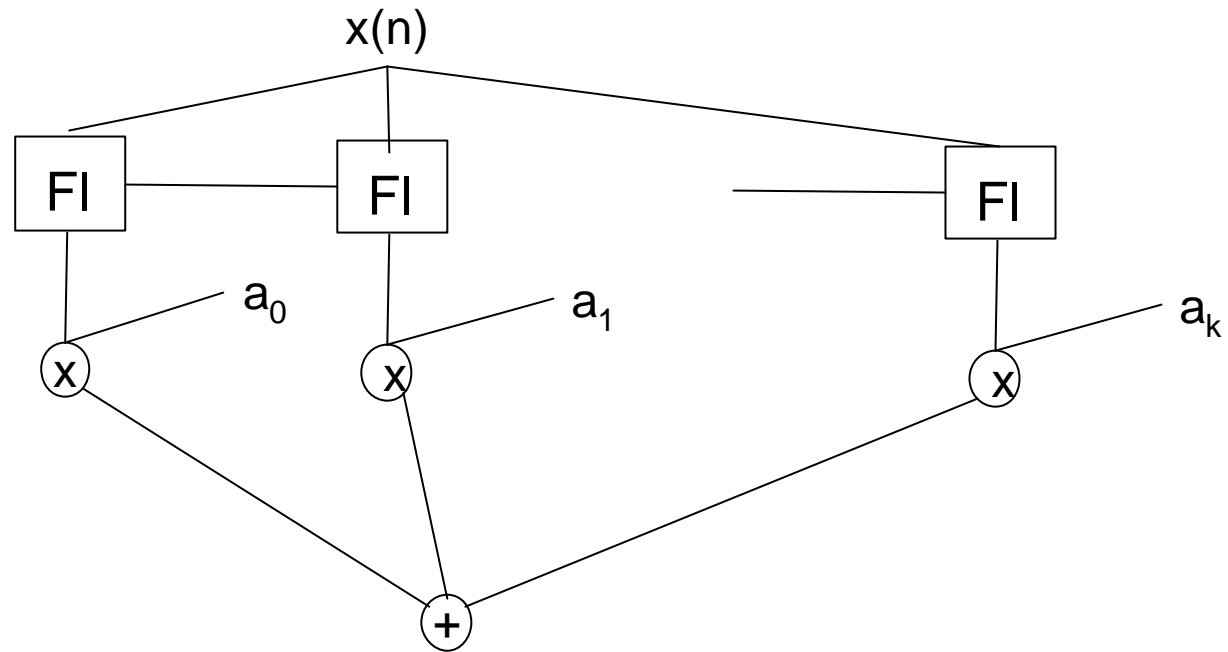


Rolling Stone Flangé

Version plus élaborée le phasing

Effet Chorus

Une
technique

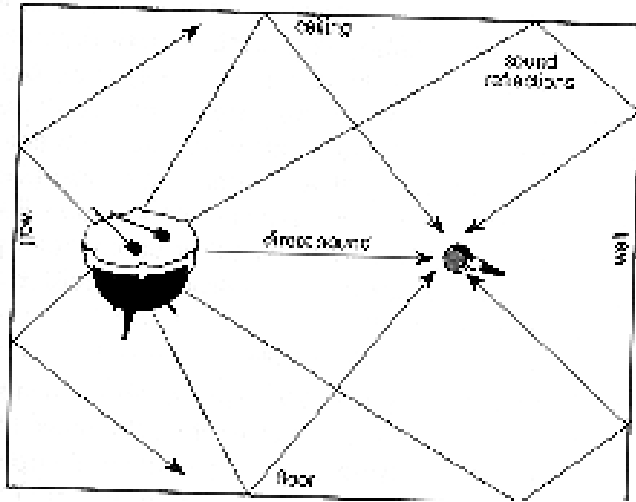


Sans

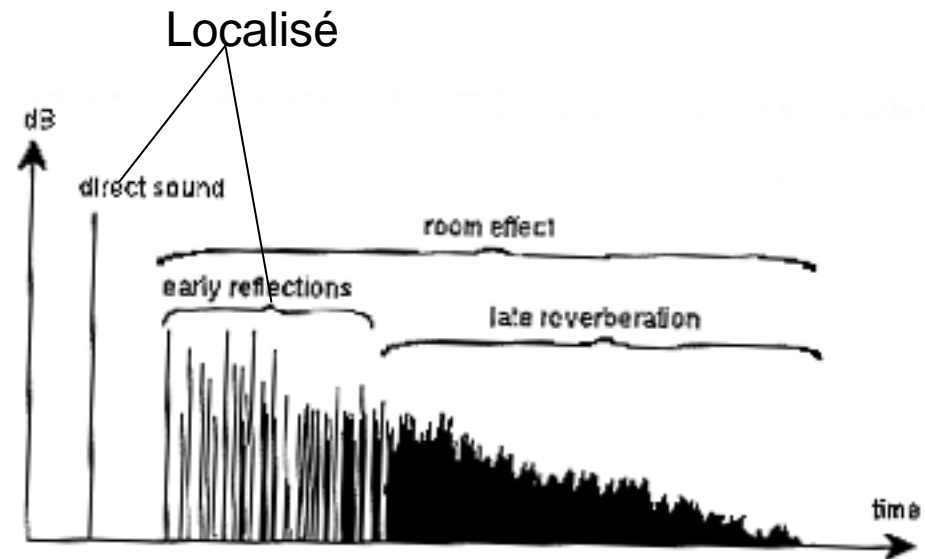


Avec

Spatialisation

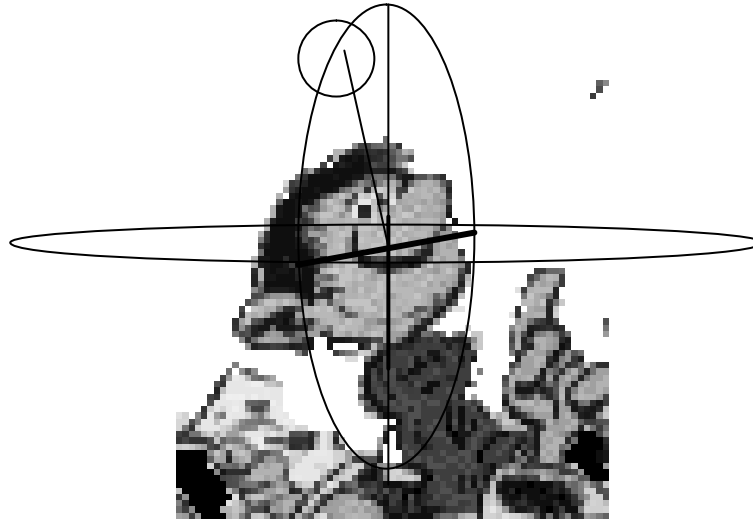


Effet de réverbération
Effet de localisation



Spectre du signal réverbéré (JMJ)

Détermination de la position d'un son dans l'espace



Détermination de l'azimut et zenith

- Différence de temps entre les deux oreilles
- Différence d'amplitude avant arrière (ombre de la tête)
- Différence spectrale liées aux réflexions asymétriques (pavillons, épaules, torses)

