



# Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG-2 Audio

José M. Martínez  
Escuela Politécnica Superior  
Universidad Autónoma de Madrid, SPAIN

JoseM.Martinez@uam.es  
tel:+34.91.497.22.58

2011-2012



# Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

## INDICE

- Introducción
- Audio MPEG-1
- Audio MPEG-2
- Audio MPEG-4
- Audio multicanal Dolby AC-3

## Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

### INDICE

- **Introducción**
- Audio MPEG-1
- Audio MPEG-2
- Audio MPEG-4
- Audio multicanal Dolby AC-3

### Introducción

#### Principios de la codificación de audio MPEG (natural)

- **Entrada Audio PCM a diversas frecuencias de muestreo**
- **Codificación subbandas**
  - o Imita el mecanismo de análisis frecuencial del oído. Modelo de banco de filtros.
  - o La señal se pasa a un dominio espectral
  - o Codifica cada banda con diferente resolución (calidad)
- **Codificación perceptual**
  - o Utiliza un modelo psico-acústico
    - Enmascaramiento auditivo (depende del sonido codificado)
    - Umbral de audición (subjetivos)
    - Se calcula en un dominio transformado (FFT, MDCT –Modified DCT, overlapped-)
  - o Idea: **MENOS bits** (o ninguno) para los sonidos **MENOS audibles**. Se codifica con menos bits con la misma calidad de audio percibida (calidad perceptual)
  - o Tiene que analizar la señal continuamente para determinar el umbral de audibilidad en cada instante
- **Otras tecnologías utilizadas**
  - o Conmutación de ventanas
  - o Asignación dinámica de bits

## Introducción: estándares de audio (I)

### MPEG-1 Audio (parte 3)

- Tres niveles de diversa complejidad
- MPEG-1 layer III: mp3

### MPEG-2 Audio (parte 3)

- Soporta el audio MPEG-1
- Lo extiende a multicanal (hasta 5 canales)
- Es lo normativo (obligatorio) en DVB

