

# Tema 4

---

## Técnicas de Compresión de audio y vídeo

# Contenido

---

- Familia MPEG
  - MPEG-1
    - Vídeo y Audio
  - MPEG-2
    - Vídeo, Audio y Datos
  - MPEG-4
    - Vídeo y Audio
  - MPEG-7
- Familia H.26x:
  - H.261 y H.263
- Resumen

# MPEG



**MPEG: familia de estándares de compresión de vídeo**

# MPEG



**MPEG: familia de estándares de compresión de vídeo**

**MPEG-1**

## MPEG (I)

---

- Motion Picture Expert Group, ISO IEC/JTC1/SC29/WG11
- Combina streams de audio y de video en un único bit stream.
- Se compone de una capa de sistema y de capas de compresión.
- La capa del sistema contiene la información para la demultiplexación del resto de las capas.

## MPEG (II)

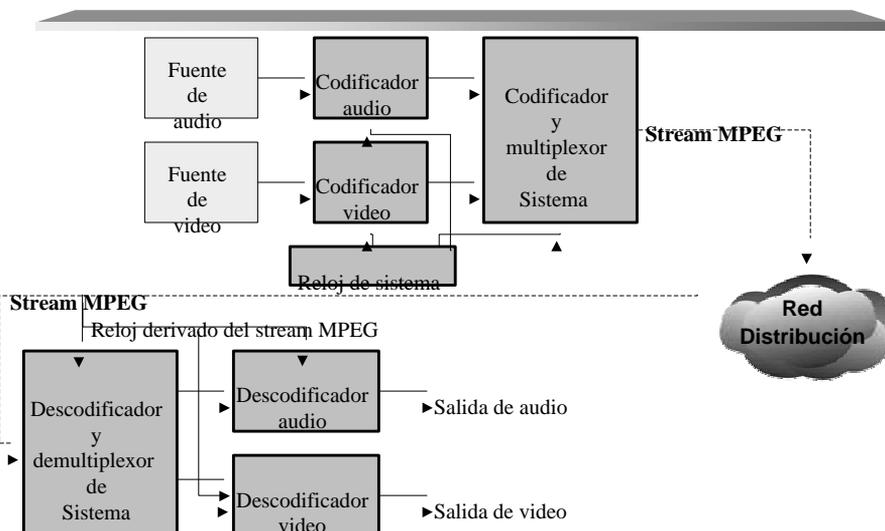
---

- MPEG es un estándar genérico que incide en compresión de video, audio y la sincronización audio-video.
- MPEG comprime audio y video a un rate de 1.5 Mbps.
- Los factores de compresión a los que llega son: 50:1 a 200:1.
- Algoritmo asimétrico: requiere más recursos para comprimir que para descomprimir

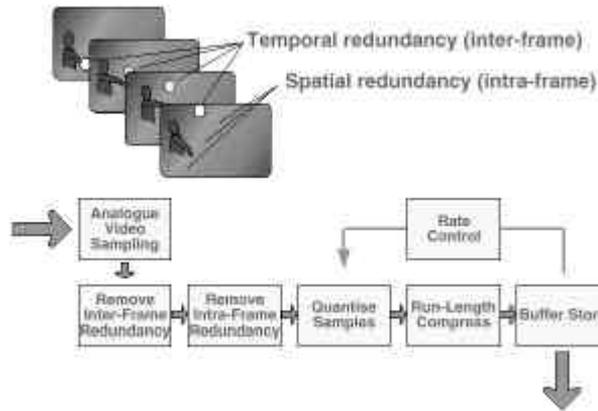
## Compresión MPEG

- Aprovecha tanto la redundancia espacial de cada imagen como la redundancia temporal del video en movimiento.
- MPEG incorpora el concepto de Group of Pictures (GOP).
- Tres tipos de frames: Intra Pictures (I), Predicted Pictures (P), Bi-directional Pictures (B).

## Estructura del Sistema MPEG



## Esquema de un codificador MPEG

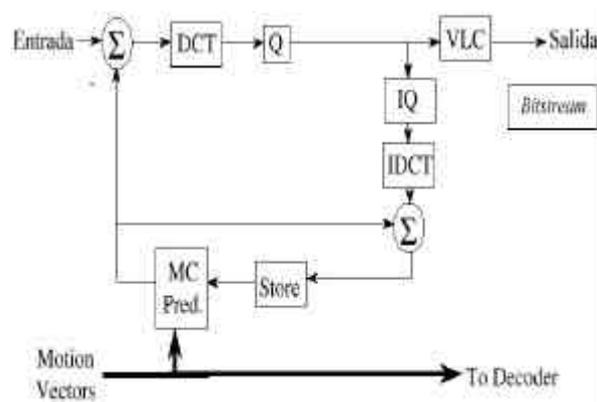


## MPEG: codificador típico

IQ: cuantificación inversa

MC Pred.: Predictor de compensación de mov.

IDCT: Transformada del coseno inversa



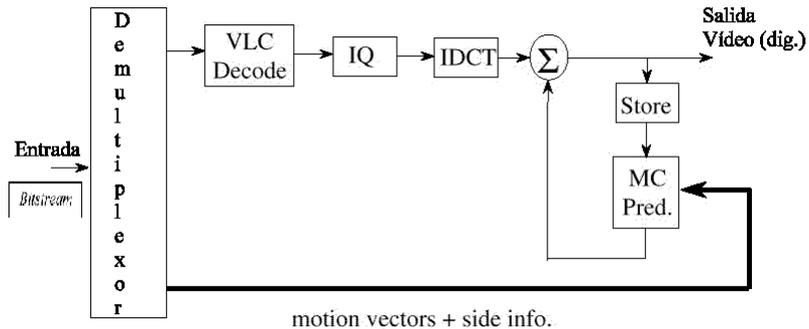
## MPEG: descodificador típico

IQ: cuantificación inversa

MC Pred.: Predictor de compensación de mov.

IDCT: Transformada del coseno inversa

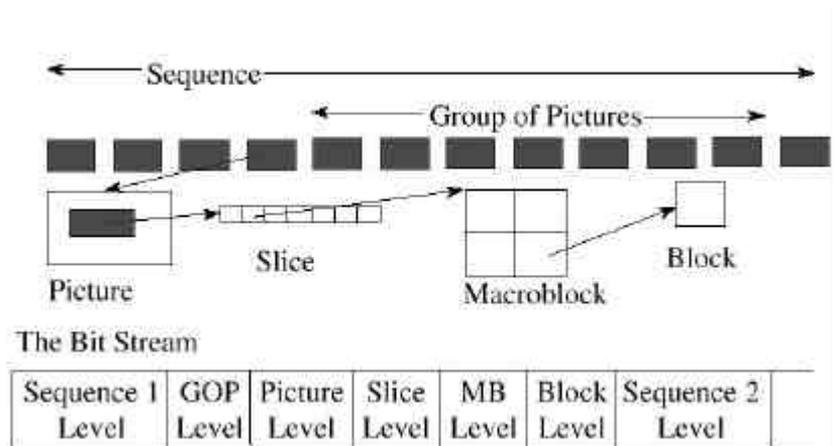
Side info: Vectores de desplazamiento, parámetros de cuantificación, tablas de codificación, ...