Détermination de la distance

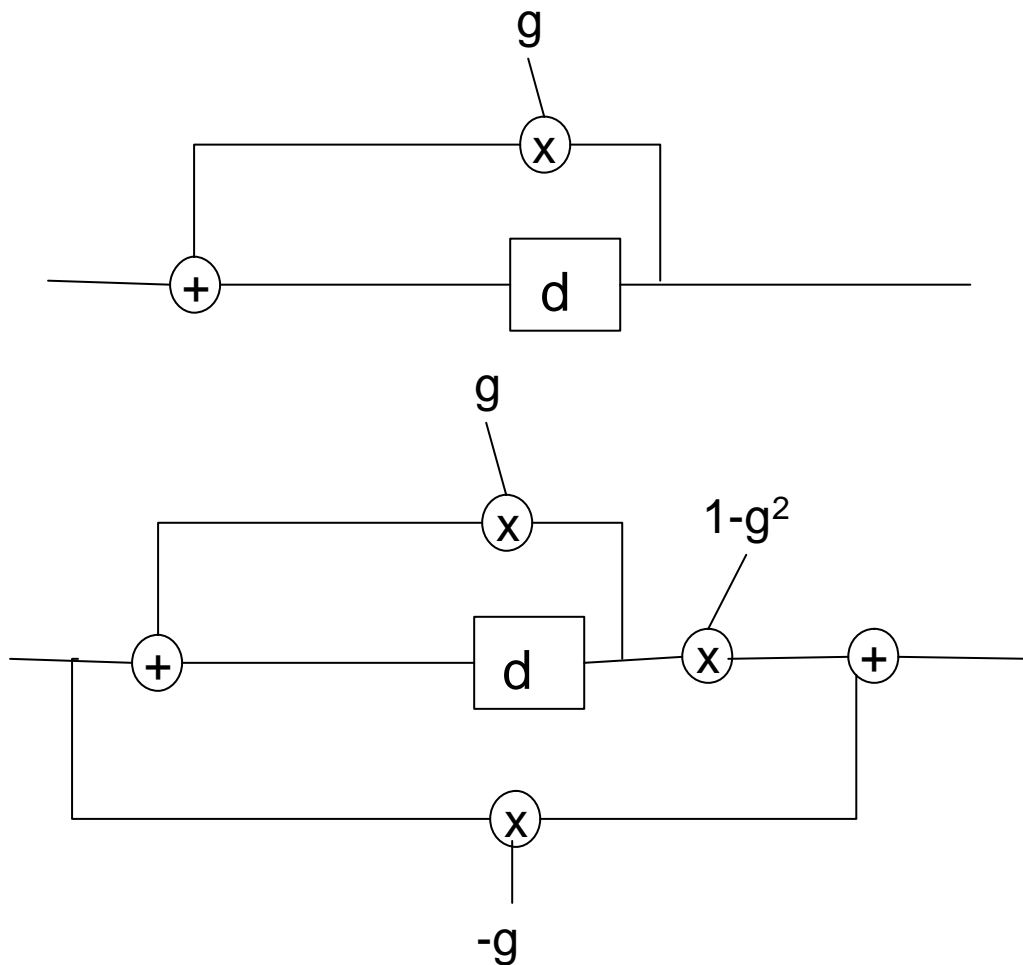
- Rapport entre le signal direct et le signal réverbéré
- Pertes des composantes hautes fréquences
- Pertes des détails

Restitution

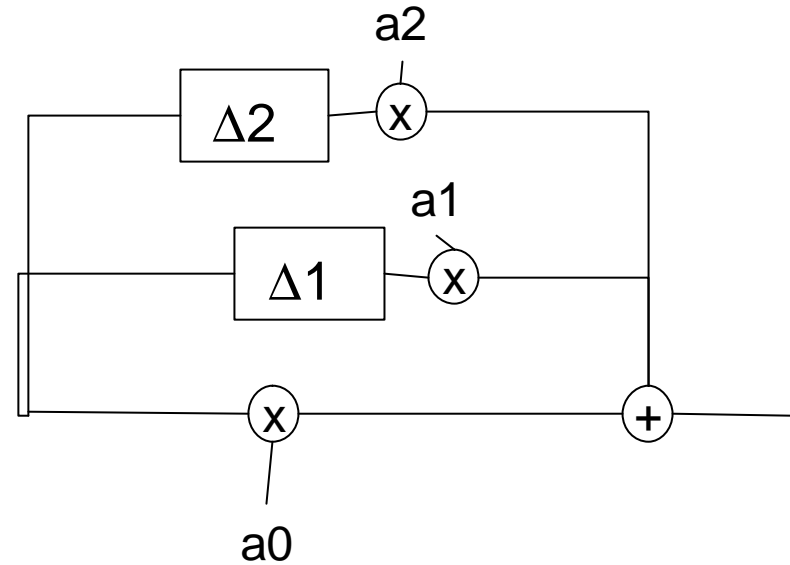
- Mono (distance)
- Stéréo (distance et localisation sur un plan)
- Spatial avec deux hauts parleurs (transaural)
- Spatial avec un casque (binaural)
- Spatial avec N haut parleurs

