



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

ÍNDICE

- Introduction
- The MPEG Family
- MPEG-1
- MPEG-2



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

INDICE

- Introduction
- The MPEG Family
- MPEG-1
- MPEG-2



Introduction: MPEG Standardisation Group



Moving Picture Experts Group

(created in 1988)

ISO/IEC JTC 1/SC 29/ WG 11

Int. Standards Organization/Int. Electrotechnical Commission

Development of standards for the coded representation of
moving pictures and associated audio

MPEG-1 MPEG-2 MPEG-4 MPEG-7 MPEG-21 MPEG-X



Introduction: Mission (Terms of Reference)

Informal Title

Moving Picture Experts Group (MPEG)

Formal Title

Coding of Moving Pictures and Audio

Area of work

Development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio, and their combination, in order to satisfy a wide variety of applications.



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

INDICE

- *Introduction*
- **The MPEG Family**
- MPEG-1
- MPEG-2



The MPEG Family

- MPEG-1 - Error free environments
- MPEG-2 - Broadcast TV (SD and HD)
- MPEG-4 - Object based coding
- MPEG-7 - Multimedia description
- MPEG-21 - Seven element initiative for multimedia deployment

- *MPEG-A – Combining parts for applications design*
- *MPEG-B – Binary Format for XML, Fragment Request Unit*
- *MPEG-C – Fixed point DCT/IDCT, Auxiliary Video Data Representation*
- *MPEG-D – MPEG-Surround*
- *MPEG-E – Multimedia Middleware*



The MPEG Family: MPEG-1

MPEG-1: ISO/IEC 11172 (1992)

- Coding of moving pictures and associated audio for digital storage media at around 1.5 Mbit/s
- For digital storage media (CD-ROM)



The MPEG Family: MPEG-2

MPEG-2: ISO/IEC 13818 (1994)

- Generic coding of moving pictures and associated audio information
- Transmission and storage of coding moving pictures and associated audio with contribution and distribution (broadcast) quality for Digital Television (3 ~ 15 Mbit/s)
- More complex and flexible than MPEG-1
 - Different formats of video input (4:2:0, 4:2:2 – progressive, interlaced).
 - More flexibility for determining movement vectors at MBs level.
 - Scalability
- Some minor extensions underway



The MPEG Family: MPEG-4

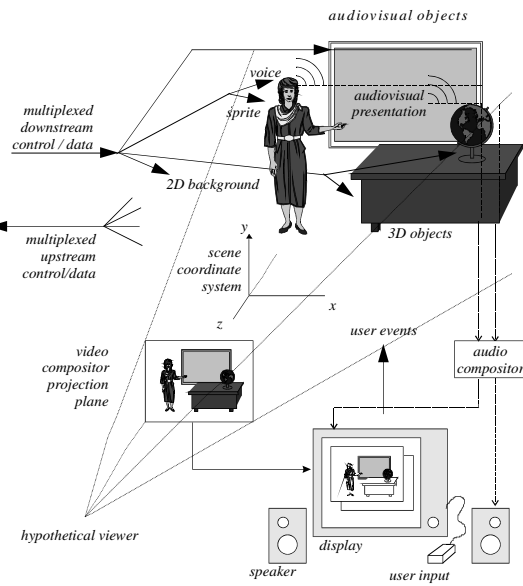
MPEG-4: ISO/IEC 14496 (1998)

- coding of natural and synthetic audiovisual objects
- Multimedia applications
- further extensions (mainly) in video and systems parts underway
 - H.264/MPEG-4 part 10: the “ultimate video standard” (not object based)

The MPEG Family: MPEG-4

MPEG-4: ISO/IEC 14496

- coding of natural and synthetic audio
- Multimedia applications
- further extensions (mainly H.264/MPEG-4 part 10)



The MPEG Family: MPEG-7

MPEG-7: ISO/IEC 15938 (2001)

- multimedia content description for AV material
- content search, filter and access (navigation, browsing, transcoding, ...)

Escuela Politécnica Superior Universidad Autónoma de Madrid **UA**

The MPEG Family: Scope of MPEG-7

Description generation → Description → Description consumption

Research and future competition → Scope of MPEG-7

The description generation (feature extraction, indexing process, annotation & authoring tools,...) and consumption (search engine, filtering tool, retrieval process, browsing device, ...) are non normative parts of MPEG-7.

The goal is to define the minimum that enables interoperability.

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

Escuela Politécnica Superior Universidad Autónoma de Madrid **UA**

The MPEG Family: Abstract MPEG-7 Architecture and Working Areas

MM Content → **Feature extraction** → **Description Generation** → **MPEG7 Description (XML)** → **Encoder for Delivery** → **Binary and Textual MPEG-7 Coded Description** → **Decoder for Delivery** → **MPEG-7 Terminal**

Users and computational systems ↔ **Filter** ↔ **Search & Browse** ↔ **DDL**

Description Schemes (DS): DS1, DS2, DS3, DS4, DS5
Description (D): D1, D2, D3, D4, D5

Description Definition Language (DDL): MPEG-7 DDL

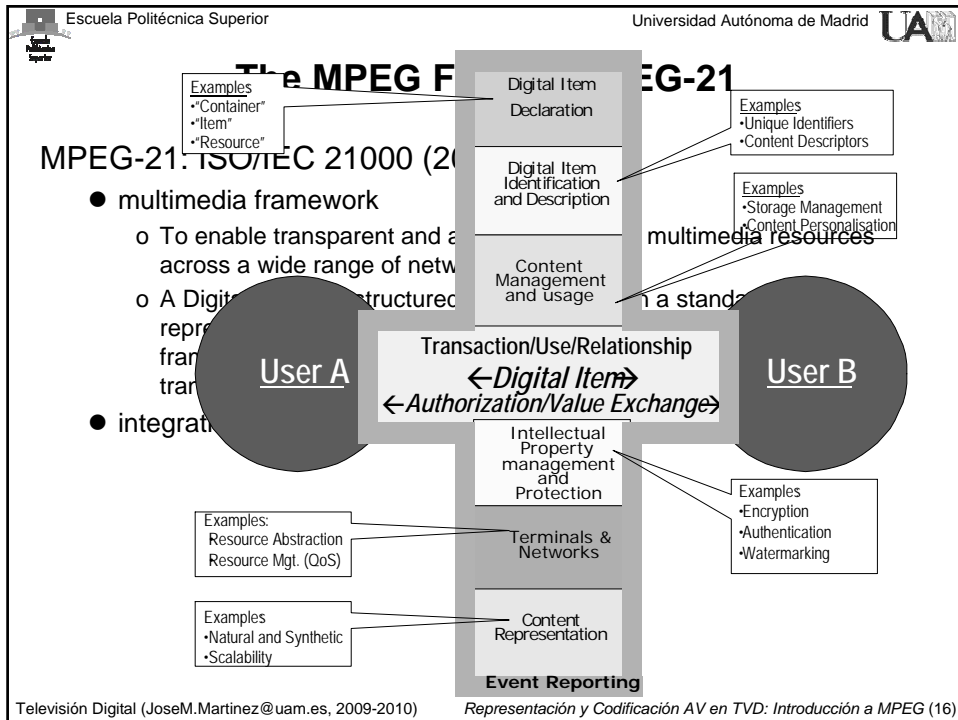
MPEG-7 Systems: BiM/TeM AU, BiM/TeM AU, BiM/TeM AU

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

The MPEG Family: MPEG-21

MPEG-21: ISO/IEC 21000 (2001)

- multimedia framework
 - To enable transparent and augmented use of multimedia resources across a wide range of networks and devices
 - A Digital Item is a structured digital object with a standard representation, identification and metadata within the MPEG-21 framework. This entity is also the fundamental unit of distribution and transaction within this framework
- integration of e-content technologies





Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

INDICE

- *Introduction*
- *The MPEG Family*
- **MPEG-1**
- MPEG-2



MPEG-1 Parts

ISO/IEC 11172: MPEG-1 (1992)

Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s

Part 1 - MPEG-1 Systems - Program Stream (1993)

Part 2 - MPEG-1 Video for CD-I (1993)

Part 3 - MPEG-1 audio (1993)

including Layers I, II, and III (a.k.a. mp3)

Part 4 – Compliance Testing (1995)

Part 5 – Software Simulation (1998)

Part 6 – Specification for implementation of IDCT (CD)

Directed primarily at error free environments

MPEG-1 Systems (*)

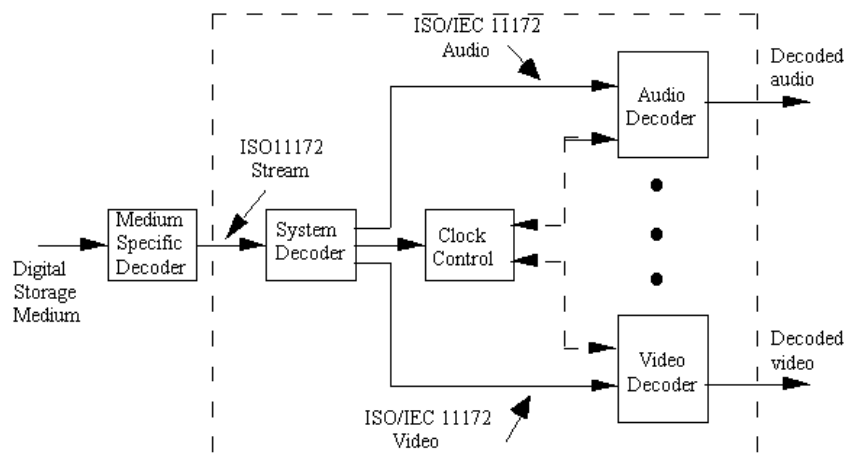
ISO/IEC 11172-1:1993 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 1: Systems

- ISO/IEC 11172-1:1993/Cor 1:1996
- ISO/IEC 11172-1:1993/Cor 2:1999

Addresses the problem of combining one or more data streams from the video and audio parts of the MPEG-1 standard with timing information to form a single stream. This is an important function because, once combined into a single stream, the data are in a form well suited to digital storage or transmission.

MPEG-1 Systems (*)

Prototypical ISO/IEC 11172 decoder



MPEG-1 Visual (*)

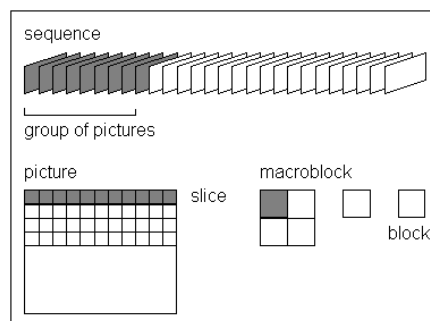
ISO/IEC 11172-2:1993 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 2: Video

- ISO/IEC 11172-2:1993/Cor 1:1996
- ISO/IEC 11172-2:1993/Cor 2:1999
- ISO/IEC 11172-2:1993/Cor 3:2003

Specifies a coded representation that can be used for compressing video sequences - both 625-line and 525-lines - to bitrates around 1,5 Mbit/s. Part 2 was developed to operate principally from storage media offering a continuous transfer rate of about 1,5 Mbit/s. Nevertheless it can be used more widely than this because the approach taken is generic.