Sistemas Multimedia lagarcia

11

## Capas de MPEG bit-stream (I)



Sistemas Multimedia lagarcia

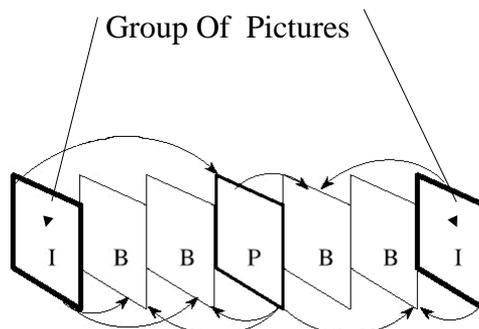
12

## Capas de MPEG bit-stream (II)

- Sequence
  - Imagen y relación de aspecto
  - Velocidad de imagen
  - Bit rate
  - Tamaño mínimo del buffer del decodificador
  - Restricciones de los parámetros
  - Opcional:
    - Tabla de cuantificación intra
    - Tabla de cuantificación no-intra
    - Datos de usuario

## Capas de MPEG bit-stream (III)

- Group of Picture (GOP)
  - Códigos de tiempo (timecode)
  - Flag de GOP contiguo
  - Flag de link perdido
  - Opcional:
    - Datos de usuario



## Capas de MPEG bit-stream (IV)

---

- Picture
  - Referencia temporal
  - Tipo de frame (I/P/B)
  - Ocupación inicial del buffer de decodificación
  - Opcional:
    - Resolución y rango del vector de estimación de movimiento (forward)
    - Resolución y rango del vector de estimación de movimiento (backward)
    - Datos de usuario

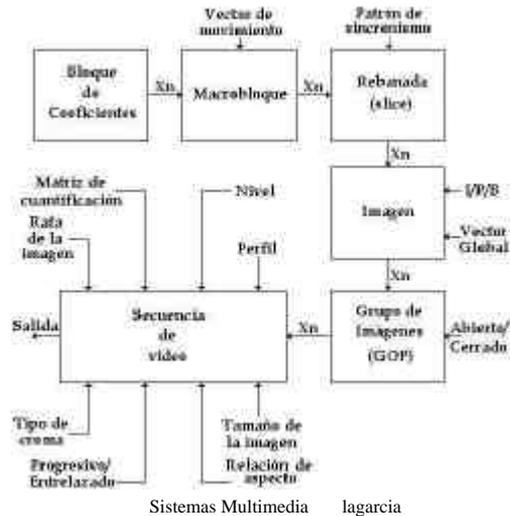
## Capas de MPEG bit-stream (V)

---

- Slice
  - Posición vertical del slice
  - Escala del cuantificador
- Macrobloque
  - Incremento de dirección
  - Tipo de macrobloque
  - Opcional:
    - Escala del cuantificador
    - Vector de movimiento (forward)
    - Vector de movimiento (backward)
    - Patrón de codificación del bloque
    - Coeficientes de la DCT de hasta 6 bloques (4+1+1)

## Capas de MPEG bit-stream (VI)

- *Esquema resumen de los parámetros de las distintas capas MPEG*



17

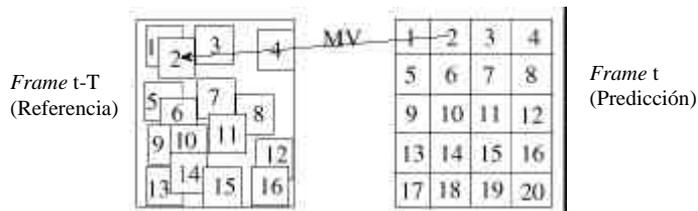
## Compresión secuencias de vídeo

- *Compresión Intraframe*: trata cada *frame* de una secuencia como una imagen fija (*still*). Sólo explota la redundancia espacial.
- *Compresión Interframe*: utiliza predicciones temporales, por lo que aprovecha las redundancias temporal y espacial, aumentando la eficiencia de la compresión.
  - Ejemplo: Compresión predictiva por compensación temporal del movimiento

## Compresión secuencias de vídeo

- **Compresión *Interframe*. Principios:**

- El *frame* de referencia debe estar disponible en el codificador/descodificador.
- Los vectores de movimiento (*MV*) y los *Prediction Error Blocks* (*PEB*) son codificados y transmitidos.
- A partir de los *MV*'s y los *PEB*'s, el *frame* puede reconstruirse en el descodificador.



## MPEG-1: Principios generales

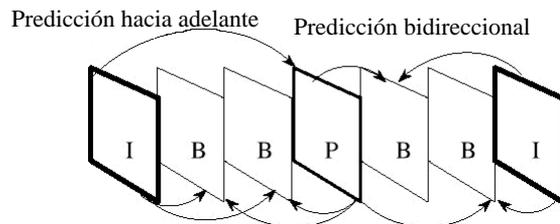
- Las actividades de estandarización de MPEG-1 se basan en la posibilidad de almacenar y recuperar vídeo, y su audio asociado, a una tasa binaria de  $\cong 1.5$  Mbps
- Soporta simultáneamente los modos de compresión *intraframe* e *interframe*.
- El estándar considera:
  - Vídeo representado por luminancia y dos canales de color. Estos canales están submuestreados por un factor 2 en ambas direcciones.
  - 8 bit/pixel

## MPEG-1: Principios generales

- MPEG-1 estandariza una sintaxis para la representación del bit-stream codificado y un método de decodificación.
  - La sintaxis soporta las operaciones de:
    - Transformada Discreta del Coseno (DCT),
    - predicción por compensación de movimiento,
    - cuantificación, y
    - codificación de longitud variable (VLC).
  - Hay una gran flexibilidad en el diseño del codificador.
    - Por ejemplo, no estandariza el algoritmo de estimación de movimiento.

## MPEG-1: Modos básicos de compresión

- Asignación típica de los modos de compresión de *frames* de una secuencia.



## MPEG-1: Modos básicos de compresión (II)

---

- Compresión *intraframe*:
  - Los *frames* I representan *frames* codificados exclusivamente con compresión *intraframe*.
  - El objetivo de estos *frames* es proporcionar puntos de acceso aleatorio a la secuencia de vídeo.

## MPEG-1: Compresión *interframe*

---

- Compresión *interframe*:
  - En los *frames* P se utiliza compresión predictiva de compensación de movimiento sobre los macrobloques.
    - Hay predicción de los macrobloques basados en los frames I y P recientemente descodificados.
    - Se codifican vectores de movimiento y señales de error.
  - En los *frames* P, está permitido utilizar compresión intra si la predicción no se considera buena.