Modèles complexes de réverbération: Composants de base

Filtre en peigne IIR

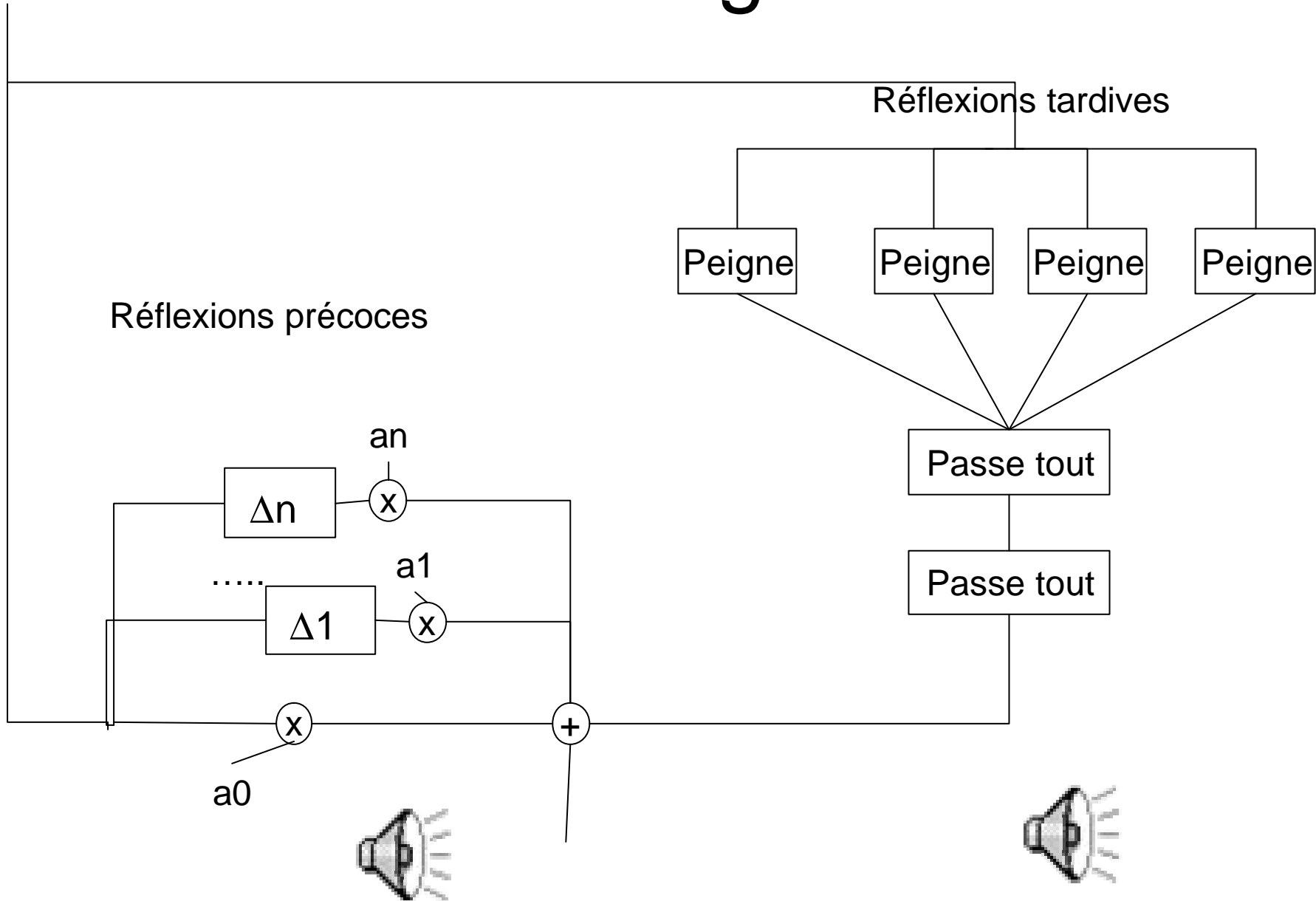


Ligne à retard multibascule

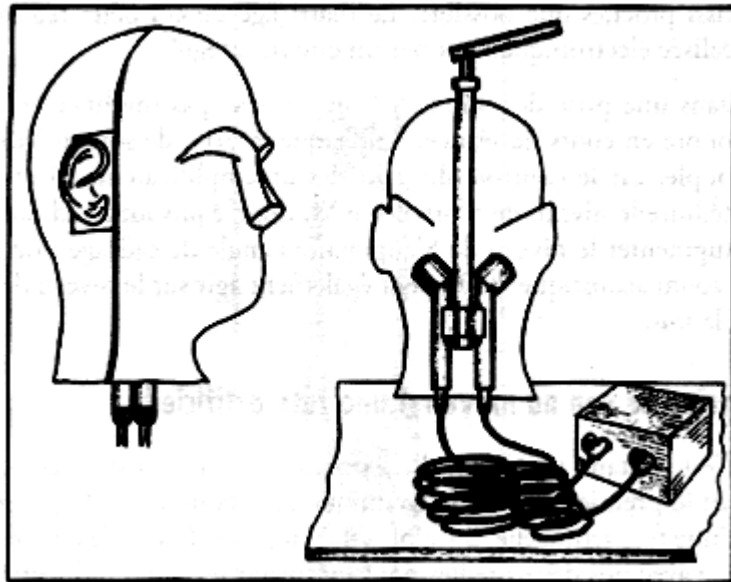


Réseau passe tout du premier ordre

Modèle global

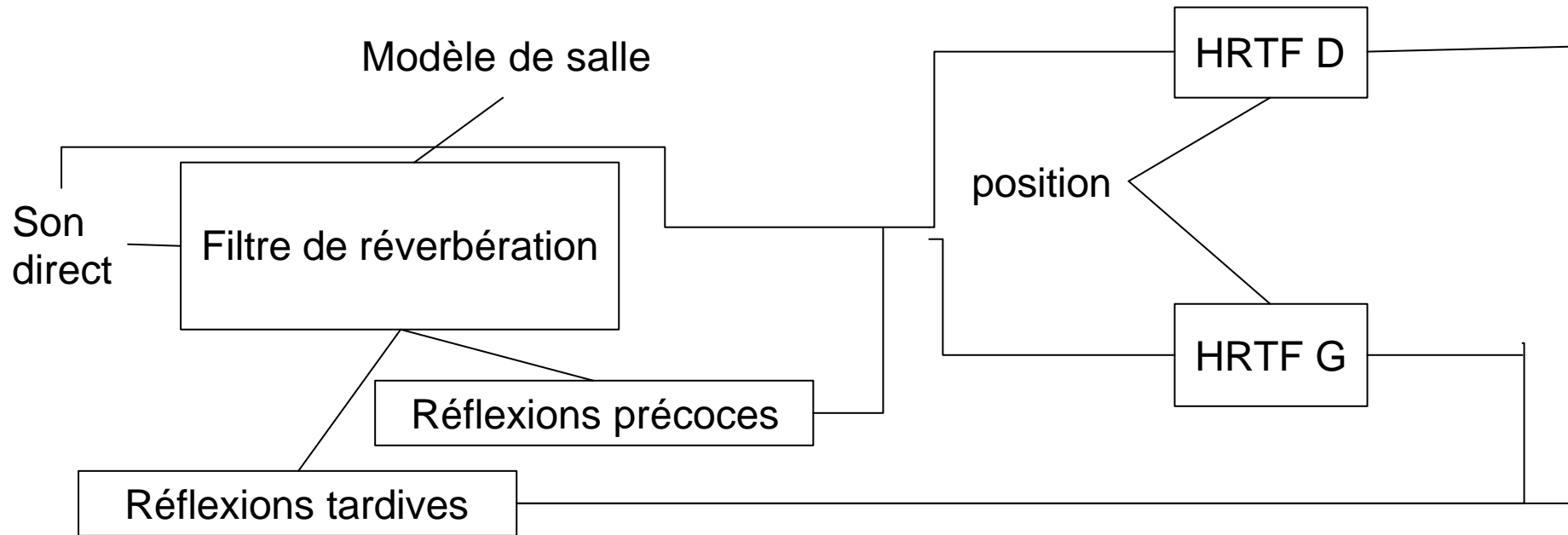


HRTF

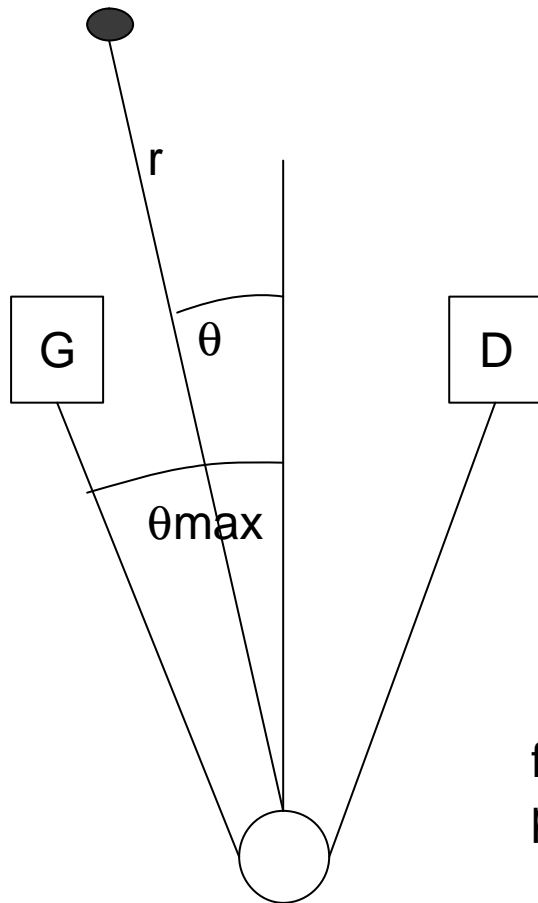


Détermination expérimentale des filtres définissant la différences de perception entre les deux oreilles pour une position d'un objet sonore.
Dépend de la tête...