### MPEG-2 Advanced Audio Coding (AAC – parte 7)

- Más eficiente y con más calidad que MPEG-1 Layer 3
- No compatible hacia atrás

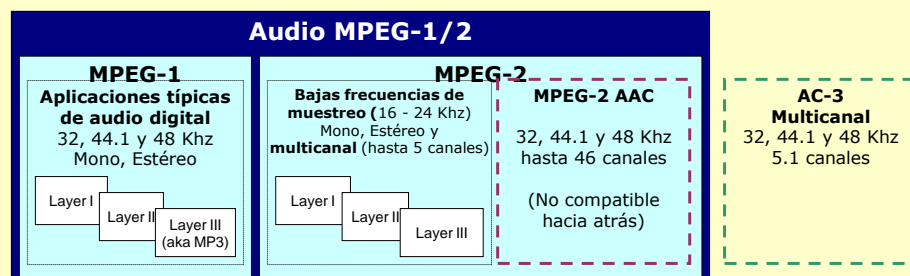
### MPEG-4 Audio

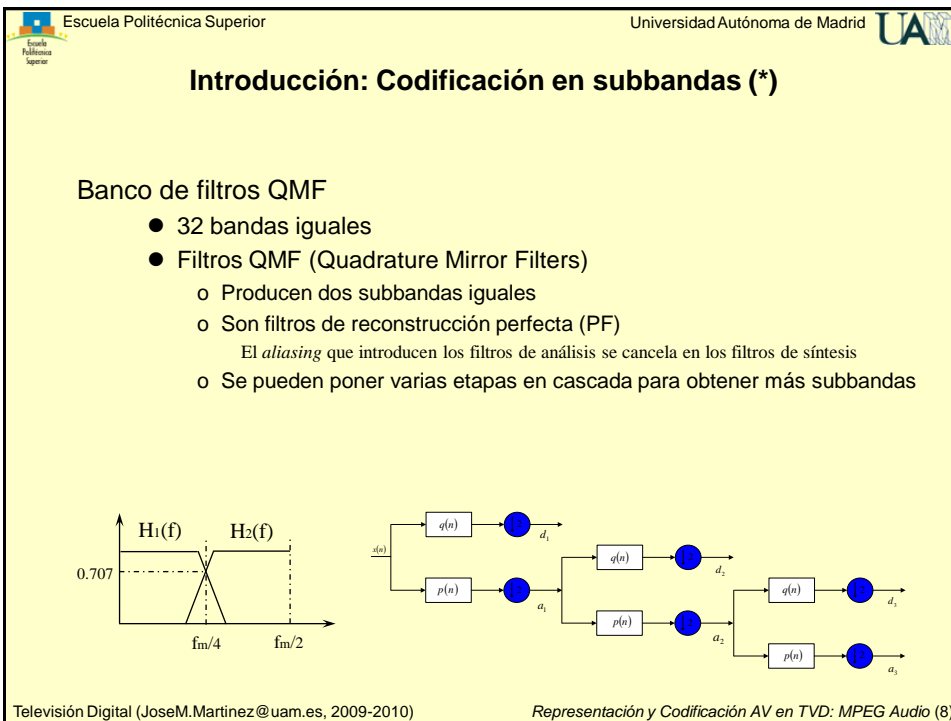
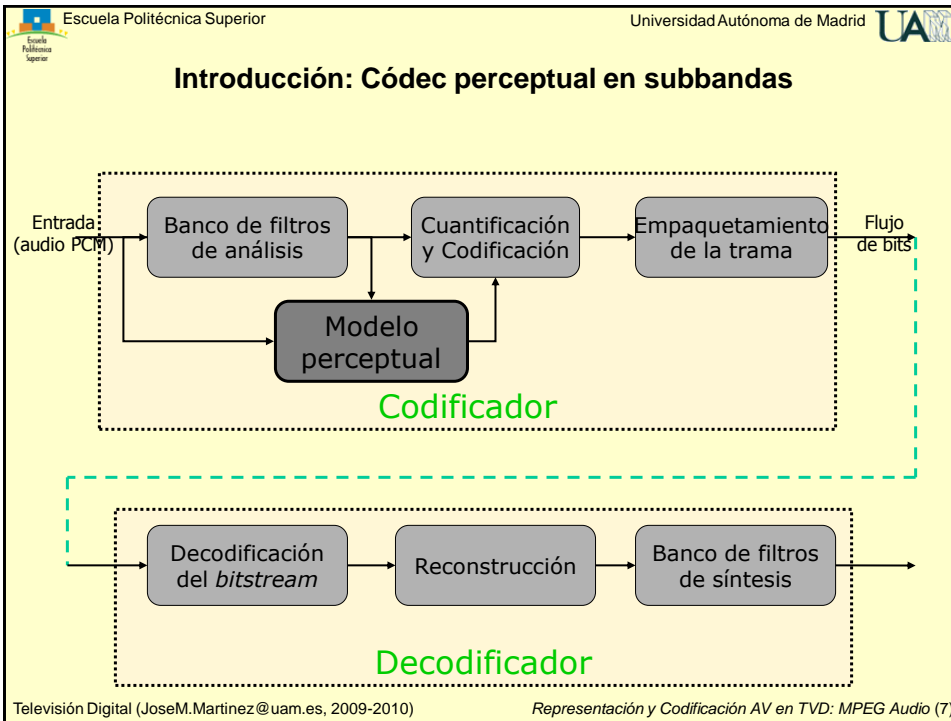
- Diferenciación según tipo de audio (audio natural, voz, audio sintético, voz sintética)

### Audio multicanal Dolby AC-3

- No es parte de MPEG-2, pero MPEG-2 systems si lo soporta
- Usado en DVD y ATSC

## Introducción: estándares de audio (II)





Escuela Politécnica Superior Universidad Autónoma de Madrid

## Introducción: Modelo psicoacústico

Modela los umbrales de enmascaramiento

- Calcula la relación señal a máscara (SMR) en cada banda
  - Subbanda/banda crítica
  - Umbral de audición con FFT, MDCT, ...
- Relación ruido a máscara (NMR = SMR - SNR, en dB)