## MPEG-1: Compresión *interframe* (II)

- Compresión *interframe*:
  - En los *frames* B se utiliza compresión predictiva bidireccional de compensación de movimiento sobre los macrobloques.
    - Hay predicción de los macrobloques basados en los frames I y P recientemente descodificados.
    - Se codifican vectores de movimiento y señales de error.
  - En los *frames* B, está permitido utilizar compresión intra si la predicción no se considera buena.

## MPEG-1: Modos básicos de compresión (III)

- El número relativo de imágenes I, P y B no está determinado.
  - Depende del tipo de aplicación.
  - La única limitación es un período máximo de aparición de un *frame* I cada 132 *frames*.
- Las imágenes B son las que proporcionan mayores niveles de compresión en condiciones favorables de predicción.

## MPEG-1: Modos básicos de compresión (IV)

- Algunos ejemplos de secuencias:

I B B P B B I B B P B B I  
I I I I I I I I I I I I I I  
I P I P I P I P I P I P I P

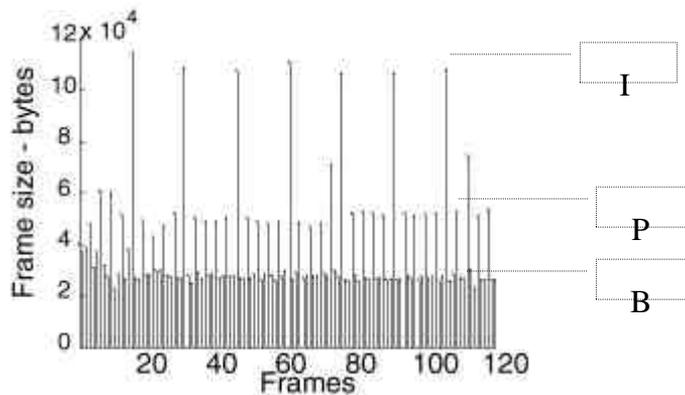
- Orden de presentación, codificación y descodificación:

- E. codificador/ S. decodificador/ Presentación: 1 2 3 4 5 6 7  
    » Tipo de *frame*: I B B P B B I
- S. codificador/ *Bitstream* / E. decodificador: 1 4 2 3 7 5 6

## MPEG-1: Tamaño *frame* por tipo

- Tamaño de una imagen según el tipo de *frame*:

- I
- P
- B



## MPEG-1: Etapas de codificación

- *Intraframe*:
  - Transformación
  - Cuantificación y codificación entropía
- *Interframe*:
  - Estimación de movimiento
  - Predicción hacia delante por compensación de mov.
  - Predicción bidireccional por compensación de mov.
  - Transformación
  - Cuantificación y codificación de entropía

## MPEG-1: codificación *intraframe*

- Transformación:
  - Transformada discreta del coseno aplicada a bloques de 8x8 de luminancia y color.

$$F(u,v) = \frac{1}{4} C(u)C(v) \sum_{m=0}^7 \sum_{n=0}^7 f(m,n) \cos\left(\frac{\pi(2m+1)u}{16}\right) \cos\left(\frac{\pi(2n+1)v}{16}\right)$$

Donde:

$$u,v, m, n = 0, 1, 2, \dots, 7$$

$$C(w) = \begin{cases} 1/\sqrt{2} & \text{para } w=0 \\ 1 & \text{en el resto} \end{cases}$$

## MPEG-1: codificación *intraframe* (II)

- Ejemplo:

Bloque inicial M	Bloque transformado NINT [ DCT (M) ]
139 144 149 153 155 155 155 155	1260 -1 -12 -5 2 -2 -3 1
144 151 153 156 159 156 156 156	-23 -17 -6 -3 -3 0 0 -1
150 155 160 163 158 156 156 156	-11 9 -2 2 0 -1 -1 0
159 161 162 160 159 159 158 159	-7 -2 0 1 1 0 0 0
159 160 161 162 162 155 155 155	-1 -1 1 2 0 -1 1 1
161 161 161 161 160 157 157 157	2 0 2 0 -1 1 1 -1
162 162 161 163 162 157 157 157	-1 0 0 -1 0 2 1 -1
162 162 161 161 163 158 158 158	-3 2 -4 -2 2 1 -1 0

NINT: truncado al entero más próximo

## MPEG-1: codificación *intraframe* (III)

- Cuantificación:
  - Los coeficientes de la transformada se cuantifican uniformemente.
  - Coeficiente DC, se divide por 8 y el resultado se trunca al entero más próximo en  $[-256, 255]$

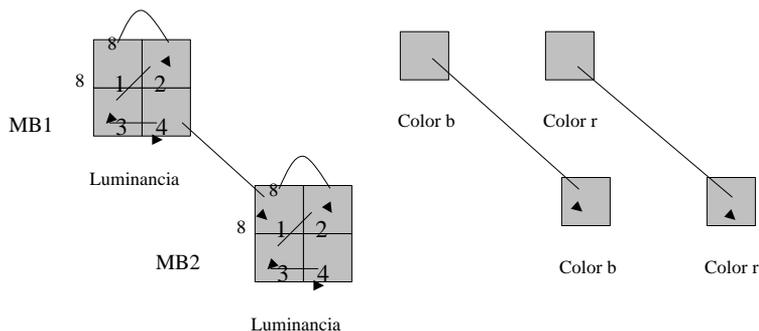
$$QF(0,0) = \text{NINT}[ F(0,0)/8 ]$$

## MPEG-1: codificación *intraframe* (IV)

- **Coefficientes DC:**
  - La redundancia entre los coeficientes DC de los bloques se reduce con codificación diferencial (DPCM)
  - La señal diferencial resultante (acotada en  $[-255, 255]$ ), se codifica con códigos de longitud variable (Huffman). Las tablas de este código están estandarizadas por MPEG-1.

## MPEG-1: codificación *intraframe* (V)

- **Coefficientes DC:**
  - Los macrobloques son de  $16 \times 16$ , pero la DCT se calcula sobre bloques de  $8 \times 8$ , la relación es:



## MPEG-1: codificación *intraframe* (VI)

- Cuantificación de los coeficientes AC:

$$QF(u,v) = \text{NINT}[ 16 \cdot F(u,v)/w(u,v) \cdot \text{escala} ]$$

Donde:

- $w(u,v)$ , matriz de pesos o de cuantificación. 8 bits
  - escala: parámetro del cuantificador, facilita la cuantificación adaptativa.
- El resultado se trunca al entero más próximo en  $[-256, 255]$
  - MPEG-1 especifica, por defecto, la siguiente matriz de cuantificación.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

## MPEG-1: codificación *intraframe* (VII)

- En los coeficientes de alta frecuencia se utiliza una cuantificación gruesa, con grandes pesos, para aprovechar la insensibilidad del sistema visual al ruido de cuantificación de alta frecuencia.