Exemple de processus de spatialisation binaural



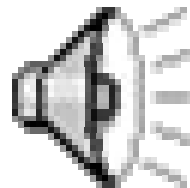
Pan Pot d'intensité (2D)



$$A_G = A \cdot f(r) \cdot \frac{\sqrt{2}}{2} [\cos \mathbf{q} + \sin \mathbf{q}]$$

$$A_D = A \cdot f(r) \cdot \frac{\sqrt{2}}{2} [\cos \mathbf{q} - \sin \mathbf{q}]$$

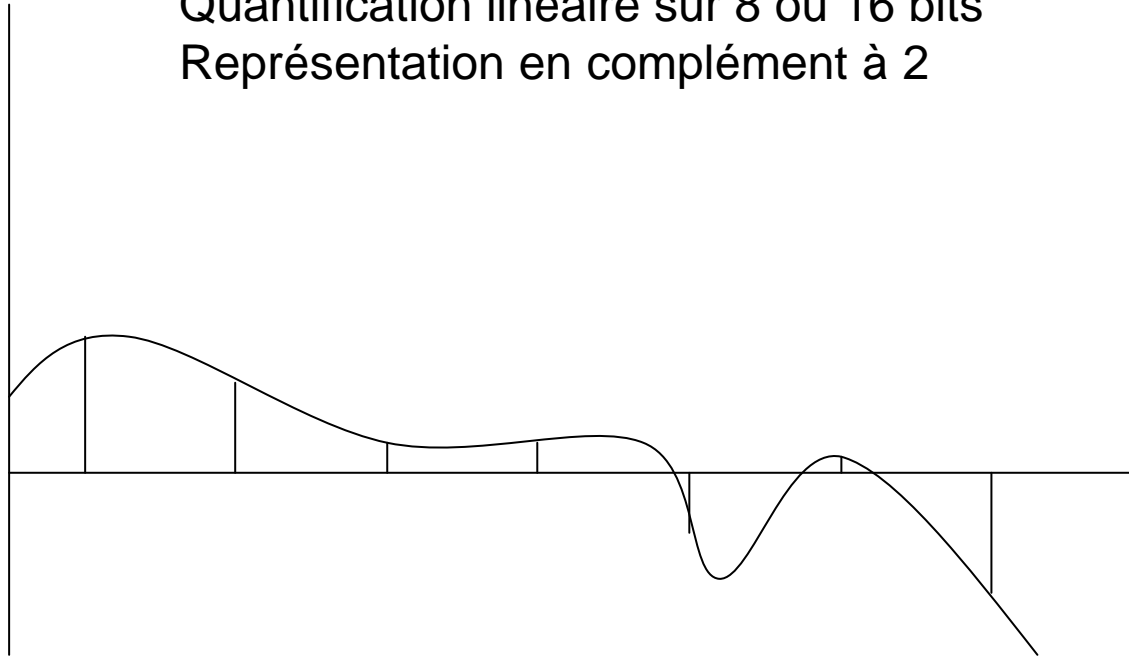
$f(r)$ décroît avec la distance. Peut être différente pour le son direct, les réflexions précoces et tardives



Codage du son numérique

Pulse Code Modulation

Représentation temporelle du signal échantillonné
Fréquences d'échantillonnage
de 4Khz à 44Khz
Quantification linéaire sur 8 ou 16 bits
Représentation en complément à 2



PCM différentiel

DCMP codage différentiel: on code $x(0)$ et

$$d_n = x(n) - x(n-1)$$

Efficace si le signal varie lentement.

Par exemple si $x(n)$ est sur 16 bits on code

la longueur de $d(n)$ sur 3 bits puis d_n

Efficace si d_n tient sur moins de 13 bits

Adaptative PCM : ADPCM

Principe: coder la variation des différences. Par exemple (schéma IMA)

$$x(n-1) = x(n-2) + d_{n-1} \text{ avec } d_{n-1} = D(k)$$

Notons $i_{n-1} = k$

$$x(n) = x(n-1) + d_n \text{ avec } d_n = D(t)$$

$i_n = t$

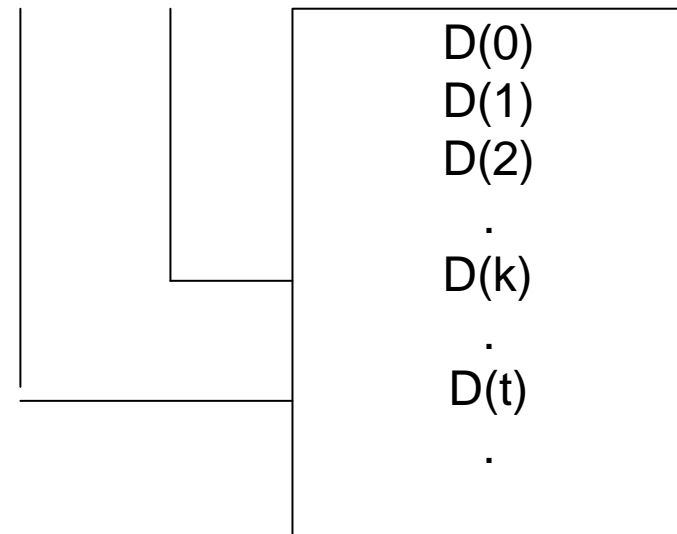
On code

$x(0)$ sur 16 bits

i_1 sur 8 bits

$d_i = i_n - i_{n-1}$ sur 4 bits

Table des différences D



Compression psycho perceptive

- La perception du son est logarithmique

Utiliser une quantification logarithmique

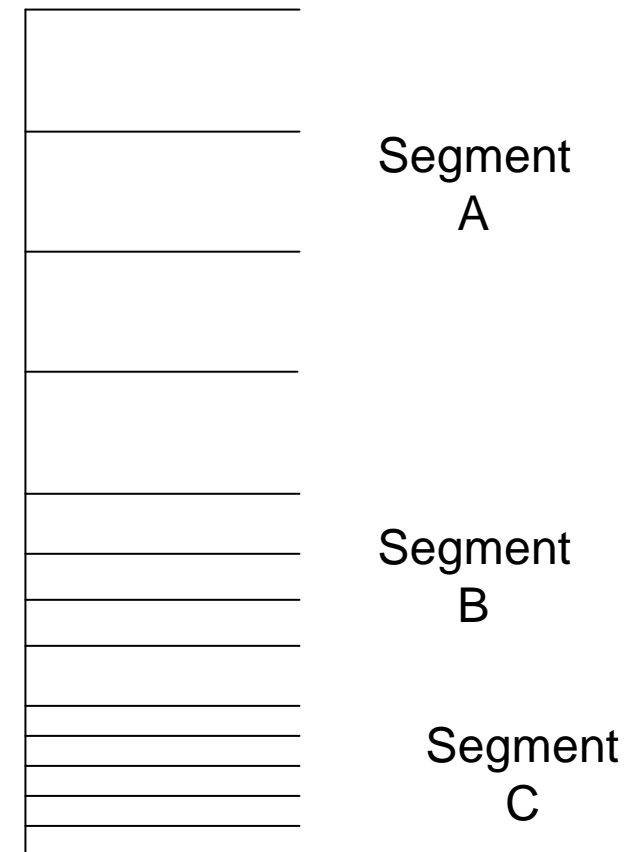
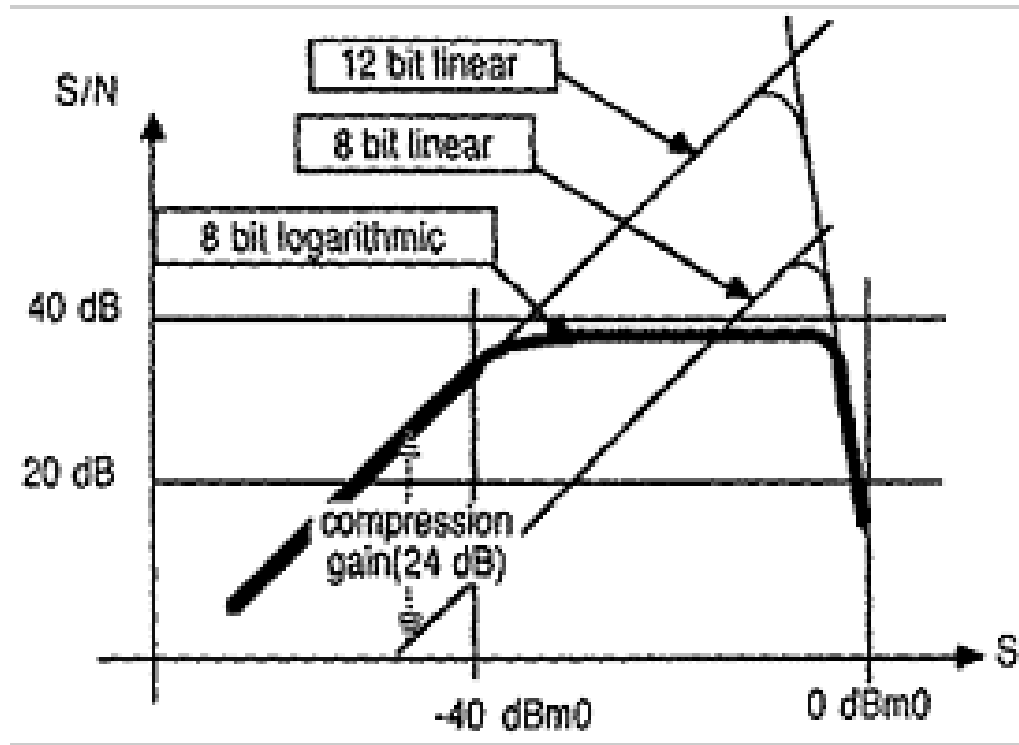
- L'oreille distingue mieux les sons dans la bande 300 – 6000 Hz