MPEG-1 Visual (*)

A number of techniques are used to achieve a high compression ratio. The first is to select an appropriate spatial resolution for the signal. The algorithm then uses block-based motion compensation to reduce the temporal redundancy. Motion compensation is used for causal prediction of the current picture from a previous picture, for non-causal prediction of the current picture from a future picture, or for interpolative prediction from past and future pictures. The difference signal, the prediction error, is further compressed using the discrete cosine transform (DCT) to remove spatial correlation and is then quantised. Finally, the motion vectors are combined with the DCT information, and coded using variable length codes.



MPEG-1 Audio (*)

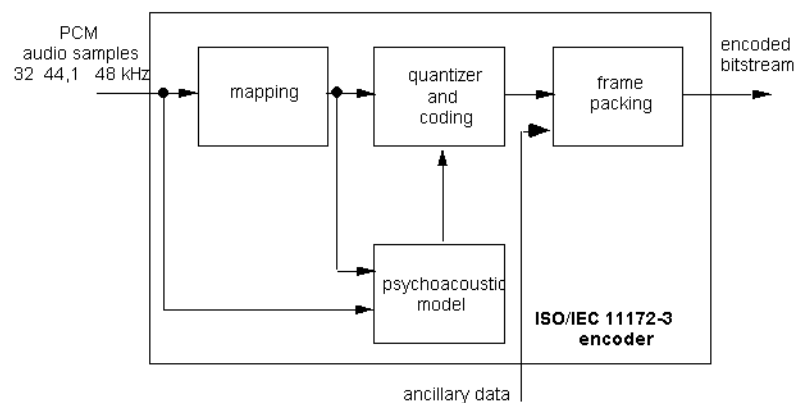
ISO/IEC 11172-3:1993 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio

- ISO/IEC 11172-3:1993/Cor 1:1996

Specifies a coded representation that can be used for compressing audio sequences - both mono and stereo.

- Input audio samples are fed into the encoder.
- The mapping creates a filtered and subsampled representation of the input audio stream.
- A psychoacoustic model creates a set of data to control the quantiser and coding.
- The quantiser and coding block creates a set of coding symbols from the mapped input samples.
- The block 'frame packing' assembles the actual bitstream from the output data of the other blocks, and adds other information (e.g. error correction) if necessary.

MPEG-1 Audio (*)





MPEG-1 Compliance (*)

ISO/IEC 11172-4:1995 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 4: Compliance testing

Specifies how tests can be designed to verify whether bitstreams and decoders meet the requirements as specified in parts 1, 2 and 3 of the MPEG-1 standard. These tests can be used by:

- manufacturers of encoders, and their customers, to verify whether the encoder produces valid bitstreams.
- manufacturers of decoders and their customers to verify whether the decoder meets the requirements specified in parts 1,2 and 3 of the standard for the claimed decoder capabilities.
- applications to verify whether the characteristics of a given bitstream meet the application requirements, for example whether the size of the coded picture does not exceed the maximum value allowed for the application.



MPEG-1 Software simulation (*)

ISO/IEC TR 11172-5:1998 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 5: Software simulation

A technical report giving a full software implementation of the first three parts of the MPEG-1 standard



MPEG-1 implementation of IDCT (*)

ISO/IEC TR 11172-6:200X Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- *Part 6: Specification for implementation of IDCT (CD)*

Will provide the specification of IDCT accuracy. The current standard refers to the IEEE 1180 standard, which has been withdrawn. This part will provide equivalent text or extension thereof.



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

INDICE

- *Introduction*
- *The MPEG Family*
- *MPEG-1*
- **MPEG-2**

MPEG-2 Parts

ISO/IEC 13818: MPEG-2 (1994)

Generic coding of moving pictures and associated audio information



1996 Emmy for technical excellence

- Part 1 Systems (joint with ITU) - 2nd edition 2000
- Part 2 Video (joint with ITU) - 2nd edition 2000
- Part 3 Audio-fwd/bwd compatible to MPEG-1 (1998)
- Part 4 Conformance testing – 2nd edition 2004
- Part 5 Software simulation/Reference software (1998)
- Part 6 DSM-CC: Digital Storage Media Cmd and Cntl (1998)
- Part 7 AAC - Advanced Audio Coding – 3rd ed. 2004
- Part 9 RTI - Real Time Interface (1996)
- Part 10 Conformance Extensions for DSM-CC (1999)
- Part 11 IPMP on MPEG-2 Systems (2004)

MPEG-2 Systems

ISO/IEC 13818-1:2000 Information technology --
Generic coding of moving pictures and associated
audio information: Systems

Addresses the combining of one or more elementary streams of video and audio, as well as, other data into single or multiple streams which are suitable for storage or transmission. This is specified in two forms: the Program Stream and the Transport Stream. Each is optimised for a different set of applications.

MPEG-2 Systems

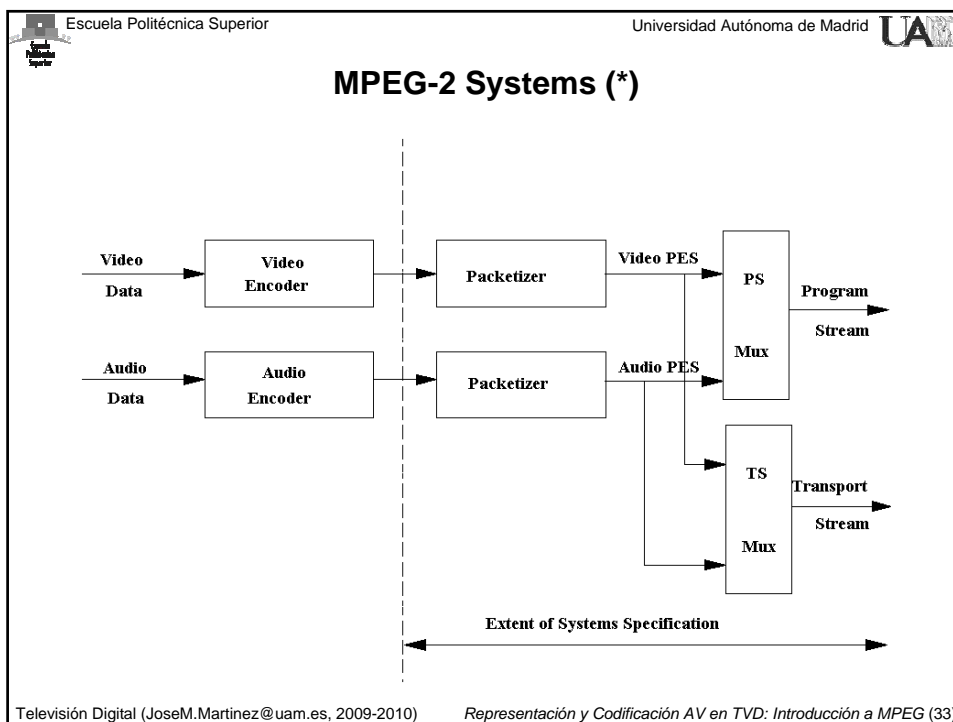
The Program Stream is similar to MPEG-1 Systems Multiplex. It results from combining one or more Packetised Elementary Streams (PES), which have a common time base, into a single stream. The Program Stream is designed for use in relatively error-free environments and is suitable for applications which may involve software processing. Program stream packets may be of variable and relatively great length.

The Transport Stream combines one or more Packetized Elementary Streams (PES) with one or more independent time bases into a single stream. Elementary streams sharing a common timebase form a program. The Transport Stream is designed for use in environments where errors are likely, such as storage or transmission in lossy or noisy media. Transport stream packets are 188 bytes long.

MPEG-2 Systems (*)

ISO/IEC 13818-1:2000 Information technology --
Generic coding of moving pictures and associated
audio information: Systems

- ISO/IEC 13818-1:2000/Cor 1:2002
- ISO/IEC 13818-1:2000/Cor 2:2002
- ISO/IEC 13818-1:2000/Cor 3:2004
- ISO/IEC 13818-1:2000/Amd 1:2003: Carriage of metadata over 13818-1 streams
- ISO/IEC 13818-1:2000/Amd 2:2004: Support of IPMP on MPEG-2 Systems
- ISO/IEC 13818-1:2000/Amd 3:2004: Transport of AVC over MPEG-streams
- ISO/IEC 13818-1:2000/Amd 4:200x: Metadata application format codepoints (FDAM)
- ISO/IEC 13818-1:2000/Amd 5:200x: New audio profile and level signaling (FDAM)



Escuela Politécnica Superior Universidad Autónoma de Madrid

MPEG-2 Video

ISO/IEC 13818-2:2000 Information technology -- Generic coding of moving pictures and associated audio information: Video

Builds on the powerful video compression capabilities of the MPEG-1 standard to offer a wide range of coding tools. These have been grouped in profiles (with levels) to offer different functionalities.

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción a MPEG (34)



MPEG-2 Video (*)

ISO/IEC 13818-2:2000 Information technology -- Generic coding of moving pictures and associated audio information: Video

- ISO/IEC 13818-2:2000/Cor 1:2002
- ISO/IEC 13818-2:2000/Amd 1:2001: Content description data

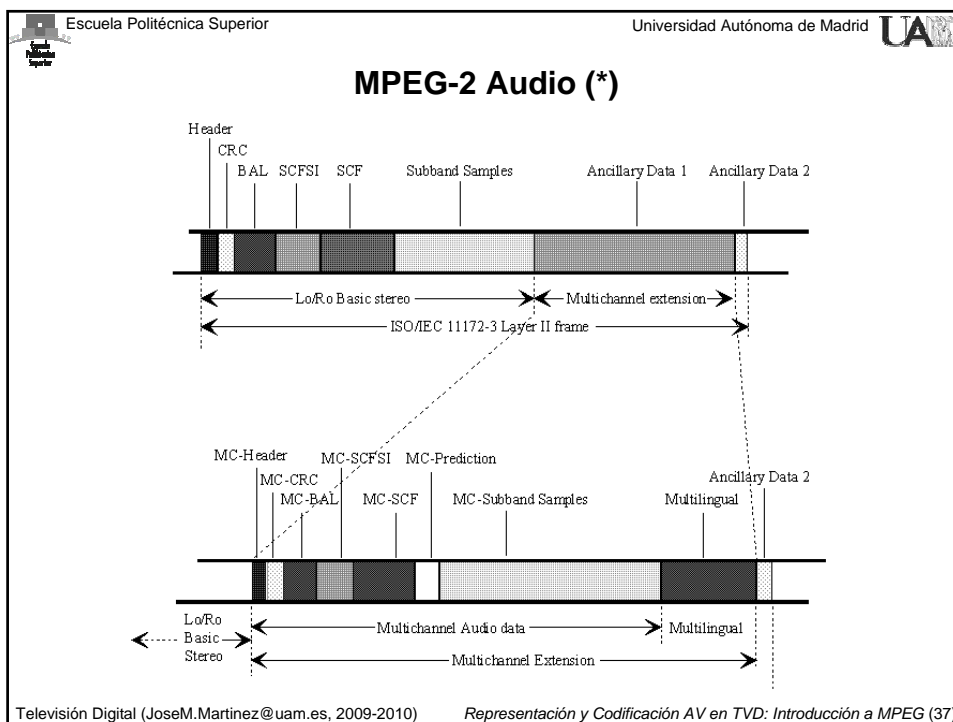


MPEG-2 Audio

ISO/IEC 13818-3:1998 Information technology -- Generic coding of moving pictures and associated audio information -- Part 3: Audio

Backwards-compatible multichannel extension of the MPEG-1 Audio standard.