Ejemplo de coeficientes cuantificados:

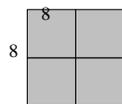
158	-1	-5	-2	1	-1	-1	0
-12	9	-2	-1	-1	0	0	0
-5	-6	-2	-1	-1	0	0	0
-3	-1	0	0	0	0	0	0
0	0	0	1	0	0	0	0
1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0
-1	1	-1	0	0	0	0	0

## MPEG-1: codificación *intraframe* (VIII)

- MPEG-1 permite el uso/descarga de cualquier matriz específica de la aplicación en el *bitstream* diferente a la definida por defecto.
- En general, la matriz de cuantificación depende del nivel de ruido de la fuente y de las condiciones de visión.

## MPEG-1: codificación *intraframe* (IX)

- Cuantificación espacial adaptativa.
  - Es posible por el parámetro escala que pondera los valores de la matriz de cuantificación ( $w(u,v)$ ).
  - Este parámetro está permitido que cambie de un macrobloque a otro dentro del mismo *frame* para ajustar la cuantificación dependiendo de los macrobloques.



Luminancia



Color b



Color r

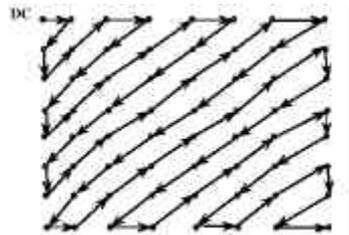
## MPEG-1: codificación *intraframe* (X)

- Cuantificación espacial adaptativa.
  - Los valores permitidos del parámetro escala son:  
2, 4, 6, 7, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60 y 62.
- Coeficientes AC:
  - La idea principal se basa en que la mayoría de los coeficientes cuantificados son cero y es más eficiente representar los datos por la posición y valor de los que no son cero.

## MPEG-1: codificación *intraframe* (XI)

- Coeficientes AC:
  - Se ordenan en zigzag y se codifican como símbolo = [run, level]
    - » run: número de coeficients de valor cero que le preceden
    - » level: valor del coeficiente
  - A continuación se codifican con un método Huffman de longitud variable (VLC). Las tablas de este método están estandarizadas.

Zig-zag de los coeficientes de AC:



## MPEG-1: codificación *intraframe* (XII)

- Ejemplo de codificación de coeficientes AC:

158	-1	-5	-2	1	-1	-1	0
-12	9	-2	-1	-1	0	0	0
-5	-6	-2	-1	-1	0	0	0
-3	-1	0	0	0	0	0	0
0	0	0	1	0	0	0	0
1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0
-1	1	-1	0	0	0	0	0

run	level
0	-1
0	-12
0	-5
0	9
0	-5
.....	
0	-1
EOB	

## MPEG-1: codificación *interframe* (I)

- Estimación de movimiento:
  - La predicción del movimiento se realiza en base a los macrobloques.
  - Los vectores de desplazamiento se consideran constantes para todo el macrobloque.
  - Por tanto, el vector de desplazamiento se estima para los 16x16 puntos de luminancia y los 2 bloques asociados de color de 8x8.
  - En el caso bidireccional (cuadros B), se calculan dos vectores una referido al pasado y el otro al futuro para cada macrobloque.

## MPEG-1: codificación *interframe* (II)

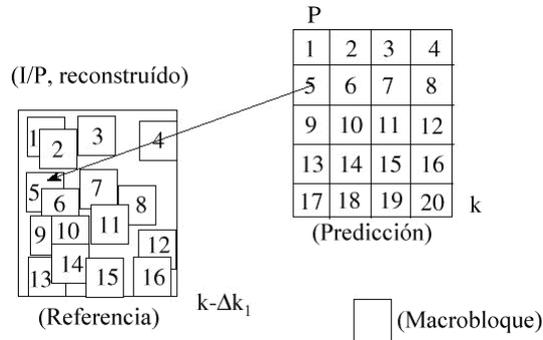
- Estimación de movimiento:
  - Se permite precisión de 1/2 pixel en los vectores de movimiento
  - La redundancia entre los VM de macrobloques vecinos se reduce codificación diferencial entre ellos (DPCM).
- Importante:
  - Aunque, la coincidencia de regiones es un método muy natural para reducir la redundancia temporal, MPEG no especifica el algoritmo de estimación de movimiento.
  - Cualquier algoritmo puede ser utilizado, dependiendo del tipo y requisitos de cada aplicación.

## Codificación *interframe*: cuadros P

- Predicción hacia delante y cuadros P:
  - El vector de desplazamiento del MB del cuadro actual  $k$  se estima en base al MB correspondiente del cuadro reconstruido anteriormente  $k-\Delta k_1$  (cuadro de referencia).
  - La diferencia (luminancia y color) entre el MB actual en el cuadro  $k$  y su predicción constituyen el MB de error de la predicción.
  - El vector desplazamiento y el término de error constituyen la información necesaria para reconstruir el MB actual del cuadro  $k$  a partir del MB correspondiente del cuadro  $k-\Delta k_1$ .
  - La DCT se aplica a los bloques de  $8 \times 8$  del término de error del MB, antes de su cuantificación y codificación (VLC).

## Codificación *interframe*: cuadros P (II)

- Predicción hacia delante y cuadros P:



## Codificación *interframe*: cuadros P (III)

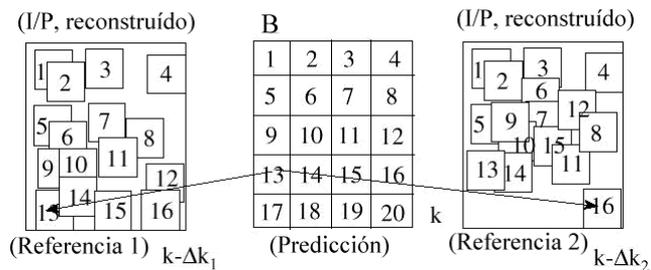
- Definición cuadros P:
  - Los cuadros P están compuestos por MB:
    - con codificación predictiva (*interframe*), o
    - con codificación *intraframe* (como los I).
  - Está permitido que el codificador decida que tipo de codificación utilizar dependiendo de la precisión de la predicción.
  - Esta decisión puede hacerse de distintas formas, y se deja a la elección del codificador.
- Ejemplo de criterio:
  - Comparar la variación de la luminancia del MB original con la del término de error.
  - Si la variación del término de error es mayor, entonces el MB se utilizaría codificación intra.

## Codificación *interframe*: cuadros B (I)

- Predicción bidireccional y cuadros P:
  - Para cada MB del cuadro actual  $k$ , se calculan dos vectores de desplazamiento (uno con la referencia en el pasado y otro en el futuro).
  - La predicción bidireccional está especialmente adaptada para descubrir regiones de predicción y consigue elevados niveles de compresión.