Utilidad

- Todo aquello que se encuentra por debajo del umbral de enmascaramiento es inaudible. Ventajas:
  - Bien no lo transmito (no es sencillo)
  - O bien lo aprovecho para meter ruido de cuantificación (es decir, utilizo **MENOS BITS** para codificar un cierta banda)
    - $SNR(m) = SNR$  (de un cuantificador de m bits)
    - $NMR(m) = SMR - SNR(m)$  en dB
- Conclusión: la precisión del modelo psicoacústico que utilice determinará el nivel de compresión del codificador

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG Audio (9)

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Introducción: Modelo psicoacústico

Modela los umbrales de enmascaramiento

- Calcula la relación señal a máscara (SMR) en cada banda

Utilidad

Diagram description: The diagram plots Sound Pressure Level (in dB) on the y-axis against Frequency on the x-axis. A central peak represents the 'Masker' tone. The area under this peak is divided into 'Critical Band' and 'Neighboring Band'. A 'masking threshold' curve is shown, which is higher than the masker tone in the neighboring band. The 'Signal-to-Mask Ratio (SMR)' is the vertical distance between the masker tone and the masking threshold in the critical band. The 'Noise-to-Mask Ratio (NMR)' is the vertical distance between the noise level of an m-bit quantizer and the masking threshold in the neighboring band. The 'Umbral mínimo de Enmascaramiento NMR = 0' is indicated at the level of the noise level. Callouts provide further details: 'Tono enmascarante' points to the masker peak; 'Umbral de enmascaramiento' points to the masking threshold; 'Umbral mínimo de Enmascaramiento NMR = 0' points to the noise level; and 'Nivel de ruido de cuantificación con m bits' points to the noise level of the m-bit quantizer.

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG Audio (10)

## Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

### INDICE

- *Introducción*
- **Audio MPEG-1**
- Audio MPEG-2
- Audio MPEG-4
- Audio multicanal Dolby AC-3

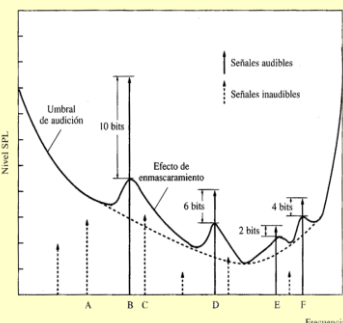
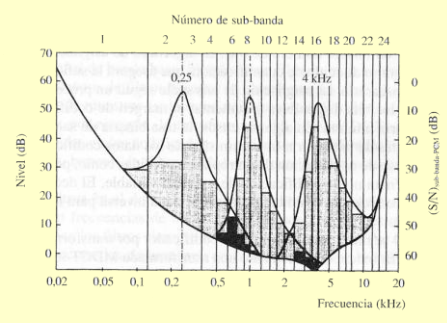
## Audio MPEG-1 (I): Modelo psicoacústico (\*)

Modelo de enmascaramiento (sigue codificador perceptual de Johnston 19988)

- Dos tipos de enmascaramiento
  - Tono enmascarando otro tono (umbral  $\approx -14.5 - P_{\text{tono}}$  dB)
  - Tono enmascarando ruido (umbral  $\approx -5.5$  dB)
- Proceso de cálculo del umbral de enmascaramiento
  - 1) Análisis en bandas críticas
    - Se calcula la potencia de cada banda a partir de coeficientes de transformadas (FFT, MDCT)
  - 2) Función de dispersión
    - Para emular la respuesta auditiva en la membrana basilar
    - Permite ver el enmascaramiento en las bandas críticas próximas
  - 3) Umbral enmascarado
    - Distinguiendo entre enmascaramiento por tono o ruido
    - Calculamos el nivel máximo de sonido enmascarado por banda
  - 4) Renormalización y umbral absoluto
    - Para deshacer el efecto de la dispersión en cada banda

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Audio MPEG-1 (II): Modelo psicoacústico (\*)