- This is the MPEG-1 specification really used in DVB



Escuela Politécnica Superior Universidad Autónoma de Madrid

MPEG-2 Conformance and Software simulation (*)

ISO/IEC 13818-4:2004 Information technology -- Generic coding of moving pictures and associated audio information -- Part 4: Conformance testing

- ISO/IEC 13818-4:2004/Amd 1:200x IPMP conformance testing (FDAM)
- ISO/IEC 13818-4:2004/Amd 2:200x Audio conformance extensions (FDAM)

ISO/IEC TR 13818-5:1997 Information technology -- Generic coding of moving pictures and associated audio information -- Part 5: Software simulation

- ISO/IEC TR 13818-5:1997/Amd 1:1999 Advanced Audio Coding (AAC)
- ISO/IEC TR 13818-5:1997/Amd 2:200x IPMP Reference software

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción a MPEG (38)

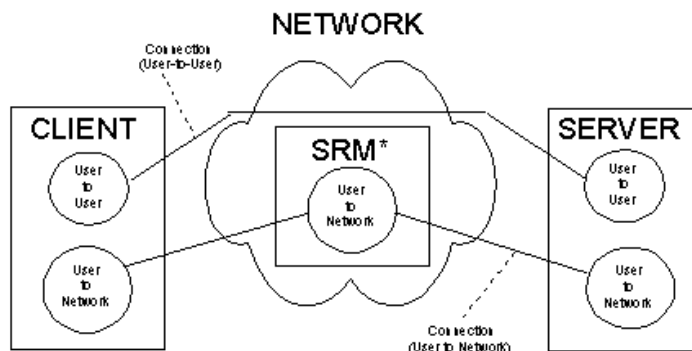
MPEG-2 DSM-CC (*)

ISO/IEC 13818-6:1998 Information technology -- Generic coding of moving pictures and associated audio information -- Part 6: Extensions for DSM-CC


- ISO/IEC 13818-6:1998/Cor 1:1999
- ISO/IEC 13818-6:1998/Amd 1:2000 Additions to support data broadcasting
- ISO/IEC 13818-6:1998/Amd 2:2000 Additions to support synchronized download services, opportunistic data services and resource announcement in broadcast and interactive services
- ISO/IEC 13818-6:1998/Amd 3:2001 Transport buffer model in support of synchronized user-to-network download protocol

Digital Storage Media Command and Control (DSM-CC) is the specification of a set of protocols which provides the control functions and operations specific to managing MPEG-1 and MPEG-2 bitstreams. These protocols may be used to support applications in both stand-alone and heterogeneous network environments. In the DSM-CC model, a stream is sourced by a Server and delivered to a Client. Both the Server and the Client are considered to be Users of the DSM-CC network. DSM-CC defines a logical entity called the Session and Resource Manager (SRM) which provides a (logically) centralized management of the DSM-CC Sessions and Resources

MPEG-2 DSM-CC (*)



*May provide session, connection, and configuration management and control


Escuela Politécnica Superior Universidad Autónoma de Madrid 

MPEG-2 AAC

ISO/IEC 13818-7:2004 Information technology -- Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC)

Specification of a multichannel audio coding algorithm not constrained to be backwards-compatible with MPEG-1 Audio.

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) *Representación y Codificación AV en TVD: Introducción a MPEG (41)*

Escuela Politécnica Superior Universidad Autónoma de Madrid 

MPEG-2 part 8 (the missing link) (*)

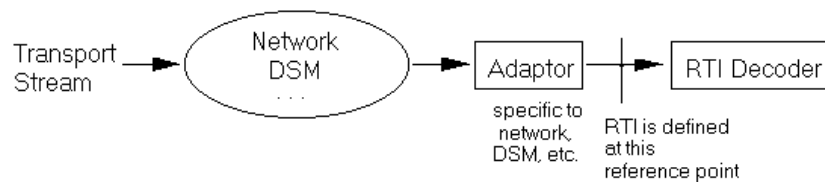
MPEG-2 part 8 was originally planned to be coding of video when input samples are 10 bits. Work on this part was discontinued when it became apparent that there was insufficient interest from industry for such a standard.

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) *Representación y Codificación AV en TVD: Introducción a MPEG (42)*

MPEG-2 RTI (*)

ISO/IEC 13818-9:1996 Information technology -- Generic coding of moving pictures and associated audio information -- Part 9: Extension for real time interface for systems decoders

Specification of the Real-time Interface (RTI) to Transport Stream decoders which may be utilised for adaptation to all appropriate networks carrying Transport Streams



MPEG-2 Conformance extensions (*)

ISO/IEC 13818-10:1999 Information technology -- Generic coding of moving pictures and associated audio information -- Part 10: Conformance Extensions for DSM-CC

Conformance testing part of DSM-CC



MPEG-2 IPMP (*)

ISO/IEC 13818-11:2004 Information technology -- Generic coding of moving pictures and associated audio information -- Part 11: IPMP on MPEG-2 Systems

IPMP specifications for MPEG-2



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción a los estándares MPEG

INDICE

- *Introduction*
- *The MPEG Family*
- *MPEG-1*
- *MPEG-2*



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

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



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

INDICE

- Introducción
- MPEG-2 Flujo de programa
- MPEG-2 Flujo de transporte
- Organización del múltiplex



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

INDICE

- **Introducción**
- MPEG-2 Flujo de programa
- MPEG-2 Flujo de transporte
- Organización del múltiplex

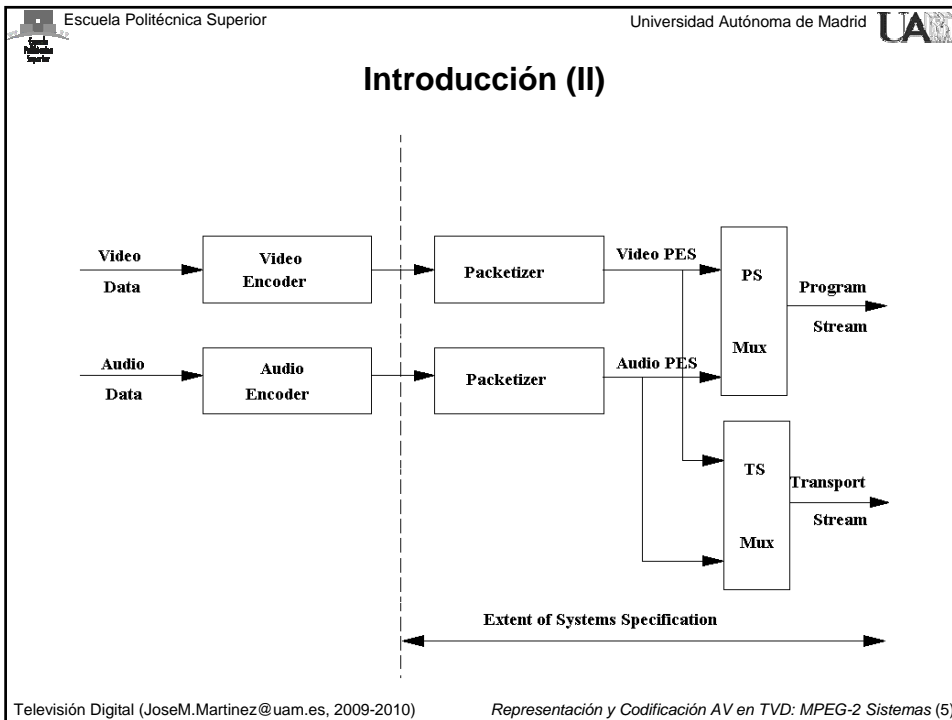


Introducción (I)

ISO/IEC 13818-1 MPEG-2, se centra en la sincronización y multiplexado de uno o más flujos elementales de vídeo, audio y otros datos en uno o varios flujos de datos para su almacenamiento o transmisión por canales digitales.

DVB asume la organización del Mux según el flujo de transporte (*Transport Stream*, TS) definido por MPEG-2.

- Salidas de los codificadores de vídeo o audio son los Flujos Elementales (*Elementary Streams*, ES).
- Los ES se disponen en paquetes (*Packetized Elementary Streams*, PES) para aumentar la eficiencia de protección contra errores y la adaptación al medio de transmisión.
- El TS combina en un único flujo binario, información que proviene de diversas fuentes (con igual o diferente base de tiempos) y que constituyen uno o varios programas.



Escuela Politécnica Superior Universidad Autónoma de Madrid

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

INDICE

- *Introducción*
- **MPEG-2 Flujo de programa**
- MPEG-2 Flujo de transporte
- Organización del múltiplex

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG-2 Sistemas (6)



MPEG-2 flujo de programa (*)

Similar a MPEG-1 Sistemas

Combina en un único flujo uno o más flujos de paquetes (PES) compartiendo una misma base de tiempos (servicio)

Es posible que cada PES genere un PS propio, y que exista una compartición de base de tiempo

El PS está diseñado para entornos con pocos errores

El tamaño de los paquetes puede ser variable y de dimensión relativamente grande



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

INDICE

- *Introducción*
- *MPEG-2 Flujo de programa*
- **MPEG-2 Flujo de transporte**
- Organización del múltiplex

MPEG-2 flujo de transporte (I)

El flujo de transporte combina en un único flujo binario, información que proviene de diversas fuentes que pueden, o no, compartir una misma base de tiempos y que constituyen uno o varios programas (un TS puede llevar un único programa). Los PES de un programa si que tienen que compartir base de tiempos.