## Codificación *interframe*: cuadros B (II)

- Predicción bidireccional y cuadros B:



## Codificación *interframe*: cuadros B (III)

- Predicción bidireccional:

$$\text{pred} = \text{NINT} [\alpha_1 \cdot \text{pred\_forward} + \alpha_2 \cdot \text{pred\_backward}]$$

Tipos:

- $\alpha_1 = 0.5$  y  $\alpha_2 = 0.5 \Rightarrow$  *bi-directional prediction*
- $\alpha_1 = 1$  y  $\alpha_2 = 0 \Rightarrow$  *forward prediction*
- $\alpha_1 = 0$  y  $\alpha_2 = 1 \Rightarrow$  *backward prediction*

## Codificación *interframe*: cuadros B (IV)

- Definición cuadros B:

- Los cuadros B están compuestos por MB:
  - con codificación predictiva *bidirectional*, o
  - con codificación predictiva *backward*, o
  - con codificación predictiva *forward*, o
  - con codificación *intraframe* (como los I)
- Un posible mecanismo para decidir es seleccionar el tipo de codificación que tenga menor variación (luminancia).
- Los MB en cuadros B no se usan como referencia.

## Codificación *interframe*: cuadros B (V)

- Cuantificación y codificación de los coeficientes de la DCT de los MB del término de error.
  - Hay diferencias respecto a los cuadros I y P:
    - Los coeficientes de continua (DC) se cuantifican igual que los de alterna (AC).
    - Se espera que después del proceso de cuantificación aparezca una “zona muerta” muy próxima a 0.
    - La matriz de cuantificación por defecto es:

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

## Compensación del Movimiento

- Dos técnicas:
  - *Predicción*.- la imagen actual puede ser modelada como una transformación de una imagen previa.
  - *Interpolación*.- técnica de multi-resolución. Se submuestra la imagen y la imagen de plena resolución se obtiene por interpolación entre la señal de baja resolución y un término de error.  
La señal a reconstruir por interpolación se obtiene añadiendo una señal de error a una combinación de referencias pasadas y futuras.

## Representación del Movimiento

- Se consideran Macro Bloques de 16x16 pixels (Motion-compensation units), que pueden ser I, B ó P.
- Hay un compromiso entre la ganancia que obtenemos por la información de movimiento y el coste de codificar el macrobloque.
- Hay uno o dos vectores de movimiento por cada macrobloque dependiendo del tipo de bloque.
- Se minimiza una función de coste con objeto de encontrar el vector de movimiento óptimo.

## MPEG/Audio

- Estándar internacional para la compresión de audio de alta calidad.
- MPEG-1: 1.5 Mbps (1.2 Mbps video, 0.3 Mbps audio aprox.)
- La frecuencia de muestreo puede ser: 32, 44.1 ó 48 KHz.
- Uno ó dos canales de audio en cuatro posibles modos (mono, dual, estéreo y estéreo conjunto)
- Rates: 32 a 224 Kbps por canal
- Factor de compresión: 2.7:1 a 24:1



## Aplicaciones MPEG/Audio

---

- Consumer Recording (DCC)
- Disc based storage (CD-i, CD-Video)
- DVD
- Disc based Editing, audio broadcasting station automation
- Solid State Storage Audio
- Cable and satellite TV (e.g. DVB, USSB, DirecTV, EchoStar)
- Cable Radio
- Digital Audio Broadcasting (e.g. ADR, DAB, US-Digital Radio, Worldspace Radio)
- Internet Radio
- Computer based Multimedia
- Contribution & Distribution Links
- ISDN Links

## Psico-Acústica

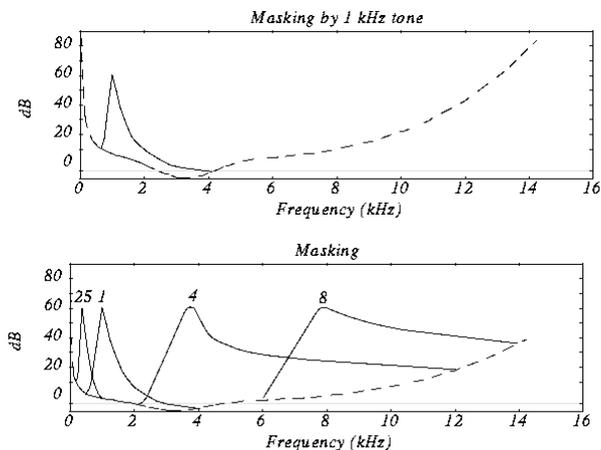
---

- La compresión MPEG/Audio se consigue en gran parte desechando partes irrelevantes de la señal de audio.
- El oído posee una resolución y un rango dinámico más amplio, si bien es más lento que el ojo.
- El rango va de 20 Hz a 20 KHz, siendo más sensible entre 2 y 4 KHz.
- Rango dinámico de 96 dB
- La voz humana va de 500 Hz a 2 KHz.

## Modelo Psico-Acústico

- El sistema auditivo humano no puede percibir ruido de cuantificación ante la presencia de enmascaramiento auditivo.
- Una señal de intensidad más alta puede enmascarar señales de menor intensidad posteriores.
- Las bandas críticas son los rangos de frecuencia en los que el oído posee las mismas propiedades perceptivas.
- El modelo psico-acústico analiza la señal de audio y calcula la relación ruido-enmascaramiento para cada una de las bandas de frecuencia.
- Se ofrecen 2 modelos psico-acústico, el Modelo 1 y el Modelo 2, que difieren en la forma en que se calculan los umbrales de enmascaramiento.

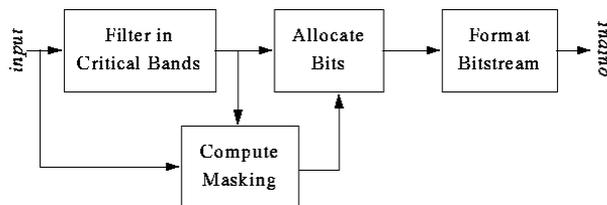
## Enmascaramiento de Señales



## Capas de Compresión

- **Tres capas de compresión independientes:**
  - Capa I, la más sencilla, para flujos por encima de 128 Kbps por canal (e.g. Philips DCC, 192 Kbps)
  - Capa II, de complejidad intermedia, flujos sobre 128 Kbps (e.g. Difusión de audio digital)
  - Capa III, la de mayor complejidad, ofrece la mejor calidad a 64 Kbps (e.g. Transmisión por línea RDSI). Coloquialmente conocido por MP3.

## Esquema codificador MPEG/Audio



## Algoritmo de Compresión MPEG/Audio

- Divido la señal de audio en 32 bandas críticas.
- Determino el umbral de enmascaramiento para cada una de las bandas vecinas a una dada.
- Si la potencia de la señal en una banda es inferior al umbral no codifico esa banda.
- Si es superior, determino el número de bits que necesito para cuantificar de modo que el ruido de cuantificación quede debajo del umbral de enmascaramiento (1 bit mejora en 6 dB la SNR)
- Empaquetó el stream de datos.