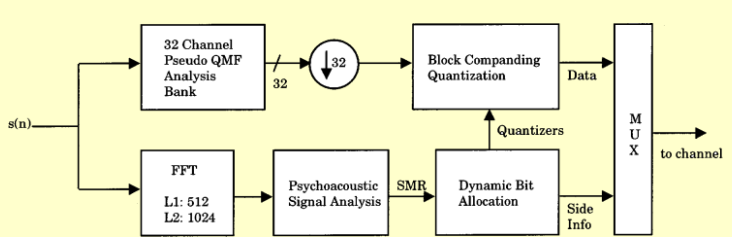
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)
Representación y Codificación AV en TVD: MPEG Audio (13)

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Audio MPEG-1 (III): Capa I (\*)

Capa I (Layer I)

- Más simple (para receptores más baratos).
- Menor compresión (384 kbps, 1:4 respecto CD)
- Codificación en 32 subbandas equiespaciadas (32 filtros QMF)
- Normaliza coeficientes (utiliza factores de escala de 6 bits)
- Codifica los coeficientes con diferentes bits (2-15 bits)
- Modelo psicoacústico I
  - FFT 512 puntos (resolución 94 Hz para frecuencia muestreo de 48 kHz)
- Utilizado por Philips en el Digital Compact Cassette (DCC)



Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)
Representación y Codificación AV en TVD: MPEG Audio (14)

## Audio MPEG-1 (IV): Capa II (\*)

### Capa II

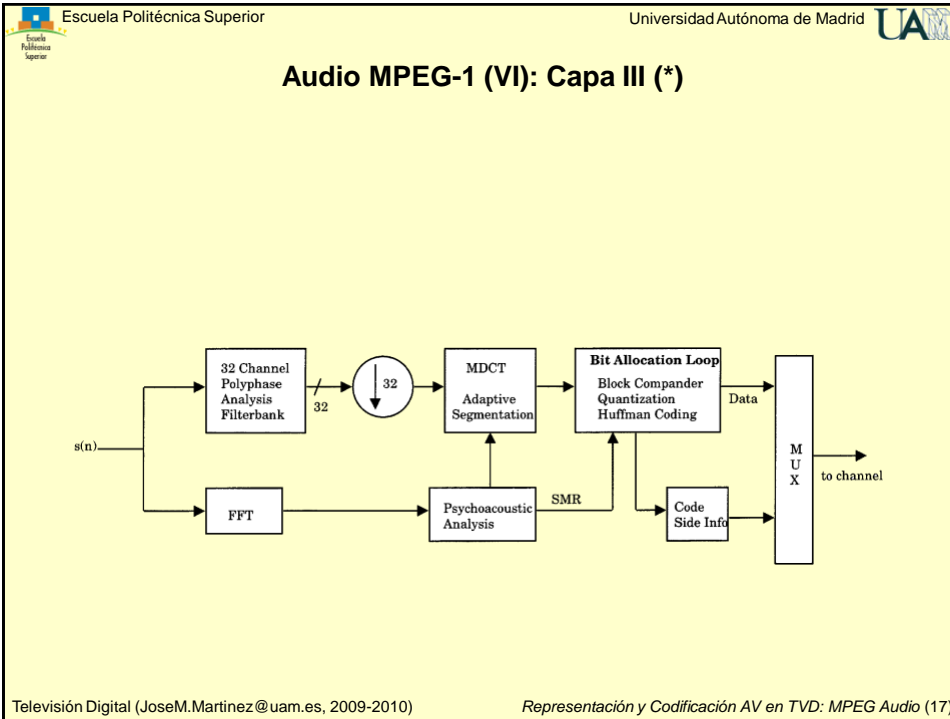
- Basado en el algoritmo MUSICAM para radiodifusión
- Mayor compresión (192-256 kbps 1:6-1:8 respecto CD)
- Algo más complejo el decodificador que el de la capa I
- Modelo psicoacústico
  - FFT 1024 puntos (resolución 47 Hz para frecuencia muestreo de 48 kHz)
- La asignación de bits y factores de escala se envía una vez cada 36 muestras de subbanda
- Cuantificación más fina
- Utilizado en radiodifusión digital (DAB) y televisión digital europea (DVB).