Echantillonner différemment selon les bandes de fréquence

- Certains sons en masque d'autre

Supprimer les informations non perçues

Quantification perceptive: lois de quantification mu et A



Mu: 8 segments de 16 Valeurs,
incrément de 2 de à 256
A : 7 segments le premier de
32 valeurs et les autres de 16
valeurs

Notion de bande critique

- Dans un intervalle de fréquence autour d'une fréquence centrale, l'oreille ne distingue pas nettement deux sons.
- Si un son a une intensité forte il masque, pendant une courte période les sons d'intensité plus faible dans sa bande critique
- La notion de bande critique est liée à un effet de filtrage en peigne à Q constant effectué par l'oreille
- En première approximation il est possible de décomposer un son audible sur une vingtaine de bande de fréquences (pendant du RVB visuel)

Compression fréquentielle perceptive (ex MP3)

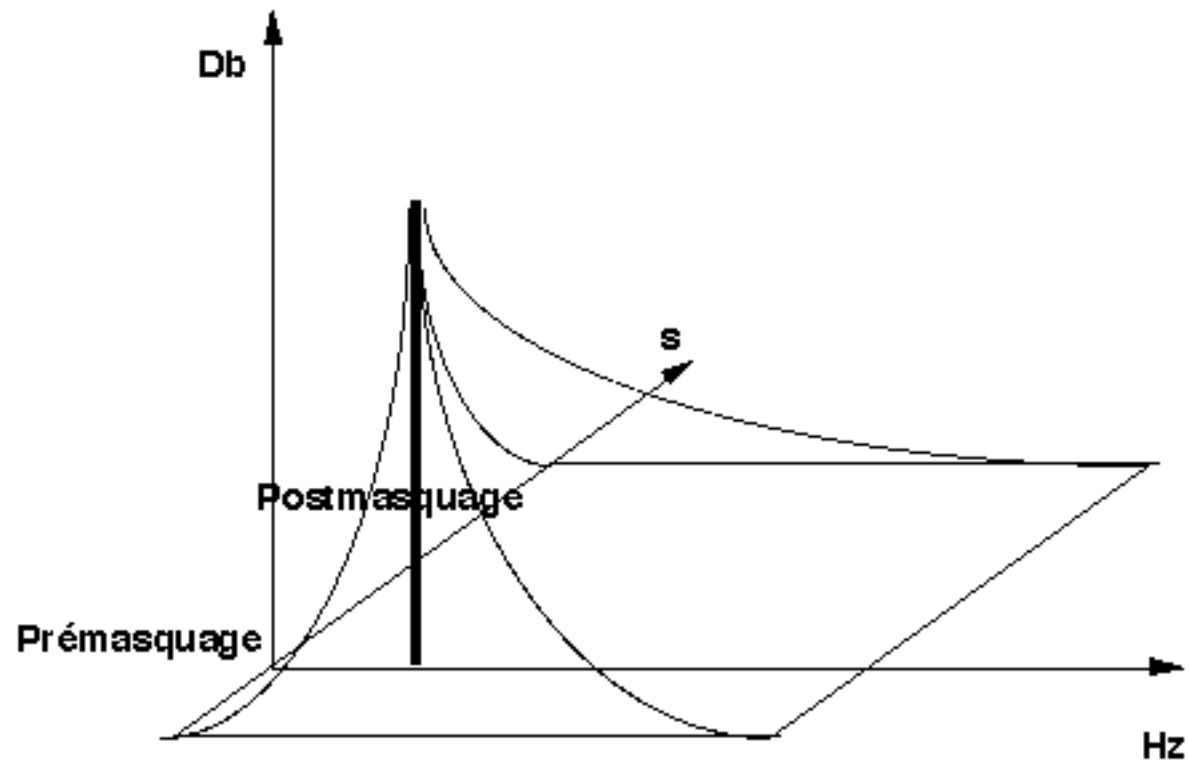
MPEG 1 est une norme de transmission de vidéo numérique (Image animée+ son) intégrant différentes qualités selon les types D'application (visiophonie...Video a la demande)

Le son est échantillonné a 32, 44,1 kHz ou 48 kHz

MPEG 1 Layer 3 (MP3) est un codage du son , il s'agit d'une compression non conservative basée sur le masquage

3 niveaux avec des codeurs de plus en plus complexes en
Fonction de la qualité de son visée

Utilisation des effets de masquage



Utilisation des effets de masquage (2)