## Banco de Filtros Polifase

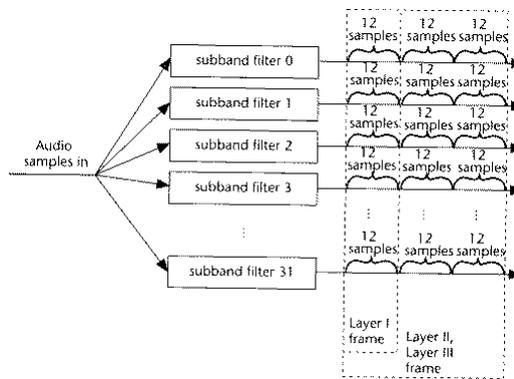


Figure: Grouping of Subband Samples for Layer 1, 2, and 3

## Parametrización codificador MPEG/Audio por capas

- **Capa 1:** Codificador en Sub-bandas y modelo psico-acústico. Filtro DCT con un frame e igual ancho de frecuencia para cada banda (385 muestras). El modelo psico-acústico sólo utiliza enmascaramiento de frecuencias.
- **Capa 2:** Mejora la resolución utilizando 3 frames en el filtro por sub-banda (anterior, actual y posterior, 1152 muestras).
- **Capa 3:** Se utiliza un filtro polifásico con bandas de diferente anchura espectral. El modelo psico-acústico incluye enmascaramiento temporal, cuantificación no uniforme, y codificación Huffman.

## Capas MPEG/Audio (II)

sound quality	bandwidth	mode	bitrate	reduction ratio
telephone sound	2.5 kHz	mono	8 kbps *	96:1
better than short-wave	4.5 kHz	mono	16 kbps	48:1
better than AM radio	7.5 kHz	mono	32 kbps	24:1
similar to FM radio	11 kHz	stereo	56..64 kbps	26..24:1
near-CD	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112..128kbps	14..12:1

# MPEG

---

## MPEG: familia de estándares de compresión de vídeo

### MPEG-2

# MPEG-2

---

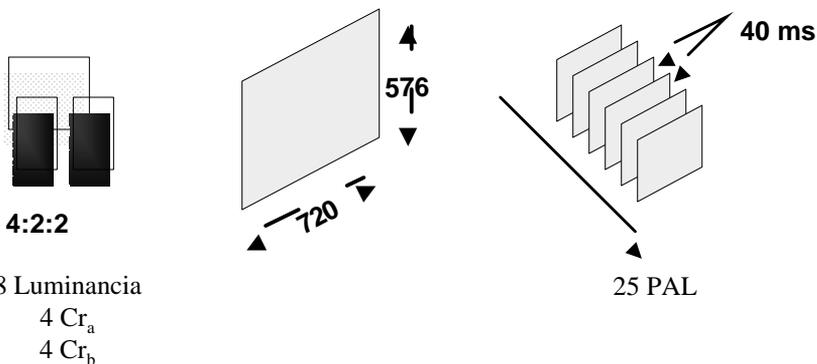
- MPEG 2: Estándar de compresión y codificación de vídeo y audio.
- Incluye:
  - MPEG-2 Video
  - MPEG-2 Audio
  - MPEG-2 Sistema **Video+Audio+Datos**
- Diseñado para soportar compresión y transmisión de video de mayor calidad que MPEG-1.
- **Objetivo básico:** difusión de vídeo digital de calidad TV, con bitrates de 4 a 9 Mbits/seg.
- Soporta vídeo entrelazado => permite manejar vídeo de alta calidad (TV convencional, HDTV, ...)

## Aplicaciones de MPEG-2

- Distribución de contenidos de vídeo de calidad:
  - Televisión digital:
    - Satélite (DBS, DVB-S, DSS)
    - Terrenal (DVB-T)
    - Cable (CATV, DVB-C)
  - HDTV (
  - DVD (Digital Versatile Disc)

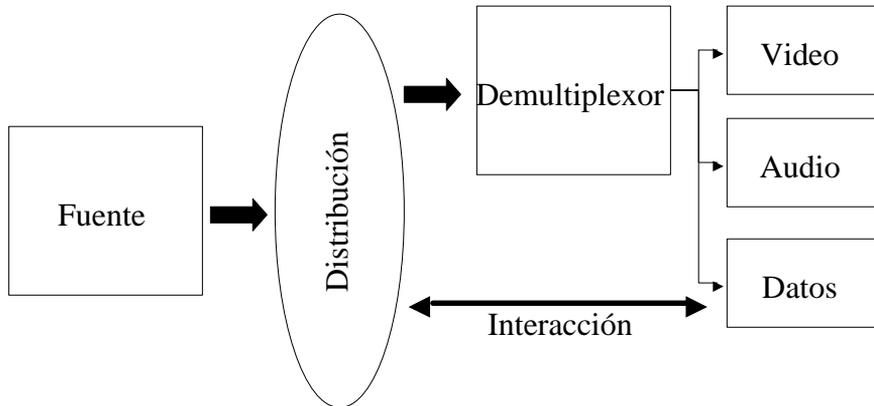
DVB: Digital Video Broadcasting  
DBS: Direct Broadcast Satellite

## Estándar de vídeo digital sin comprimir: ITU-R BT.601



$$16 \text{ bits} \quad \times \quad 720 \times 576 \text{ pixels} \quad \times \quad 25 \text{ fps} \quad = \quad 165,888 \text{ Mbps}$$

## Modelo de Referencia MPEG-2



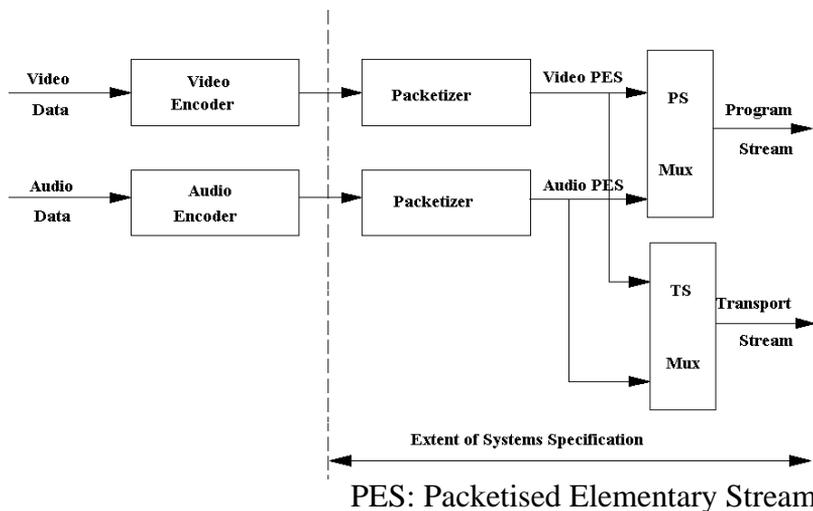
## MPEG-2

- Estándar para la codificación de video, audio y datos.
- Dividido en secciones
  - Parte 1 – Sistemas
  - Parte 2 – Vídeo
  - Parte 3 – Audio
  - Parte 4 – Pruebas de conformidad
  - Parte 5 – Software de simulación
  - Parte 6 – Extensiones de sistema: DSM-CC
  - Parte 7 – Extensión de audio: modo NBC
  - Parte 8 – Anulada
  - Parte 9 – Extensión de sistema: RTI
  - Parte 10 – Extensión de conformidad: DSM-CC