Entornos con errores

Longitud de los paquetes es fija = 188 bytes

- permite aplicar fuertes mecanismos de codificación de canal que actúan más eficientemente en bloques pequeños
- Compatible con ATM
- 4 bytes de cabecera y 184 de carga útil

MPEG-2 flujo de transporte (II)

TS con programa único

- Diversos flujos elementales
 - Audio, video, datos, ...
 - Podrían ser PS, aunque no se suele utilizar (y nunca en DVB)
- Tabla PMT (Program Map Table)
 - Contiene los identificadores PID de los PES que constituyen el programa
 - Permite al receptor localizar y extraer del multiplex los distintos PES a entregar a cada decodificador

TS multiprograma

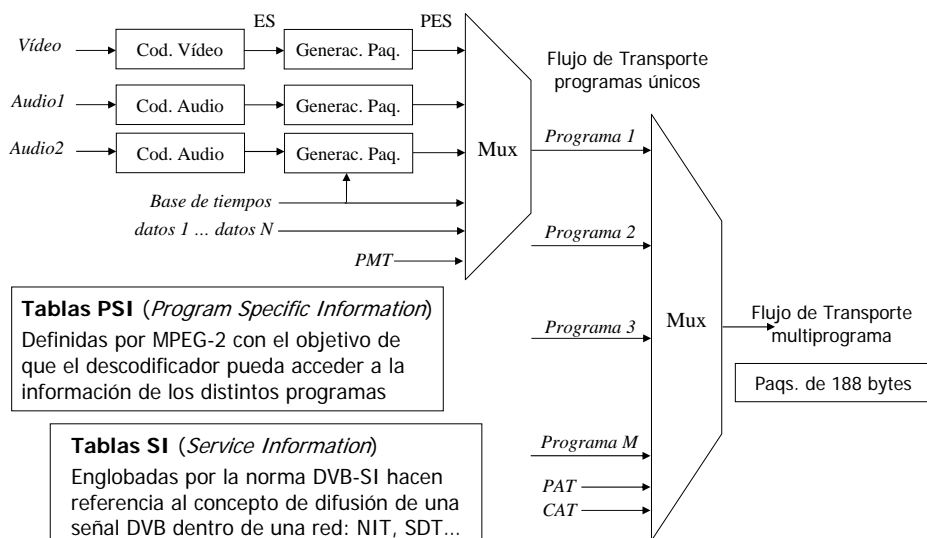
- Diversos TS de programa único
 - Pueden incluir "programas de TXT", "programas de interactivos", ... comunes al multiplex como si fuese un canal/programa más
- Tabla PAT (Program Association Table)
 - Contiene la información necesaria para poder extraer del flujo multiprograma los paquetes que contienen datos de la PMT asociada a cada programa específico
- Tabla CAT (Conditional Access Table)
 - Contiene el PID de los paquetes que contienen información relativa al acceso condicional

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

INDICE

- *Introducción*
- *MPEG-2 Flujo de programa*
- *MPEG-2 Flujo de transporte*
- **Organización del múltiplex**

Organización del Múltiplex





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

INDICE

- *Introducción*
- *MPEG-2 Flujo de programa*
- *MPEG-2 Flujo de transporte*
- *Organización del múltiplex*



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

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



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

INDICE

- Introducción
- Estructuras de datos
- Modos de predicción
- Modos de codificación de cuadro
- Escalabilidad
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía

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

INDICE

- **Introducción**
- Estructuras de datos
- Modos de predicción
- Modos de codificación de cuadro
- Escalabilidad
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía

Introducción (I)

Su campo principal de aplicación es la TVD (contribución y distribución)

- Pero su carácter genérico lo hace útil en diversos entornos

Perfiles y niveles

Se sigue un esquema híbrido predictivo

- Separación en MB/Bs, Predicción temporal (estimación movimiento a nivel MB), DCT de la señal residual (a nivel B), recorrido en zig-zag (diferente según Y o C), cuantificación mediante matriz de visibilidad, VLC, control de buffer
- Como MPEG-1 (IPB) pero con señal tanto progresiva como entrelazada (la de TV)
 - o MPEG-1 solamente SIF (288x352) y 4:2:0 (otros formatos mediante conversión)
- Formatos de entrada 4:2:0, 4:2:2, 4:4:4 (no todos tienen perfil actualmente)
- Dimensiones iniciales BT.601: Extensiones desde SIF hasta HD

Introducción (II): Formato 4:2:0 en MPEG-2

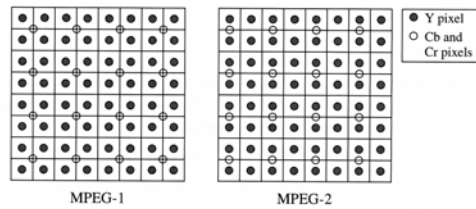


Figure 13.14 Luminance and chrominance samples in a 4:2:0 progressive frame.

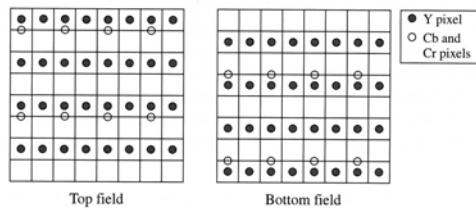


Figure 13.15 Luminance and chrominance samples in a 4:2:0 interlaced frame, in which the top field is temporally first.

Imagen extraída de "Video processing and communications"; Wang, Ostermann, Zhang, Prentice Hall, 2002

Representación y Codificación de Señales Audiovisuales en Televisión Digital

MPEG-2 Vídeo

INDICE

- *Introducción*
- **Estructuras de datos**
- Modos de predicción
- Modos de codificación de cuadro
- Escalabilidad
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía



Estructuras de datos (I)

Estructuras de datos organizadas jerárquicamente y directamente relacionadas con la estructuración del flujo binario resultante de la codificación de la secuencia

- Secuencia
- GOP
- Cuadro
- Tira
- Macrobloque
- Bloque



Estructuras de datos (II)

Secuencia: conjunto de cuadros consecutivos en el tiempo

- Sintaxis: Código de cabecera e Información general: dimensión horizontal y vertical de cuadros, Relación de aspecto, Número de cuadros por segundo, Velocidad de canal fija o variable, tamaño del buffer, tablas empleadas para la cuantificación

Grupo de Cuadros (Group of Pictures –GOP–): secuencia de cuadros correspondiente a un periodo de la estructura de modos de procesamiento. Su función principal es permitir el acceso aleatorio a partes de la secuencia

- Sintaxis: Código de cabecera y código de tiempos correspondiente al primer cuadro del GOP

Cuadro: una componente de luminancia y dos de crominancia. Unidad primaria de codificación

- Sintaxis: Código de cabecera e Información particular: referencia temporal del cuadro, modo de codificación, dimensiones máximas de los vectores de movimiento, etc
- Si el modo es entrelazado un cuadro tiene dos campos

Estructuras de datos (III)

Tira: conjunto de macrobloques consecutivos (de izquierda a derecha y de arriba abajo) en la imagen.

- La definición de la longitud y número de tiras en la imagen es libre, pudiendo llegarse a una tira por imagen (todos sus MBs)
- La tira es la unidad de resincronización en el caso de errores
- Sintaxis: Código de cabecera e información particular: posición del primer MB de la tira en la imagen, parámetros globales de cuantificación de los MBs de la tira

Macrobloque: mínimo conjunto de bloques enteros tanto Y como Cr, Cb que ocupan la misma posición espacial

- Unidad de estimación y compensación de movimiento
- Sintaxis: información sobre el tipo de predicción empleado, el vector de movimiento (en su caso) o el parámetro que controla el nivel de cuantificación de los coeficientes transformados.
- 16x16 en Y (en CrCb depende de la estructura de muestreo)

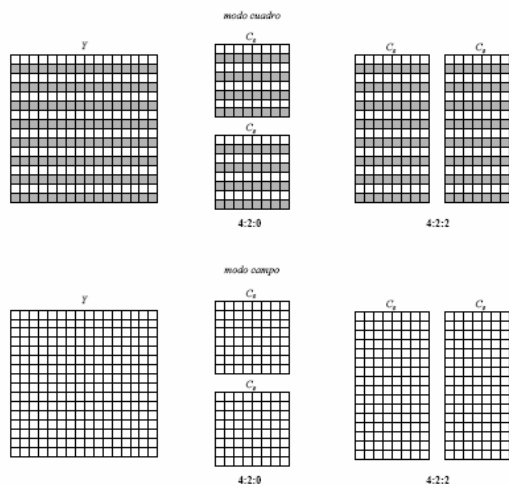
Bloque: grupo de 8x8 muestras (Y, Cr o Cb) que constituyen la unidad mínima de tratamiento

- Unidad de codificación DCT
- Sintaxis: información codificada de los coeficiente cuantificados

Estructuras de datos (IV)

Los MBs de luminancia tiene siempre 16x16, mientras que los de crominancias dependen de la estructura de muestreo

- 4:2:0: cada croma 8x8
- 4:2:2 cada croma 8x16
- 4:4:4 cada croma 16x16



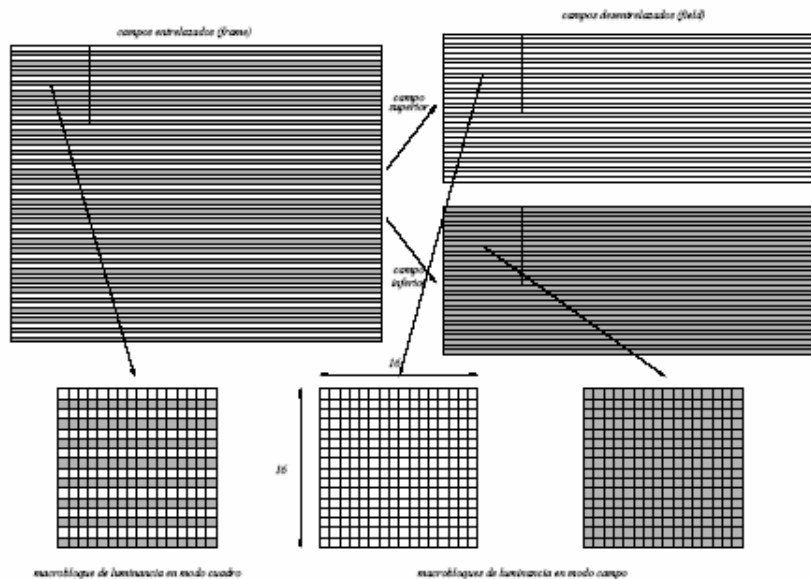
Estructuras de datos (V)

Si la señal es progresiva (en el macrobloque) las muestras de luminancia y crominancia pertenecerán todas al mismo cuadro

Si la señal es entrelazada se decide a priori si en la codificación de los cuadros de la secuencia se van a tratar los campos que los constituyen ...