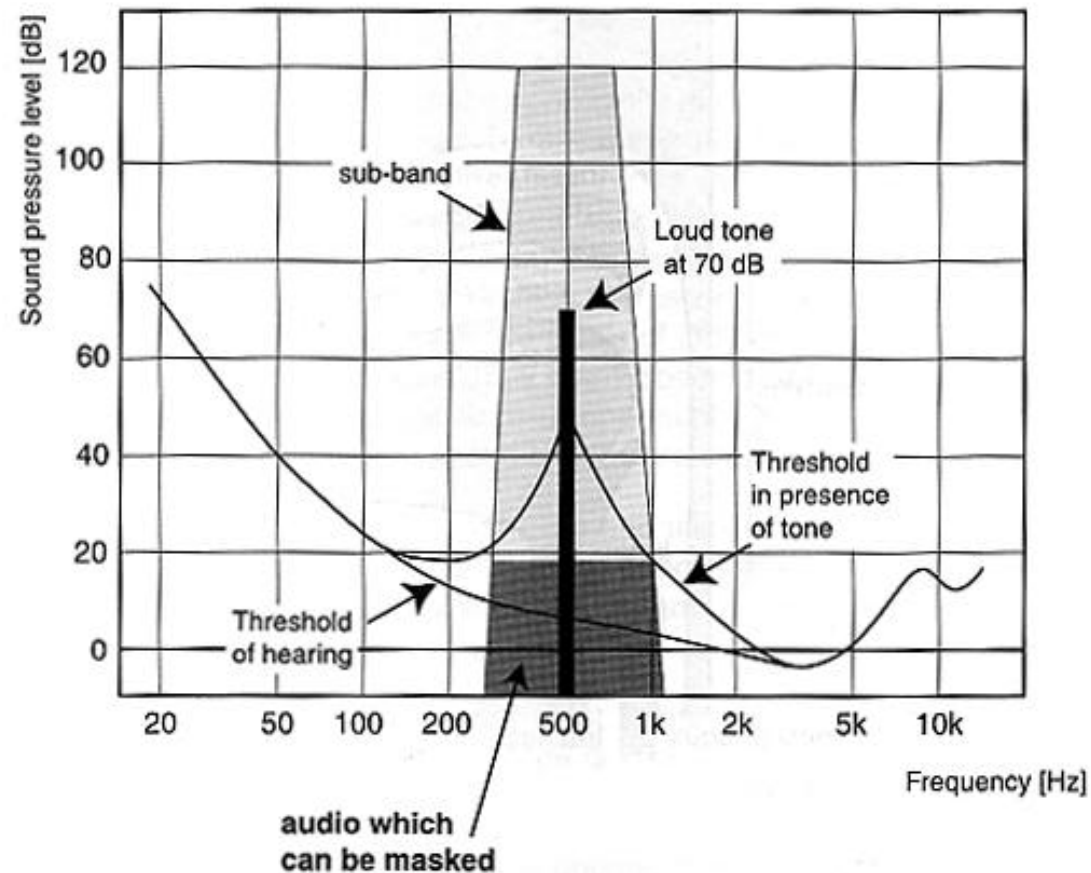
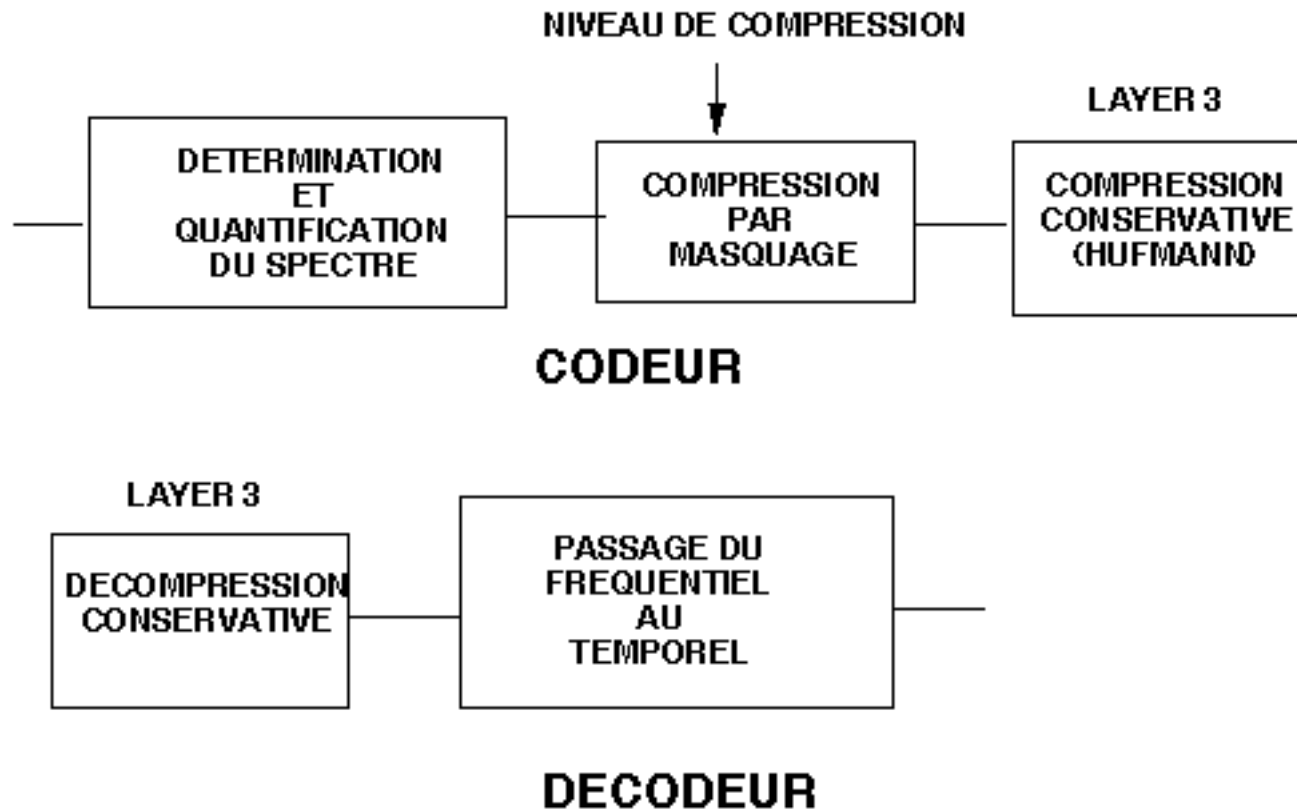


Figure 6.6 Noise masking threshold.

Codeur MP3



Compression MPEG1

Table 6.6 Summary of MPEG audio

	<i>Bit rate range (Kbit/s)</i>	<i>Target bit rate (Kbit/s)</i>	<i>Typical compression ratio</i>
Layer 1	32–448	192 (mono)	1:4
Layer 2	32–384	128 (mono)	1:6 to 1:8
Layer 3 (MP3)	32–320	64 (mono) 128 (stereo)	1:10 to 1:12

Compression dynamique de la voix

Mode de compression utilisé pour un traitement en temps réel de la voix sur un canal à faible débit (GSM, RealAudio...)

Basé sur un modèle acoustique de la gorge qui permet de caractériser le son initial et l'effet de résonance (formants)

LPC: Linear Predictive coding

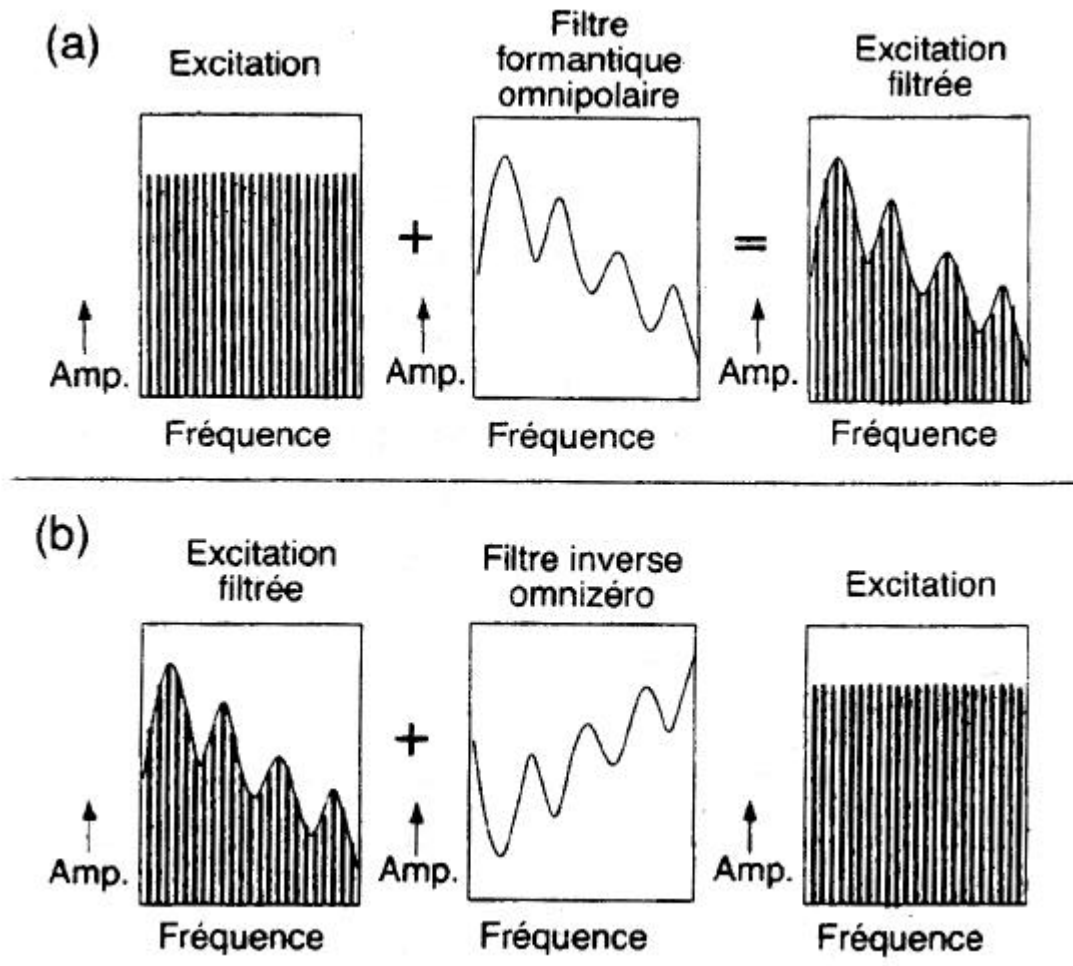
Les formants sont prédits par un filtre linéaire adaptatif qui induit le formant courant par rapport aux précédents.