## Parte 1: MPEG-2 Sistemas

- Define la combinación de uno o más streams elementales de audio, video y datos, en uno o varios streams válidos para el almacenamiento o transmisión.
  - Dos tipos de streams definidos:
    - Program stream
      - Un solo “canal”. Ejemplo: DVD
    - Transport Stream
      - Varios “canales”. Ejemplo: distribución de canales de TV por satélite
- Cada uno está optimizado para un conjunto de aplicaciones

## Parte 1: Modelo MPEG-2 Sistemas



## Parte 1: MPEG-2 Sistemas. Tipos de stream

- Program Stream:
  - Parecido al múltiplex de sistema MPEG-1
  - Combinación de uno o varios PES con un reloj común en un solo stream
  - Tamaño de paquete variable y, normalmente, grandes
  - Diseñado para:
    - Entornos de transmisión sin errores
    - Válido para aplicaciones software
- Transport Stream:
  - Combinación de uno o varios PES con uno o más relojes independientes en un solo stream
  - TS puede contener varios programas
  - Tamaño de paquete fijo de 188 bytes
  - Diseñado para:
    - Entornos de transmisión con errores

## Parte 2: MPEG-2 video

- Se construye en base a las potentes capacidades de compresión del MPEG-1 para ofrecer un amplio repertorio de herramientas de codificación.
- Las herramientas se agrupan en perfiles para ofrecer diferentes funcionalidades.
  - Perfiles: conjunto de funcionalidad
    - Definen la sintaxis => algoritmos
  - Niveles: restricciones numéricas a la complejidad
    - Resolución
    - Muestreo (luminancia, crominancia)
    - Bitrate

## Parte 2: MPEG-2 video perfiles y niveles

Perfil \ Nivel	Simple	Main	Escalable (SNR)	Escalable (espacial)	High	Multivista	4:2:2
High		■			■		
High-1440		■		■	■		
Main	■	■	■		■	■	■
Low		■					

## Parte 2: MPEG-2 video perfiles

- Profiles (Perfiles)
  - Simple
    - Sólo frames I y P
    - Aplicaciones software y CATV
  - Main
    - Muy similar al MPEG-1
    - Más extendido (95%): satélite, CATV.
    - El más utilizado es: “Main Level-Main Profile” ó ML@MP
  - Escalable (SNR)
    - Perfil Main con escalabilidad en calidad
  - Escalable (espacial)
    - Perfil Main con escalabilidad espacial
  - High
    - Perfil Main, SNR, Espacial más otras opciones

## Parte 2: MPEG-2 video niveles

<i>Nivel</i>	<i>Resolución</i>	<i>fps</i> <i>(max)</i>	<i>Bitrate</i> <i>máximo</i>	<i>Aplicación</i>
Low	352x240	30	4 Mbps	SIF, cintas vídeo Equivalente a MPEG-1
Main	720x576	30	15 Mbps	CCIR 601, TV estudio. Resolución 4:2:0 "normal"
High 1440	1440x1152	60	60 Mbps	4x601, HDTV de usuario
High	1920x1080	60	80 Mbps	Producción alta calidad. Optimizado para la HDTV

## Parte 2: MPEG-2 video (I)

- Mejoras respecto a MPEG 1:
  - codificación de vídeo entrelazado
  - mayor precisión para cuantificar el coeficiente DC (8, 9, 10 ó 11 bits).
  - cuantificación no-lineal
  - tablas de Variable-Length-Coding
  - escalabilidad: representación de video a varias resoluciones.

## Parte 2: MPEG-2 video (II)

<i>Estándar</i>	<b>MPEG-1</b>	<b>MPEG-2</b>	
<i>Publicación</i>	1992	1994	(HDTV)
<i>Formatos de imagen</i>	CIF (1/4 TV) 352x288 4:2:0 Vídeo progresivo	TV 720x576 4:2:0 y 4:2:2 Vídeo progresivo y entrelazado	4 x TV 1440x1152 4:2:0 y 4:2:2 Vídeo progresivo y entrelazado
<i>Resolución temporal</i>	25-30 imágenes/s	50-60 cuadros/s	100-120 cuadros/s
<i>Bitrate</i>	1.5 Mbps	4-15 Mbps	18-60 Mbps
<i>Calidad</i>	Comparable VHS	Comparable NTSC/PAL	Cine digital

## Parte 2: MPEG-2 video. Frames

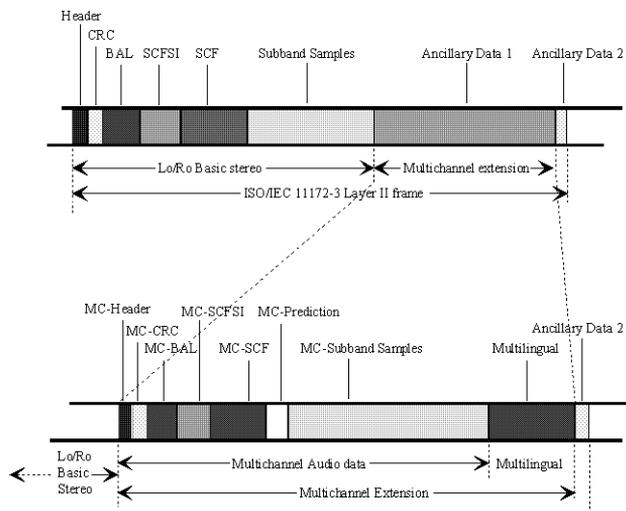
Tamaño típico de los frames codificados

	<b>I</b>	<b>P</b>	<b>B</b>	<b>Media</b>
<b>MPEG-1</b>				
1.15 Mb/s	150	50	20	38
<b>MPEG-2</b>				
4.00 Mb/s	400	200	80	130

## Parte 3: MPEG-2 Audio

- Compatible con MPEG-1 Audio.
- Extiende el rango de frecuencias de muestreo incluyendo 16 KHz, 22.05 KHz y 24 KHz, para anchos de banda por debajo de 64 Kbps.
- Se añaden 3 canales adicionales hasta un total de 5: izq., der., centro, y dos canales *surround*. Opcionalmente puede añadir hasta 7 canales multiidioma.