## Audio MPEG-1 (V): Capa III (\*)

### Capa III (conocida como MP3)

- Más complejo. Mayor compresión (112-128 kbps 1:10-1:12 respecto CD)
- Descompone cada una de las 32 subbandas en otras 18 (en total 576 coef.) mediante MDCT. Mayor resolución espectral (42 Hz)
- Control de longitud de ventanas (para evitar preecos en las transiciones)
  - 6/18 muestras – 4/12 ms (ventana corta/larga)
- Utiliza cuantificación no uniforme
- Utiliza codificación entrópica (Huffman)
- Control de distorsión mediante bucles iterativos de análisis-síntesis
- Es el estándar *de facto* para la transmisión y almacenamiento de audio comprimido (Internet, lectores MP3, ...)





Escuela Politécnica Superior Universidad Autónoma de Madrid

## Representaci3n y Codificaci3n de Se~ales Audiovisuales en Televisi3n Digital

### MPEG Audio

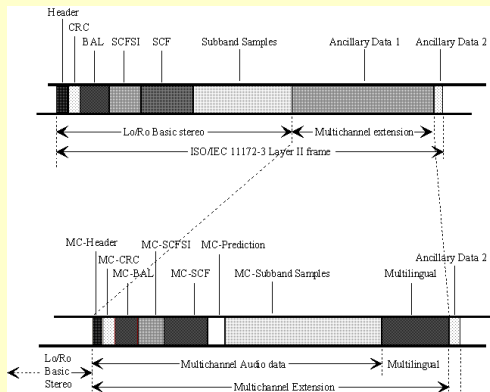
INDICE

- *Introducci3n*
- *Audio MPEG-1*
- **Audio MPEG-2**
- Audio MPEG-4
- Audio multicanal Dolby AC-3

Televi3n Digital (JoseM.Martinez@uam.es, 2009-2010) Representaci3n y Codificaci3n AV en TVD: MPEG Audio (18)

### Audio MPEG-2 (I): MPEG-2 parte 3

- Soporta el audio MPEG-1
- Lo extiende a frecuencias de muestreo más bajas
- Lo extiende hasta 5 canales
- Es lo realmente usado en DVB



Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)

Representación y Codificación AV en TVD: MPEG Audio (19)

### Audio MPEG-2 (II): Advanced Audio Coding (AAC) (\*)

Sigue el modelo de MPEG-1 Audio Layer III, pero con mejoras

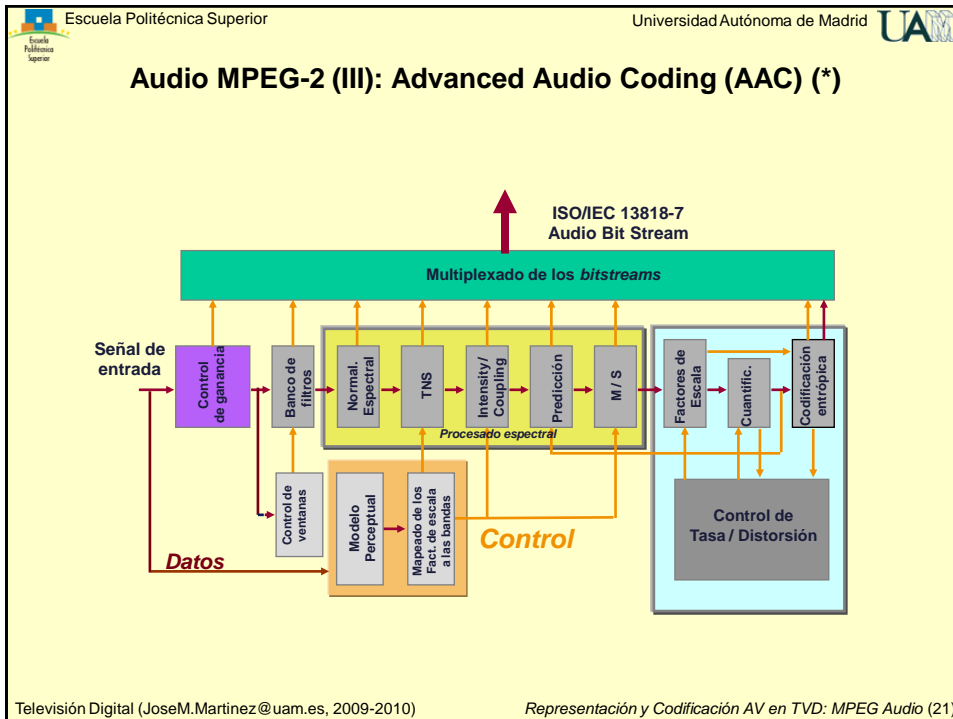
- Mejoras en eficiencia de codificación
  - Mayor resolución en frecuencia (1024 líneas de frecuencia frente a 576 en MPEG-1 Layer III)
  - Permite realizar predicción hacia atrás (opcional)
  - Estéreo conjunto (*joint stereo*) mejorado
  - Mejor y más flexible codificación Huffman
- Mejoras en la calidad de audio
  - Mejor gestión de bloques. Utiliza MDCT (en lugar del híbrido de MPEG-1 Layer III) y un filtro con respuesta al impulso más corta. Esto reduce la presencia de preecos
  - Temporal Noise Shaping (TNS). Conformar el ruido