Une fois le formant calculé il est soustrait du signal initial ce qui donne le résidu qui doit en théorie contenir le son direct

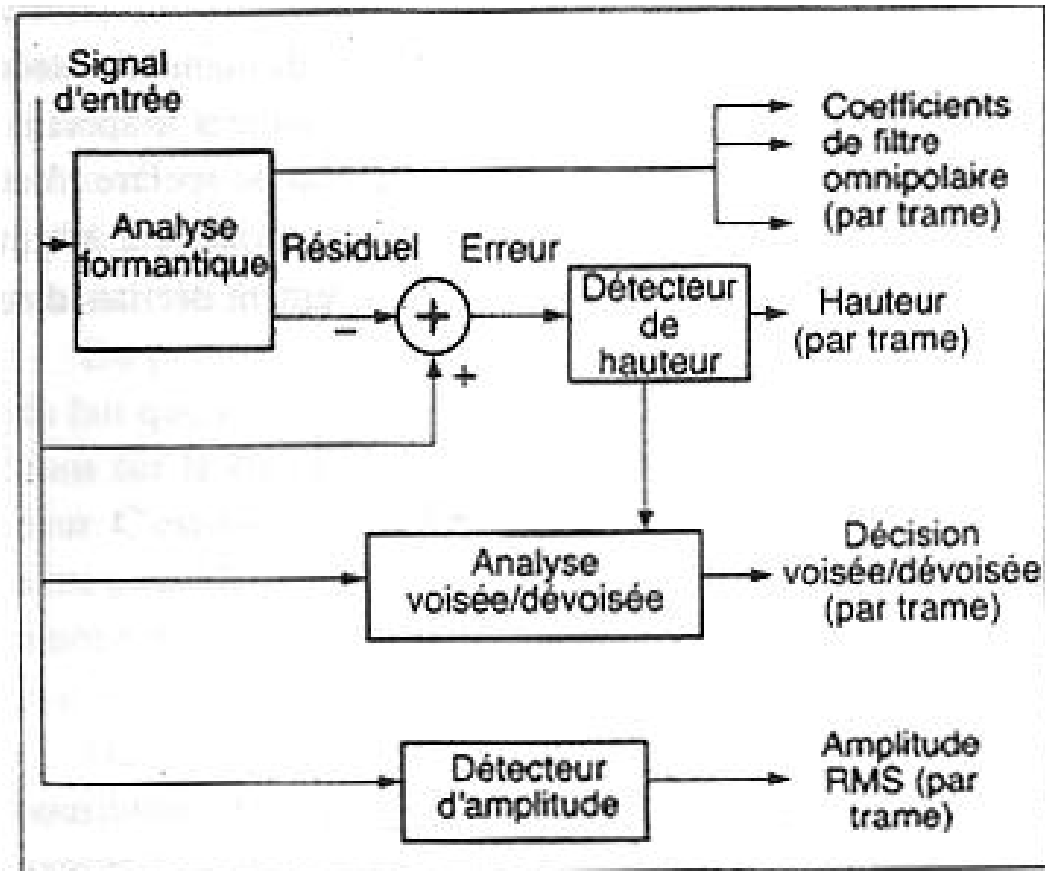
En pratique il faut transmettre une partie du formant ou une référence à un formant voisin