- ... por separado (desentrelazándolos)
 - o modo campo (field mode)
 - o Habrá un MB para cada uno de los dos campos
 - o Solamente es posible si la secuencia de entrada es entrelazada
- ... juntos (manteniendo el entrelazado)
 - o modo cuadro (frame mode)
 - o El MB tiene muestras que pertenecen a ambos campos
 - o Es posible para secuencias de entrada entrelazadas y progresivas

Estructuras de datos (VI)





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

INDICE

- *Introducción*
- *Estructuras de datos*
- **Modos de predicción**
- Modos de codificación de cuadro
- Escalabilidad
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía



Modos de predicción (I)

Existen 4 modos de predicción (1 más que en MPEG-1)

Intracadro: no se hace uso de predicción. El MB se codifica directamente mediante la DCT

Predictivo: un MB se codifica tras realizar una predicción obtenida a partir de otro cuadro o campo de la secuencia.

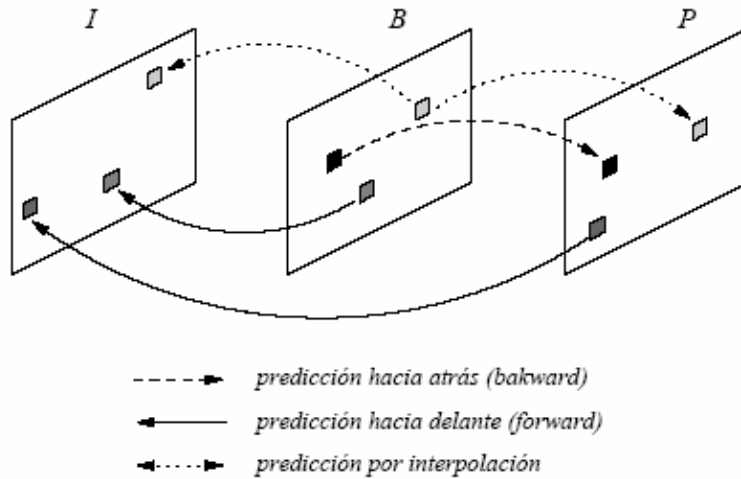
- Si es anterior: predicción hacia delante (*forward*)
 - La referencia (campo/cuadro) ya pasada en el tiempo
- Si es posterior: predicción hacia atrás (*backward*)
 - La referencia (campo/cuadro) "todavía no ha llegado"

Interpolado: el MB se codifica en modo predictivo pero haciendo uso de predicciones de un cuadro o campo anterior y otro posterior de la secuencia

Dual Prime: es una variación del interpolado, donde los dos MBs a interpolar se corresponden a predicciones *forward*.

- Para vídeo entrelazado genera cuadros P con la eficiencia de cuadros B, pero sin su retardo

Modos de predicción (II)



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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- **Modos de codificación de cuadro**
- Escalabilidad
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía



Modos de codificación de cuadro (I)

En MPEG-2 existen tres modos de codificación a nivel de cuadro

- Modo Intracuadro: cuadros I
 - No se hacen uso de predicciones.
 - Todos los MBs van con modo de predicción intracuadro
 - En MPEG-1 existe un caso particular de este modo en el cual solamente se transmite el coeficiente DC codificado (cuadros tipo D): permiten visionado rápido a baja calidad
- Modo Predictivo: cuadros P
 - Se codifican haciendo uso de predicciones de un cuadro anterior de tipo I o P.
 - Los MBs pueden predecirse en modo intracuadro (MB intra), modo predictivo con el último I o P anterior (MB FW)
- Modo Interpolado Bidireccional: cuadros B
 - Se codifican haciendo uso de predicciones tomadas del cuadro anterior I o P y/o del cuadro posterior I o P.
 - Los MBs pueden predecirse en modo intracuadro (MB intra), modo predictivo con el último I o P anterior (MB FW), modo predictivo con el siguiente I o P posterior (MB BW), modo interpolado (calculando el promedio de una predicción del cuadro anterior I o P y del cuadro posterior I o P – MB interpolado-)



Modos de codificación de cuadro (II)

Grupo de cuadros (Group of Pictures – GOP-)

- conjunto de cuadros correspondientes a un periodo de la estructura de modos de procesamiento.
- a cada uno de los cuadros del GOP se le asigna un modo de procesamiento (cuadro tipo I, P o B) en función de su posición en el GOP.

Una definición más típica de GOP es la de un conjunto de cuadros consecutivos que comienza en los cuadros B anteriores a un cuadro I y que termina con el cuadro P anterior al siguiente cuadro I.

- No puede tener más que un cuadro I
- La longitud mínima es 1 (I) y la máxima no está limitada en el estándar
- Empieza por I o B (no puede haber Ps antes del I, si long 1 => I)
- Terminar por P (o I si es de longitud 1)

Modos de codificación de cuadro (III)

GOPs

- I
- I P
- I B P B P
- B B I B P B P

GOP cerrado

- Se puede codificar (y por tanto decodificar) de manera independiente de los GOPs anteriores y posteriores de la secuencia
 - o Empiezan por cuadro I
 - o Empiezan por cuadros B que solamente tienen MB intra o MB BW (no pueden tener MB FW ni MB interpolados): esto es, solamente predicciones respecto al I del propio GOP
 - o Permiten edición "libre" a nivel GOP.

Modos de codificación de cuadro (IV)

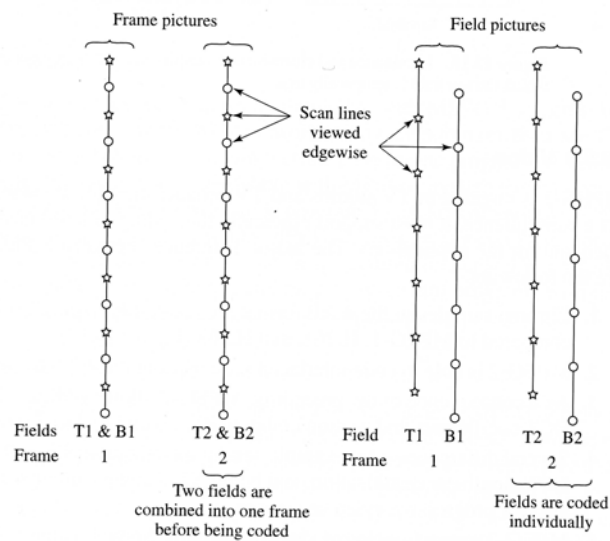


Imagen extraída de "Video processing and communications", Wang, Ostermann, Zhang, Prentice Hall, 2002

Modos de codificación de cuadro (V): modo cuadro I

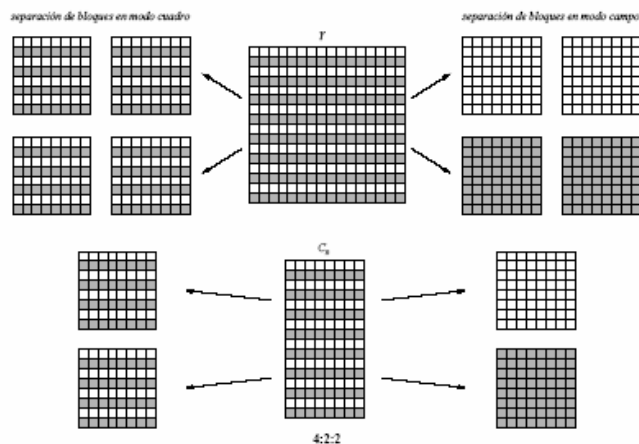
Cuadros tipo I: Se codifica sin predicción

- Idéntico a MPEG-1 (similar a JPEG)
 - Matrices de visibilidad
 - Coeficientes DC en modo predictivo
 - Coeficientes AC recorridos en zig-zag
 - Para cada MB se decide si se mantiene el nivel de cuantificación o si se modifica en función de la ocupación del buffer: factor de escala del cuantificador
Todos los bloques de un MB igual factor de escala
- Ventajas
 - Facilitan la decodificación (se puede decodificar nada más llegar el I)
 - Elimina (periódicamente) la propagación de errores de transmisión
 - Permiten la existencia de GOPs cerrados
 - Permiten la reproducción hacia atrás: entresacando y visualizando cuadros I en orden inverso
- Inconvenientes
 - Menor eficiencia (no reducción de redundancia por predicción)
 - Retardo adicional en buffer (por aumento de cantidad de información)

Modos de codificación de cuadro (VI): modo cuadro I

Para el procesamiento DCT el MB entrelazado se divide de forma diferente según se elija el modo cuadro o campo (modo cuadro/campo DCT)

- Modo cuadro o campo: a nivel MB, de forma que una misma imagen puede tener MB en modo cuadro y modo campo
- Si 4:2:0 no se puede desentrelazar la cromaancia



Modos de codificación de cuadro (VII): modo cuadro P

Cuadros tipo P: hacen predicciones con el cuadro anterior I o P.

- MB con predicción intracuadro. Si la señal es entrelazada se puede elegir en aplicar la DCT en modo cuadro o campo
- MB con modo predictivo: MB FW
 - Modo cuadro DCT: macrobloque de 16x16 para la predicción sobre I o P (si entrelazado, todos los MB usados en entrelazado). 1 único VM.
 - Modo campo DCT: macrobloque de luminancia en dos bloques de 16x8. Predicción con campo de igual paridad del cuadro I o P anterior o bien el campo anteriormente codificado en el cuadro. 2 VM (uno por bloque de 16x8)
- Se tiene que decidir (el como se decide depende del codec)
 - MC/no MC
 - Intra/no Intra
 - Cambio factor escala o no
 - Para modo no intra (MB FW), puede darse el caso que no haya nada que codificar (todos los coeficientes nulos –predicción “perfecta”-).
Codificar solamente VM o nada (skipped)

Modos de codificación de cuadro (VIII): modo cuadro P (*)

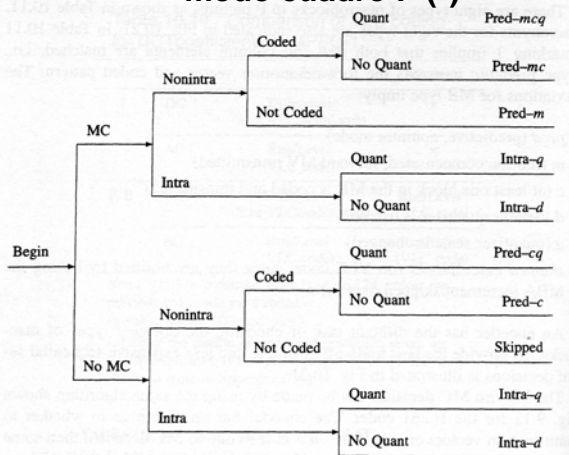


Figure 10.21 The macroblock type decision procedure in P pictures. *m*: motion compensated, *c*: at least one block coded, *q*: quantizer changed, *d*: default quantizer used, *pred*: predictive (forward) mode, *MC*: motion compensation mode

Imagen extraída de "Techniques and Standards for Image, Video and Audio Coding, Rao, Hwang, Prentice Hall, 1996

Modos de codificación de cuadro (IX): modo cuadro B

Cuadros tipo B: hacen uso de predicciones del cuadro anterior I o P y/o del cuadro posterior I o P.