Resultados:

- Calidad similar a MPEG-1 Layer III utilizando el 70% de la tasa binaria
  - 8-64 kbps/canal
- A igual tasa binaria mucha mejor calidad

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)

Representación y Codificación AV en TVD: MPEG Audio (20)



Escuela Politécnica Superior Universidad Autónoma de Madrid

## Audio MPEG-2 (IV): Advanced Audio Coding (AAC) (\*)

### Temporal Noise Shaping (TNS)

- Motivado por la dificultad para manejar transitorios
  - El ruido de cuantificación introducido se distribuye uniformemente en la duración de la trama.
  - Aparece el problema de los preecos en los transitorios
- Permite modelar de forma más fina la forma temporal del ruido de cuantificación
- Adapta la forma temporal del ruido de cuantificación a la forma temporal de la señal de entrada
- La forma del ruido se codifica predictivamente
  - Se utiliza un predictor lineal mediante análisis LPC

### Predictor

- Se utiliza un predictor de los coeficientes espectrales a partir de los coeficientes de la trama anterior
- Sólo codifica el residuo respecto a la predicción

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG Audio (22)


Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

**INDICE**

- *Introducción*
- *Audio MPEG-1*
- *Audio MPEG-2*
- **Audio MPEG-4**
- Audio multicanal Dolby AC-3

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) *Representación y Codificación AV en TVD: MPEG Audio (23)*

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Audio MPEG-4 (\*)

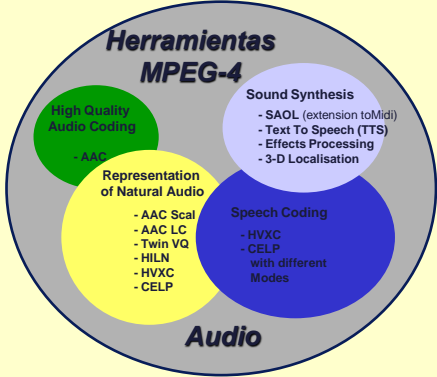
Diferentes codificaciones (herramientas) según los tipos de objetos de audio

- Audio en general (audio natural)
- Voz
- Audio sintético
- Voz sintetizada

Codificación escalable  
Perfiles (según las herramientas utilizadas)

- *Speech Profile*
- *Scalable Profile*
- *Synthesis Profile*
- *Main Profile*

En MPEG-4 se pueden codificar diferentes objetos de audio con diferentes códecs



**Herramientas MPEG-4**

**Audio**

- High Quality Audio Coding (AAC)
- Representation of Natural Audio
  - AAC Scal
  - AAC LC
  - Twin VQ
  - HILN
  - HVXC
  - CELP
- Speech Coding
  - HVXC
  - CELP with different Modes