Amélioration le CELP

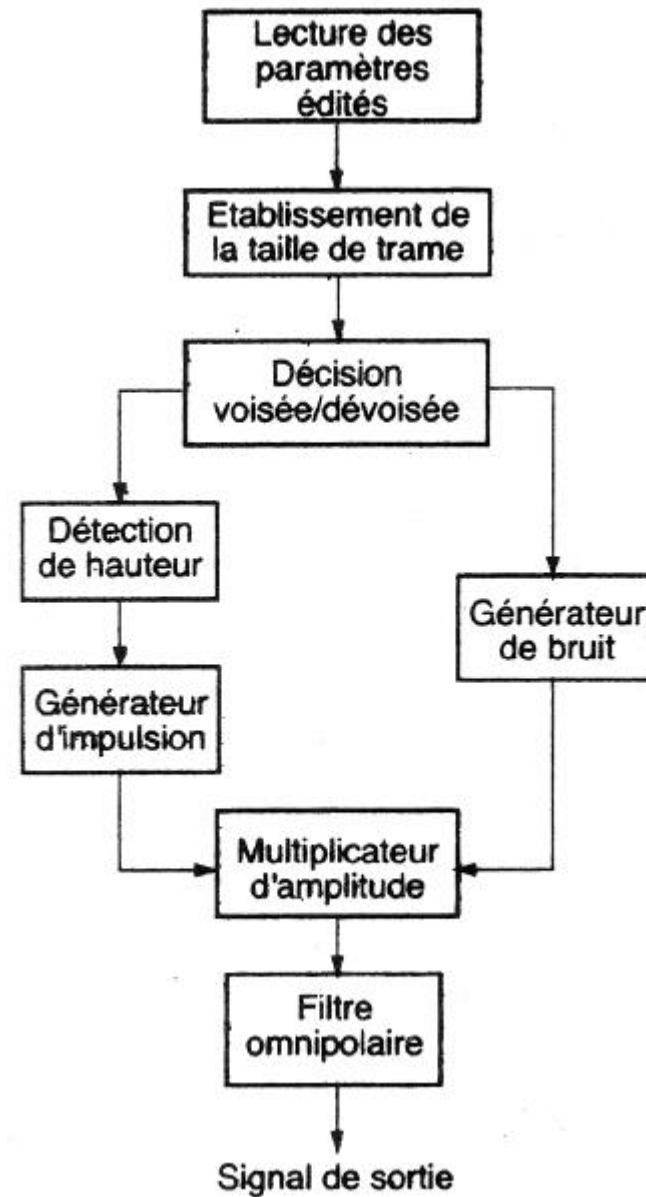
Filtre formantique



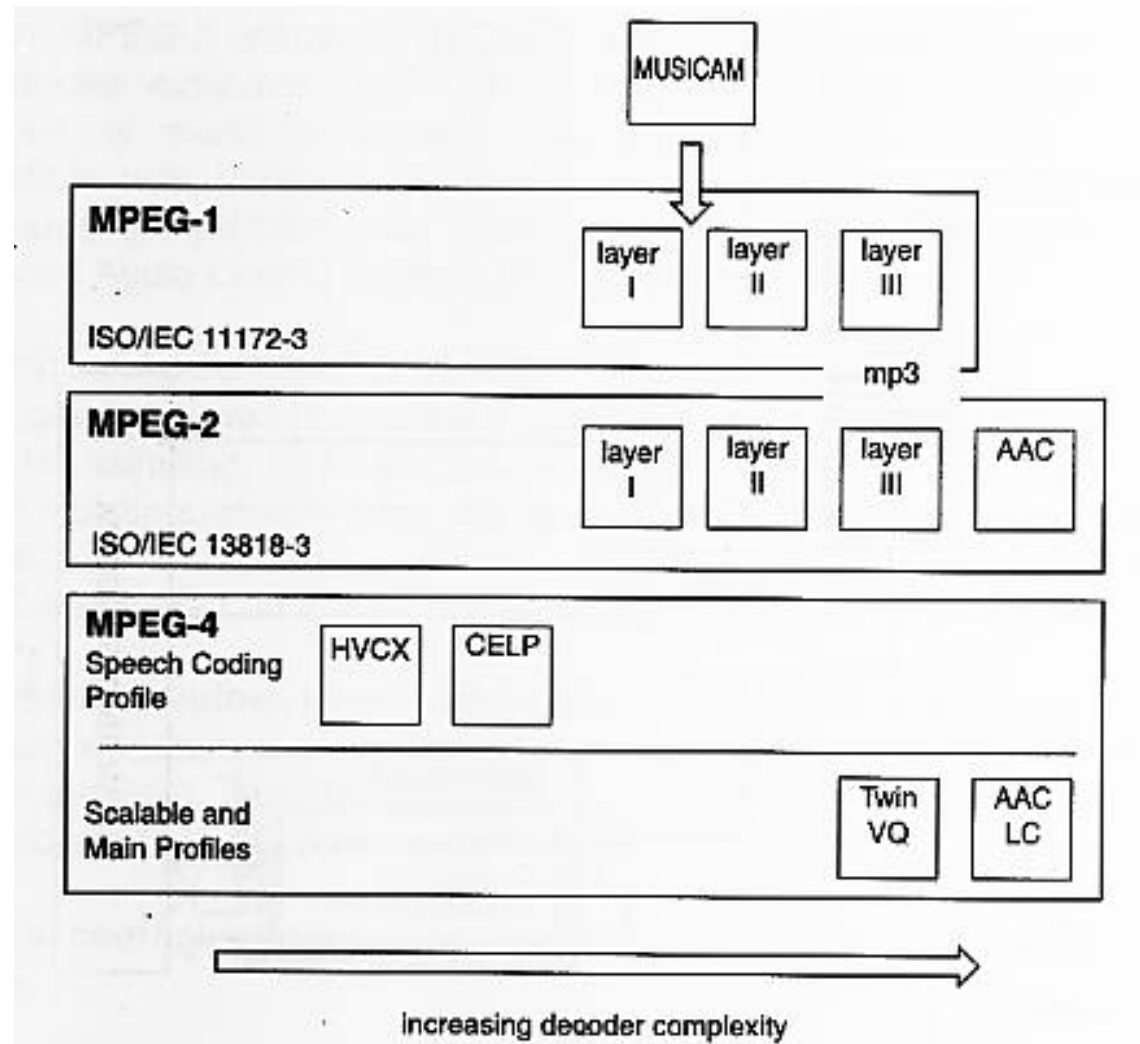
Processus de codage



Processus de décodage



MPEG Audio



Codec MPEG 4

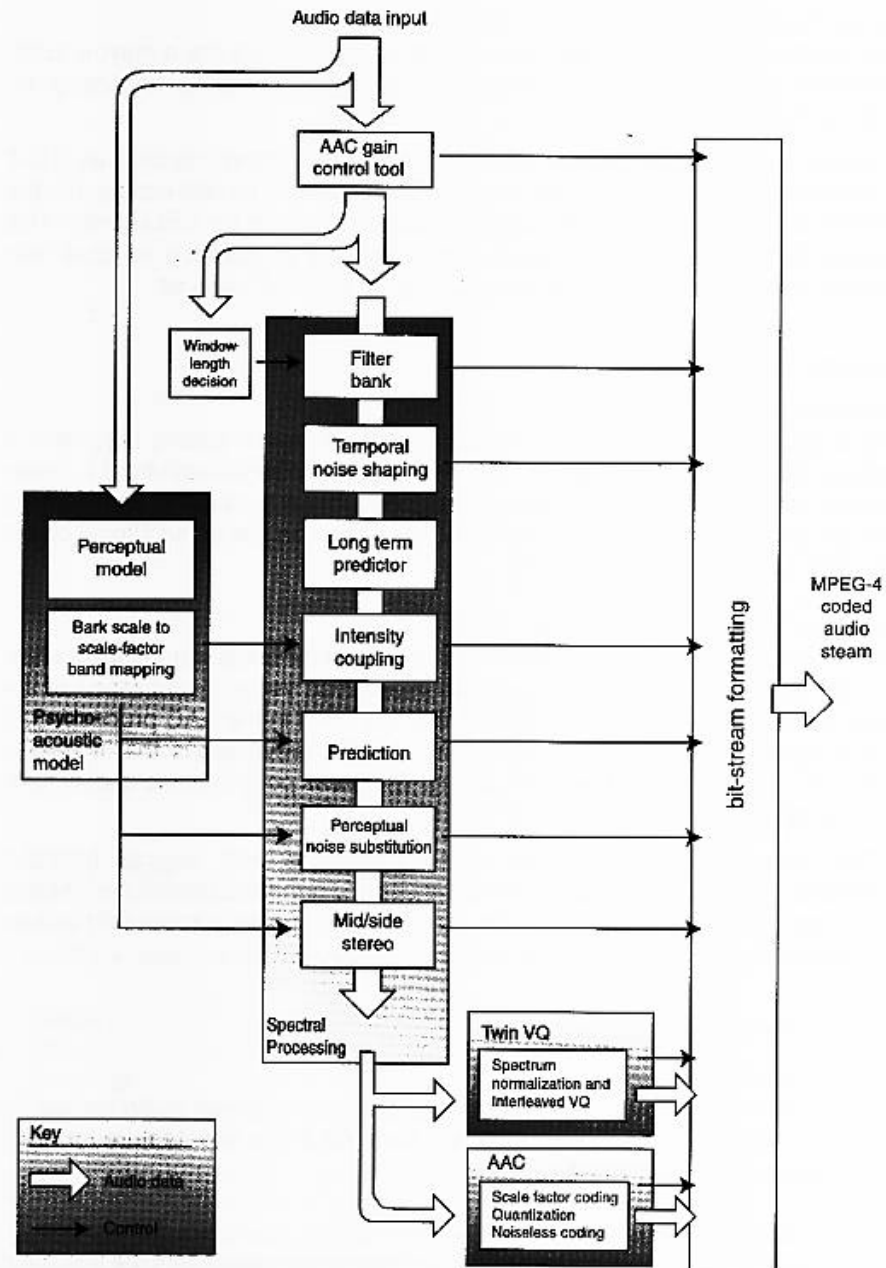


Figure 6.15 MPEG-4 general audio coder.

Débits MPEG 4

Data rate

64 kbits/s

48 kbits/s

Sampling rate

24 kbits/s

18 kbits/s

16 kHz

6 kbits/s

4 kbits/s

8 kHz

2 kbits/s

1.2 kbits/s

vbr : narrow : wide
HVCX
CELP

speech

Twin VQ
AAC

general audio

Figure 6.14 MPEG-4 audio coding algorithms.

Formats Audio en général

Table 6.5 Summary of codec algorithms

<i>Organization/family</i>	<i>Algorithm</i>	<i>Codec</i>
MPEG		MPEG-1, layers 1, 2, 3 MPEG-2, layer 2 and AAC MPEG-4, layer
ITU-T	PCM ADPCM	G.711 PCM (μ -law, A-law) G.726 ADPCM at 16, 24, 32 and 40 Kbit/s (replaces G.721 at 32 Kbit/s and G.723 for 24 and 40 Kbit/s)
Interactive Multimedia Association	ADPCM	IMA/DVI
US DOD Federal Standard	Linear predictive CELP	LPC-10E (1015) (1016)
ETSI	Linear predictive	GSM 06.10 (global système mobile)
Proprietary	ACELP ACELP ATRACS ADPCM	Voiceage Sipro Sony ATRAC3 Microsoft ADPCM

Bibliographie

- L'audionumérique, Curtis Raod, Dunod 1998
- The technology of video and audio streaming, David Auterberry, Focal Press, 2002
- Le son musical, John Pierce, Pour la Science, Belin 1984