- MB intracuadro (como cuadros P)
- MB FW (como cuadros P)
- MB BW: igual que MB FW pero con cuadro posterior I o P (1 o 2 VM)
- MB interpolado: con el promedio de la predicción FW y BW, cada una con un VM (o 2 si Modo Campo DCT)
- Nunca sirven de referencia
- Se tiene que decidir (el como se decide depende del codec)
 - o Mejor modo de compensación (se asume siempre MC inicialmente) y luego como P
 - Intra/no Intra
 - Cambio factor escala o no
 - Para modo no intra (MB FW), puede darse el caso que no haya nada que codificar (todos los coeficientes nulos –predicción “perfecta”-): Codificar solamente VM o nada (skipped)

Modos de codificación de cuadro (X): modo cuadro B (*)

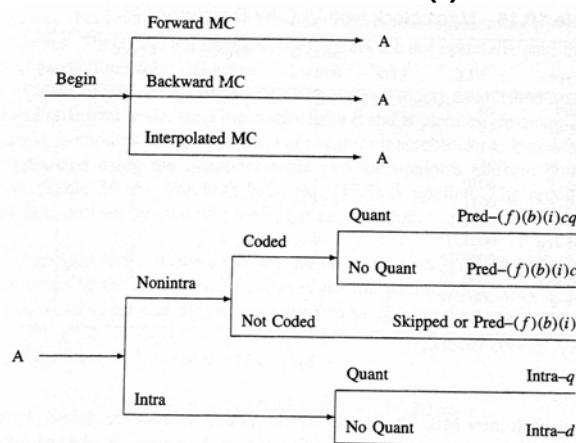


Figure 10.22 Macroblock type decision procedure in B pictures (*f*: forward; *b*: backward; *i*: interpolative).

Modos de codificación de cuadro (XI): modo cuadro B


- Ventajas cuadros B:
 - Mejora de calidad de las predicciones en el modo interpolado (se demuestra que la varianza del error –y por tanto la energía- se reduce a la mitad)
 - “Facilitan” la reproducción a velocidad rápida ya que ésta puede basarse solamente en cuadros I o P (que no dependen de los Bs)
- Inconvenientes cuadros B:
 - Aumento coste computacional (cálculo siempre de MC y tres modos de compensación)
 - Duplicar memoria de predicción (cuadro anterior y posterior)
 - Reducción de la eficiencia de predicción en los cuadros P, al aumentar la distancia entre ellos para la predicción con el I o P anterior, así como aumento de la dimensión de los VM para cubrir la posibilidad de mayores desplazamientos
 - Introducción de un retardo adicional en el proceso de recepción y decodificación (orden temporal, orden de proceso, orden de presentación)

Modos de codificación de cuadro (XII): modo cuadro



Adicionalmente hay que considerar

- el retardo de procesamiento – líneas no verticales
 - si el tiempo entre Ps es mayor que el procesamiento de las Bs, se añade un retardo–con jitter-
- la necesidad adicional de memoria para imágenes
- la necesidad de memoria/procesamiento en la tarjeta gráfica para compensar el “jitter” en presentación
 - Imágenes decodificadas

Escuela Politécnica Superior Universidad Autónoma de Madrid 

Modos de codificación de cuadro (XIII): modo campo

Modo 16/8

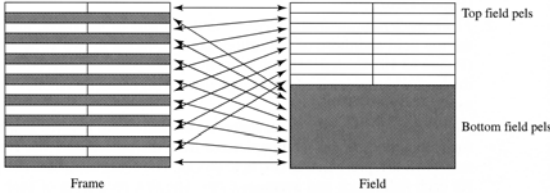


Figure 13.18 Field prediction for frame pictures: the MB to be predicted is split into top field pels and bottom field pels. Each 16×8 field block is predicted separately with its own motion vector (P-frame) or two motion vectors (B-frame).

Alternate Scan

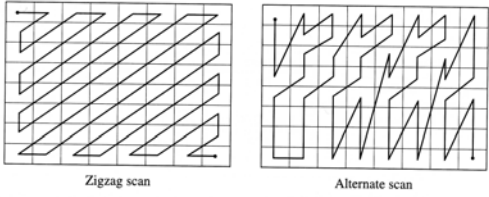



Figure 13.19 The zigzag scan as known from H.261, H.263, and MPEG-1 is augmented by the alternate scan in MPEG-2, in order to code interlaced blocks that have more correlation in the horizontal than in the vertical direction.

Imagen extraída de "Techniques and Standards for Image, Video and Audio Coding, Rao, Hwang, Prentice Hall, 1996"

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG-2 Vídeo (29)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

Modos de codificación de cuadro (XIV): modo campo

Campos I:

- Para el primer campo como Cuadros I en modo cuadro
- Para el segundo campo
 - Como cuadros I en modo cuadro
 - Con predicción FW con el campo anterior (sigue siendo decodificable a nivel imagen –requisito cuadros I-)

Campos P:

- MB intra
- MB FW
 - Para cada MB el MB de igual paridad del I o P anterior si campo T, o MB de igual paridad del I o P anterior y el de paridad contrario del mismo cuadro recién codificado si B
 - Para cada MB empleado en la predicción
 - MB 16×16 (1 VM)
 - Modo 16/8: mitad superior/inferior (2 VM)
- MB Dual Prime (en modo campo solamente)
 - Se refinan los VM transmitidos: uno a MB de campo de igual paridad de I o P anterior y otro a MB de campo de distinta paridad de I o P anterior
 - 2 VM

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: MPEG-2 Vídeo (30)

Modos de codificación de cuadro (XV): modo campo

Campos tipo B:

- MB intra (como en campos P)
- MB FW (como en campos P)
- MB BW (como en campos P MB FW, pero con I o P posterior)
- MB interpolado (promediando MB BW y FW). 2 VM si bloques 16x16, 4 si modo 16/8
 - o Pero con campos de frames anteriores, nunca del anterior en caso B

Modos de codificación de cuadro (XVI): modo campo

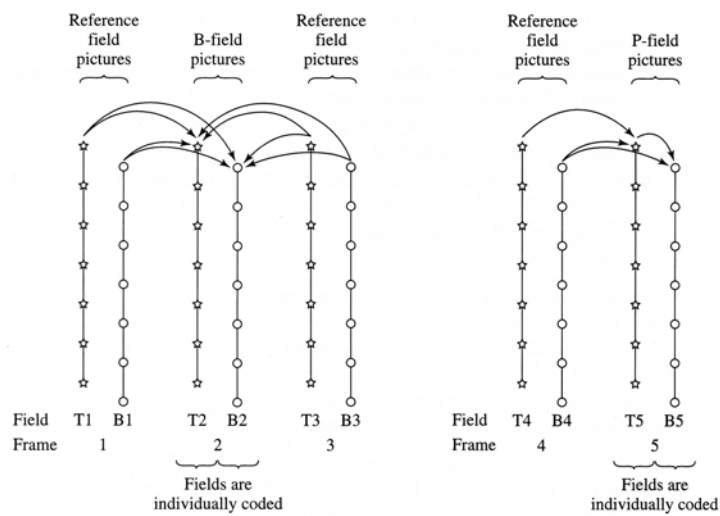


Imagen extraída de "Video processing and communications", Wang, Ostermann, Zhang, Prentice Hall, 2002



Modos de codificación de cuadro (XVII): procesamientos a nivel de MB

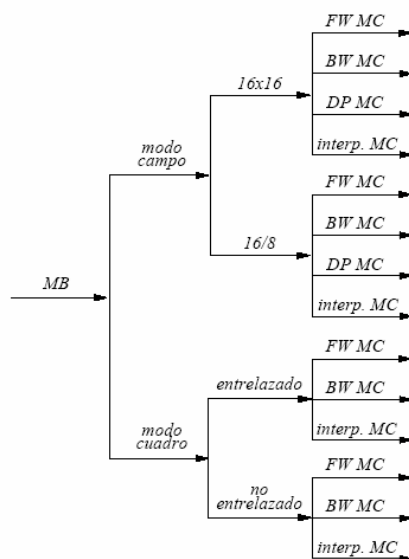
Todas las posibles decisiones a nivel de MB, si bien en función del tipo de cuadro (I,P,B) solamente algunas estarán permitidas

No se muestra por simplificar

- La primera decisión es si se usa MC o no MC
- Tras la decisión de los resultados de la compensación a utilizar: Intra/no Intra, Coded/No Coded, Quant/no Quant
- Skipped MB



Modos de codificación de cuadro (XVIII): procesamientos a nivel de MB (*)





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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- *Modos de codificación de cuadro*
- **Escalabilidad**
- Perfiles y niveles
- Calidades MP@ML
- Bibliografía



Escalabilidad (I)

Un flujo de datos se define como escalable cuando tiene la propiedad de que parte del mismo puede recuperarse y decodificarse de manera independiente del resto.

Una razón de la escalabilidad es permitir que existan receptores de menor coste (capacidad de procesamiento) que permitan mostrar el vídeo a menor resolución espacial, temporal o de menor calidad.

Adicionalmente permite la división del flujo binario en función de la prioridad de transmisión en presencia de errores (multiplexación por canales con distinta QoS)

MPEG-2 considera un máximo de 3 niveles de escalabilidad

- 1 Nivel base y 1 ó 2 de mejora

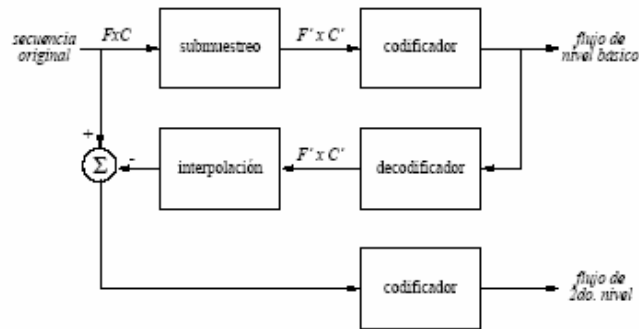
y distintos tipos

- Espacial, SNR, Temporal, Partición de datos, Híbrida

Escalabilidad (II)

Escalabilidad \Leftrightarrow distintas resoluciones

- El nivel superior (mejora) se genera a partir de la imagen de nivel inferior
 - Si espacial/temporal reconstruida e interpolada



Escalabilidad (III)

Scaled for slides (full res.)



bitstream

Escalabilidad (IV)

Escalabilidad espacial: distintas resoluciones espaciales

- El nivel superior (mejora) se genera a partir de la imagen de nivel inferior reconstruida e interpolada
- Existe flexibilidad para definir la resolución de las imágenes en los distintos niveles, por ejemplo usar una subimagen de la completa como origen para el nivel base.
 - Uso para transmisión de HDTV con TV insertada

Escalabilidad SNR (Signal to Noise Ratio)

- Cada nivel contiene mayor precisión de los coeficientes de la DCT. Todos los niveles igual resolución.