**Sound Synthesis**

- SAOL (extension toMidi)
- Text To Speech (TTS)
- Effects Processing
- 3-D Localisation

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) *Representación y Codificación AV en TVD: MPEG Audio (24)*

## Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

### INDICE

- *Introducción*
- *Audio MPEG-1*
- *Audio MPEG-2*
- *Audio MPEG-4*
- **Audio multicanal Dolby AC-3**

## Audio multicanal Dolby AC-3 (I) (\*)

Permite codificar 5 canales (48 kHz) + canal de efectos (120 Hz) en unos 320 kbps: 5.1 canales

Explota las redundancias entre canales (canal de acoplo)

- Se codifica un canal promedio y el resto de canales respecto a éste (con menos resolución)

Modelo psicoacústico

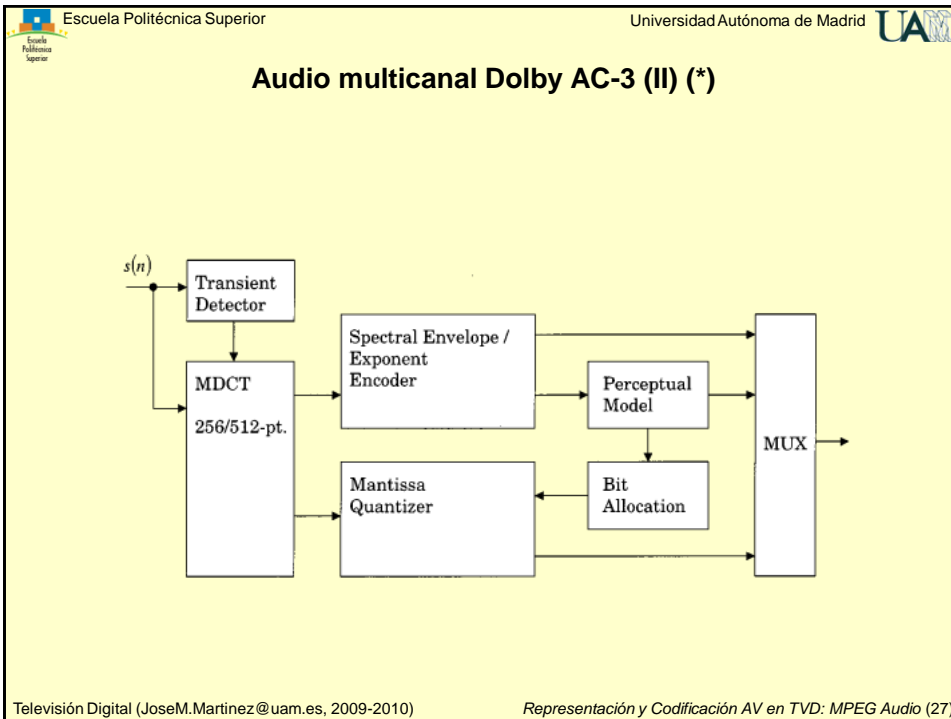
- Banco de filtros MDCT de 512 puntos
- Híbrido bidireccional (*backward/forward*) adaptativo

Codifica exponente + mantisa de cada subbanda

- El exponente indica la envolvente espectral

Soporta un modo estéreo que codifica la suma y diferencia de canales

Utilizado en DVD, televisión digital en USA (ATSC), cines, ...



Escuela Politécnica Superior Universidad Autónoma de Madrid

## Representaci3n y Codificaci3n de Se~ales Audiovisuales en Televisi3n Digital

### MPEG Audio

#### INDICE

- *Introducci3n*
- *Audio MPEG-1*
- *Audio MPEG-2*
- *Audio MPEG-4*
- *Audio multicanal Dolby AC-3*

Televi3n Digital (JoseM.Martinez@uam.es, 2009-2010) Representaci3n y Codificaci3n AV en TVD: MPEG Audio (28)



## Créditos

Estas transparencias están basadas en un trabajo de Luis Herranz.