Escalabilidad temporal

- Similar a la espacial, pero submuestreando e interpolando en el tiempo (eliminando imágenes de la secuencia)

Partición de datos

- Codificación en modo fundamental MPEG-2 (sin escalabilidad)
- Organización de los datos en el flujo binario por niveles de prioridad
 - Nivel básico: cabeceras, VMs, coeficientes DC, ...
 - Nivel superior: coeficientes DCT de alta frecuencia

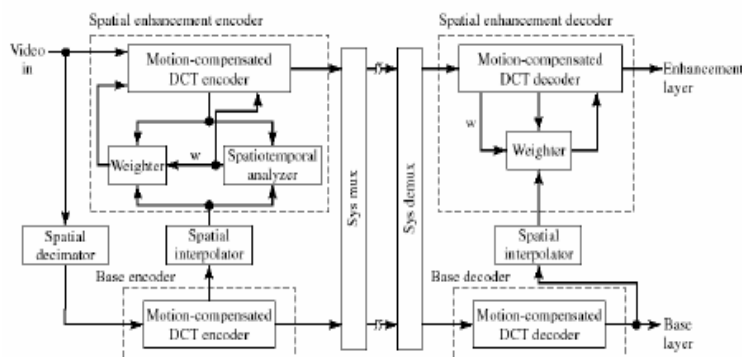
Híbrida

- Combinación de escalabilidad SNR, espacial y temporal
- 1 nivel básico y 2 de mejor

Escalabilidad (V) (*)

Escalabilidad espacial: distintas resoluciones espaciales

- El nivel superior (mejora) se genera a partir de la imagen de nivel inferior reconstruida e interpolada



Escalabilidad (VI) (*)

Escalabilidad SNR (Signal to Noise Ratio)

- Cada nivel contiene mayor precisión de los coeficientes de la DCT. Todos los niveles igual resolución.

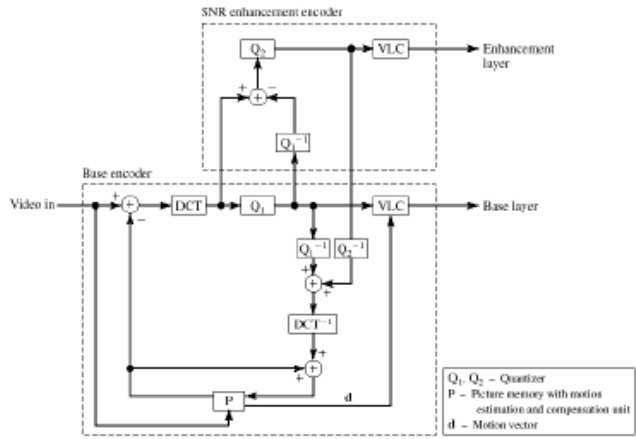


Imagen extraída de "Video processing and communications", Wang, Ostermann, Zhang, Prentice Hall, 2002

Escalabilidad (VII) (*)

Escalabilidad temporal

- Similar a la espacial, pero submuestreando e interpolando en el tiempo (eliminando imágenes de la secuencia)

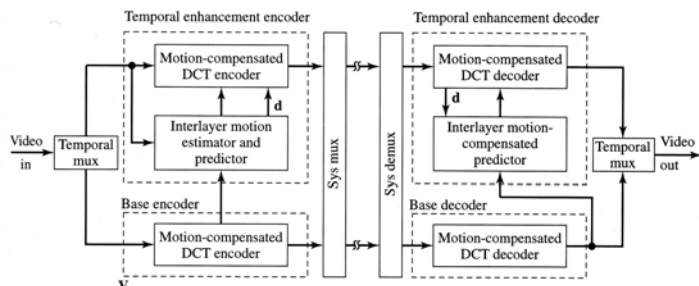


Figure 13.25 A temporal scalability encoder consists of two complete encoders, with the enhancement encoder using the base-layer video as an additional reference for prediction. The temporal demux sends pictures alternately to the base encoder and the enhancement encoder.

Imagen extraída de "Video processing and communications", Wang, Ostermann, Zhang, Prentice Hall, 2002

Escalabilidad (VIII) (*)

Partición de datos

- Codificación en modo fundamental MPEG-2 (sin escalabilidad)
- Organización de los datos en el flujo binario por niveles de prioridad
 - o Nivel básico: cabeceras, VMS, coeficientes DC, ...
 - o Nivel superior: coeficientes DCT de alta frecuencia

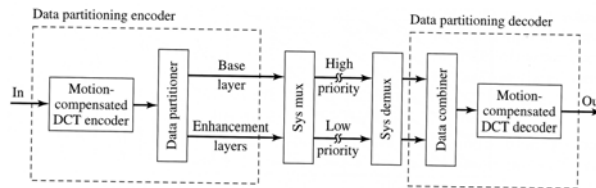


Figure 13.20 A data partitioning codec suited for ATM networks that support two degrees of quality of service.

Imagen extraída de "Video processing and communications", Wang, Ostermann, Zhang, Prentice Hall, 2002

Híbrida

- Combinación de escalabilidad SNR, espacial y temporal
- 1 nivel básico y 2 de mejora

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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- *Modos de codificación de cuadro*
- *Escalabilidad*
- **Perfiles y niveles**
- *Calidades MP@ML*
- *Bibliografía*

Perfiles y niveles

MPEG-2 Vídeo debe dar solución a aplicaciones muy variadas.

Al igual que en otras partes de MPEG-2 (y en MPEG-X; $X \geq 2$) se crean una serie de subconjuntos de la especificación orientados a la creación de productos orientados a aplicación y con coste acorde al mercado.

Estos subconjuntos se definen mediante los conceptos de:

- **Perfil:** complejidad a nivel de número de herramientas
 - conjunto de herramientas de compresión empleado
 - compromiso entre compresión y coste del decodificador.
 - MPEG-2 vídeo: formato de muestreo, cuadros B, escalabilidad.
 - MPEG-2 vídeo: 5 (Simple, Principal, SNR, Espacial, Alto) +2 (4:2:2 –similar a principal-, MultiViewProfile -para codificación conjunta de la misma escena con múltiples cámaras-):
- **Nivel:** complejidad a nivel de prestaciones
 - conjunto de valores máximos de parámetros soportados por la implementación
 - MPEG-2 vídeo: dimensiones, cuadros/s, velocidad binaria
 - MPEG-2 vídeo: 4 (bajo, principal, alto-1440, alto)
 - No todos los niveles se soportan en todos los perfiles

TVD: Main Profile at Main Level (MP@ML)

Perfiles

Elemento Sintáctico	PERFIL						
	Simple	Principal	SNR	Espacial	Alto	4:2:2	MVP
Muestreo Crominancia	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0 4:2:2	4:2:0 4:2:2	4:2:0
Cuadros B	NO	SI	SI	SI	SI	SI	SI
Escalabilidad	NO	NO	SNR	SNR Espacial	SNR Espacial	NO	Temporal
Nr. Niveles Escalabilidad	0	0	2	3	3	0	2

Niveles

Parámetro	NIVEL			
	Alto	Alto-1440	Principal	Bajo
Columnas	1920	1440	720	352
Filas	1152	1152	576	288
Cuadros/seg.	60	60	30	30
Velocidad (Mbps)	90	60	15	4

Combinaciones Perfiles-Niveles

		Profile						
		Simple (L, P)	Main (L, P, B)	SNR (L, P, B)	Spatial (L, P, B)	High (L, P, B)	Multiview (L, P, B)	4:2:2 (L, P, B)
		(4:2:0)	(4:2:0)	(4:2:0)	(4:2:0)	(4:2:0; 4:2:2)	(4:2:0)	(4:2:0; 4:2:2)
Low	Pels/line			352	352		352	
	Lines/frame			288	288		288	
	fps			30	30		30	
	mbps			4	4		8	
Main	Pels/line	720	720	720		720	720	720
	Lines/frame	576	576	576		576	576	512/608
	fps	30	30	30		30	30	30
	mbps	15	15	15		20	25	50
Level	High-1440	Pels/line	1440		1440	1440	1440	
	Lines/frame		1152		1152	1152	1152	
	fps		60		60	60	60	
	mbps		60		60	80	100	
High	Pels/line	1920			1920	1920	1920	
	Lines/frame		1152		1152	1152	1152	
	fps		60		60	60	60	
	mbps		80		100	130	300	

L, P, B: allowable picture types. Maximum bit rates include all layers in case of scalable bit streams.



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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- *Modos de codificación de cuadro*
- *Escalabilidad*
- *Perfiles y niveles*
- **Calidades MP@ML**
- *Bibliografía*



Calidades MPEG-2 MP@ML

Calidad objetiva:

- SNR o PSNR como medidas objetivas.
- Evaluables directamente sobre las imágenes transmitidas.

Calidad subjetiva:

- Dependiente del tipo de programa (contenidos).
- Evaluable a través de pruebas con individuos siguiendo protocolos establecidos (estándares).
- Relación relativa con las medidas de calidad objetiva.

Velocidades binarias (orientativas):

- 2 Mb/s: apto para señales muy simples (ej. dibujos animados).
- 4-6 Mb/s: calidad PAL. Programación típica.
- 8-9 Mb/s: calidad de estudio. Programas especiales.



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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- *Modos de codificación de cuadro*
- *Escalabilidad*
- *Perfiles y niveles*
- *Calidades MP@ML*
- *Bibliografía*



Bibliografía

- ISO/IEC 11172, Information Technology: Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbps, 1993*
- ISO/IEC 13818, Information Technology: Coding of moving pictures and associated audio, 1994*
- K.R. Rao, J.J Hwang, *Techniques and standards for image, vídeo and audio coding*, Prentice-Hall, 1996
- L. Salgado, C. Muñoz, *Apuntes de Televisión Digital*, ETSIT-UPM, 2001
- Y. Wang, J. Ostermann, Y.Q. Zhang, *Vídeo Processing and Communications*, Prentice-Hall, 2002



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

INDICE

- *Introducción*
- *Estructuras de datos*
- *Modos de predicción*
- *Modos de codificación de cuadro*
- *Escalabilidad*
- *Perfiles y niveles*
- *Calidades MP@ML*
- *Bibliografía*



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

2009-2010



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 (I)

Principios de la codificación de audio MPEG (natural)

- **Codificación subbandas**
 - Imita el mecanismo de análisis frecuencial del oído. Modelo de banco de filtros.
 - La señal se pasa a un dominio espectral
 - Codifica cada banda con diferente resolución (calidad)
- **Codificación perceptual**
 - 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-)
 - Idea: MENOS bits (o ninguno) para los sonidos MENOS audibles. Se codifica con menos bits con la misma calidad de audio percibida (calidad perceptual)
 - Tiene que analizar la señal continuamente para determinar el umbral de audibilidad en cada instante
- **Otras tecnologías utilizadas**
 - Conmutación de ventanas
 - Asignación dinámica de bits

Introducción (II)

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)
- Lo realmente usado 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

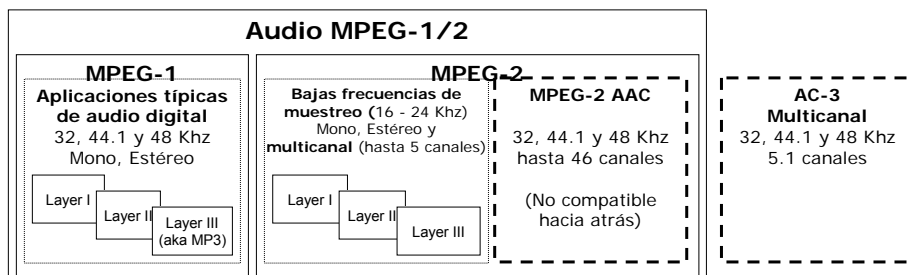
MPEG-4 Audio

- 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 (III)





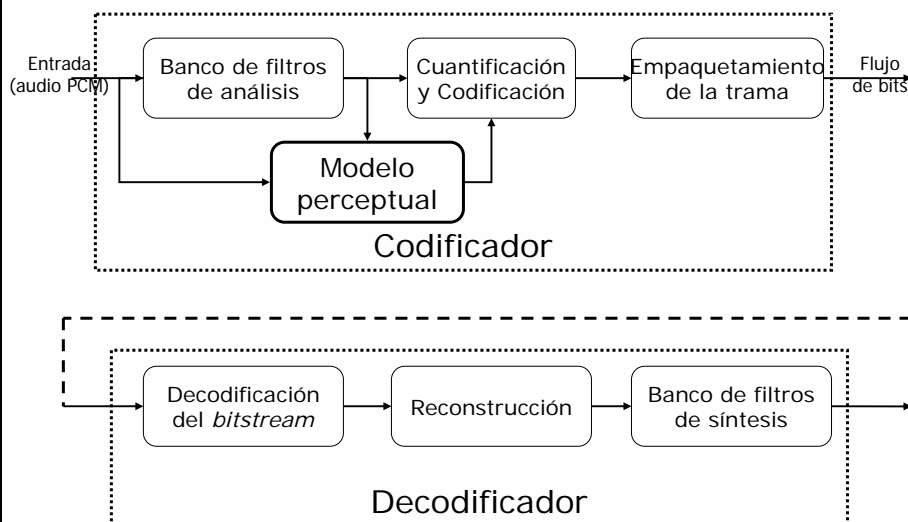
Representación y Codificación de Señales Audiovisuales en Televisión Digital MPEG Audio

INDICE

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



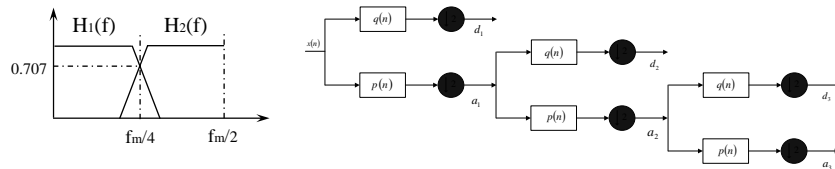
Audio MPEG-1 (I): Códec perceptual en subbandas (*)



Audio MPEG-1 (II): Codificación en subbandas (*)

Banco de filtros QMF

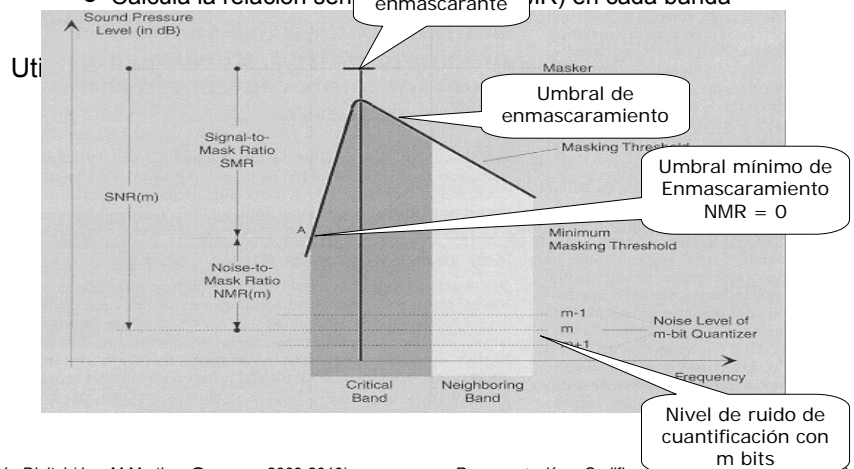
- 32 bandas iguales
- Filtros QMF (Quadrature Mirror Filters)
 - Producen dos subbandas iguales
 - Son filtros de reconstrucción perfecta (PF)
 - El *aliasing* que introducen los filtros de análisis se cancela en los filtros de síntesis
 - Se pueden poner varias etapas en cascada para obtener más subbandas



Audio MPEG-1 (III): Modelo psicoacústico (*)

Modela los umbrales de enmascaramiento

- Calcula la relación señal (SNR) enmascarante (MR) en cada banda





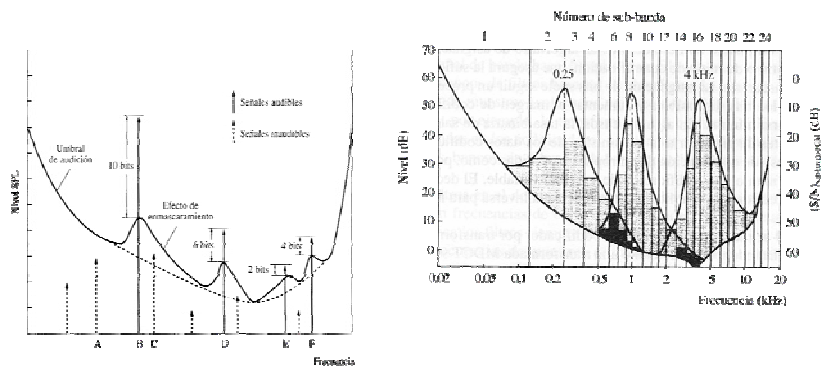
Audio MPEG-1 (IV): Modelo psicoacústico (*)

Modelo de enmascaramiento

- Dos tipos de enmascaramiento
 - Tono enmascarando otro tono (umbral $\approx -14.5 - P_{\text{tono}}$ dB)
 - Tono enmascarando ruido (umbral ≈ -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



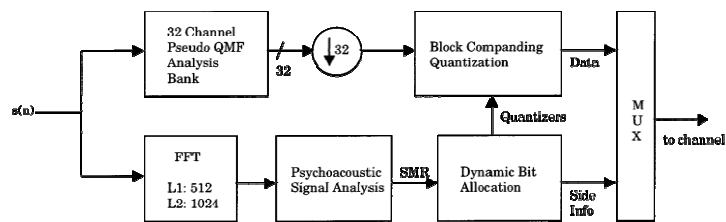
Audio MPEG-1 (V): Modelo psicoacústico (*)



Audio MPEG-1 (VI): 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
 - o FFT 512 puntos (resolución 94 Hz para frecuencia muestreo de 48 kHz)
- Utilizado por Philips en el Digital Compact Cassette (DCC)



Audio MPEG-1 (VII): 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
 - o 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).



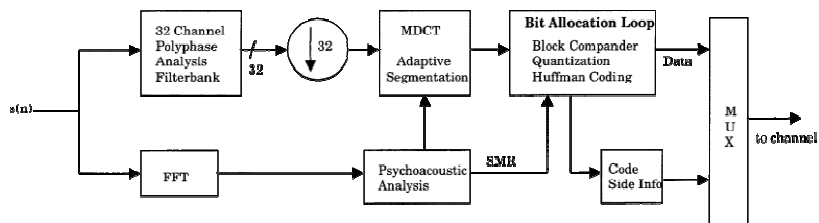
Audio MPEG-1 (VIII): Capa III (*)

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)
 - o 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, ...)



Audio MPEG-1 (IX): Capa III (*)





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-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

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

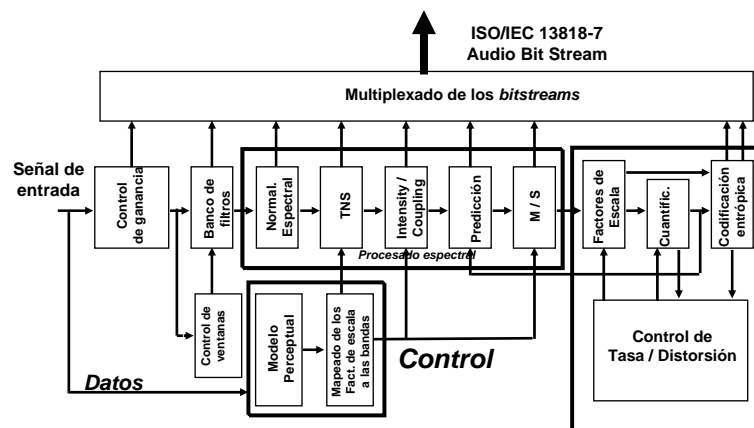
Sigue el modelo de MPEG-1 Audio Layer III, pero con mejoras

- 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

Audio MPEG-2 (III): Advanced Audio Coding (AAC) (*)





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



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-4 (*)

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

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

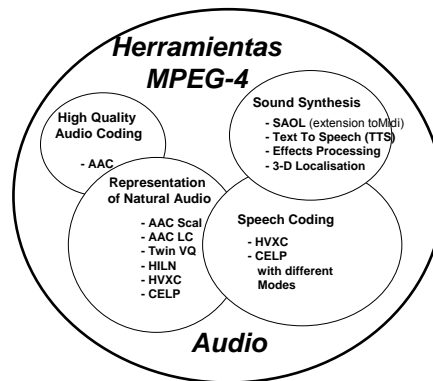
Codificación escalable

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



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

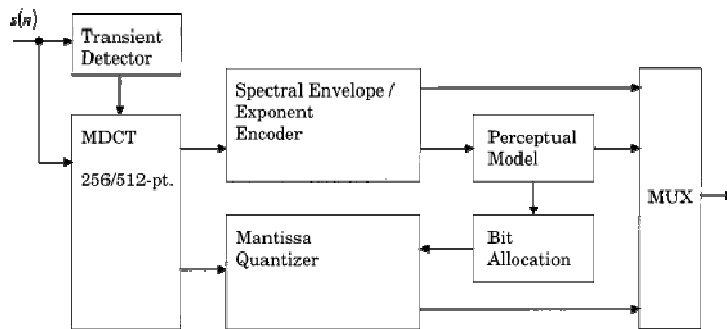
- El exponente indica la envolvente espectral

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

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



Audio multicanal Dolby AC-3 (II) (*)





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*



Créditos

Estas transparencias están basadas en un trabajo de Luis Herranz.