## Parte 3: MPEG-2 Audio. Stream audio



## Parte 3: MPEG-2 Audio. Códecs

- ISO 13818-3: compatible con MPEG-1 audio
  - Audio multicanal: 5.1
  - Audio multilingüe: 7 canales
  - Tasas binarias más bajas: hasta 8kbps
  - Frecuencias de muestreo menores: 16 KHz, 22.05 KHz y 24 KHz
  - Compatible en los dos sentidos
- ISO 13818-7 códec avanzado
  - AAC: Advanced Audio Coding
  - No es compatible con MPEG-1
- Dolby AC-3: códec alternativo
  - ATSC
  - DVD

## Parte 7: MPEG-2 AAC

- AAC: Advanced Audio Coding (ISO 13818-7)
  - Nuevos algoritmos de codificación
  - No compatible con MPEG-1 audio
  - Diseñado para la codificación de señales estereo y multicanal.
  - Basado en los mismos paradigmas que MPEG-1 Capa 3:
    - Banco de filtros de alta resolución en frecuencia
    - Cuantificación no uniforme
    - Codificación Huffman
    - Estructura de bucle interno de control y bit-rate
  - AAC consigue calidad similar a MP3 y con un 70% del bitrate

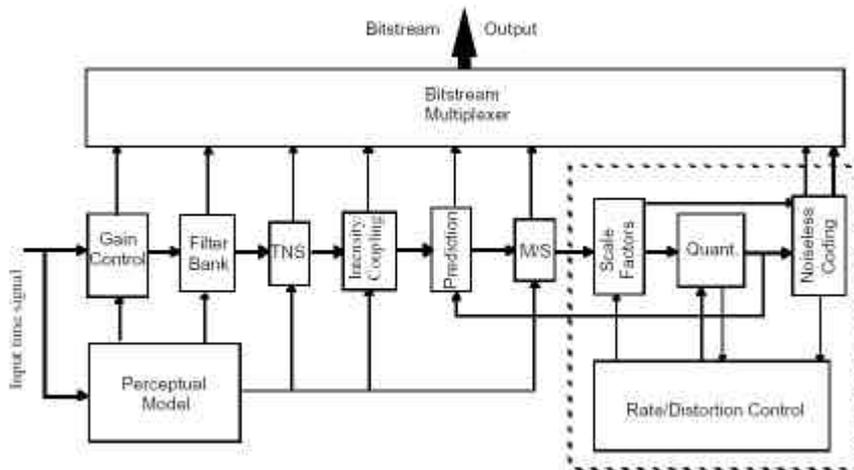
## Parte 7: MPEG-2 AAC (II)

- Cambios frente a MPEG-1 Capa 3. Eficiencia de codificación:
  - Mayor resolución frecuencial
    - 1024 líneas de frecuencia frente a 576 en MP3
  - Predicción
    - Predicción hacia atrás (opcional) calculada línea a línea
  - Mejora en la codificación joint stereo
    - Más flexible, permite aplicarla más a menudo para bajar bitrate
  - Mejora en la codificación Huffman
    - Se utiliza con mayor frecuencia la codificación por de 4 líneas de frecuencia.
    - La asignación de las tablas de códigos es más flexible

## Parte 7: MPEG-2 AAC (III)

- Cambios frente a MPEG-1 Capa 3. Mejorar en la calidad del audio:
  - Mejora del bloque de conmutación
    - Reemplaza el banco de filtros híbrido (cascada) de MP3
    - Utiliza banco de filtros conmutado MDCT (Modified Discrete Cosine Transform)
    - Respuesta al impulso 3 veces mejor que en MP3
  - TNS: Temporal Noise Shaping:
    - Conformado de ruido temporal en función de una predicción en el dominio de frecuencia
    - Es muy útil para mejorar la calidad de los diálogos a bajo bitrate

## Parte 7: MPEG-2 AAC (IV)



## Parte 6: DSM-CC

- ¿Qué es?
  - Digital Storage Media – Command & Control
  - ISO/IEC 13818-6
  - Situado entre la capa de transporte y la de aplicación
- Características
  - Independencia del medio de almacenamiento
  - Independencia del lugar de acceso (local o remoto)
  - Independencia de los protocolos de capas inferiores
  - Independencia del sistema operativo

## Parte 6: DSM-CC (II)

---

- Entorno
  - Multi-servidor
  - Multi-sesión
  - Multi-cliente
  - Conexión flexible: difusión (broadcast), punto a punto, multipunto a multipunto, multicast
- Aplicaciones
  - Vídeo Bajo de Demanda (VoD)
  - Servicios interactivos de vídeo
  - Servicios audio/vídeo en red

## Parte 6: DSM-CC (III)

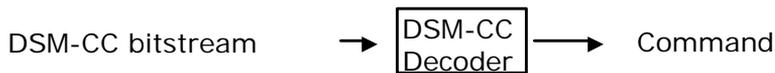
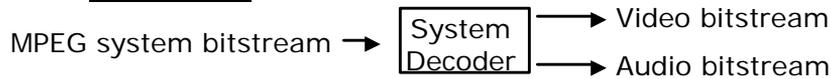
---

- Funciones básicas:
  - Acceso
    - Play / resume un stream MPEG
    - Determinar la velocidad de reproducción
    - Determinar el inicio de reproducción y de parada+
    - Determinar la dirección de reproducción
    - Pausa/Salto/Parada
  - Almacenamiento
    - Almacenar un intervalo de stream
    - Borrar

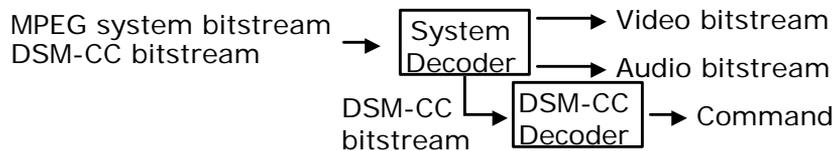
## Parte 6: DSM-CC (IV)

- **Transmisión:**

- **Individual**



- **Conjunta**



## Parte 6: DSM-CC (V)

- Modelo funcional de referencia

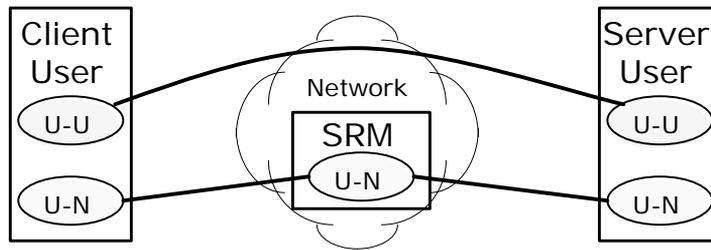
- Usuario

- Cliente: consumidor del contenido multimedia
      - STB
    - Servidor: proporciona el contenido multimedia y los servicios

- Red

- No homogénea
      - Ejemplo: servidor: red con backbone ATM
        - » Cliente: HFC o FTTC
    - SRM (Session and Resource Manager)
      - Información de configuración de usuarios
      - Autenticación de usuarios

## Parte 6: DSM-CC (VI)



- Modelo funcional de referencia (II)
  - Flujo de información entre usuarios: User-to-User
  - Flujo de información entre usuario y re: User-to-Network
    - Sesión/Conexión/Configuración

## MPEG

**MPEG: familia de estándares de compresión de vídeo**

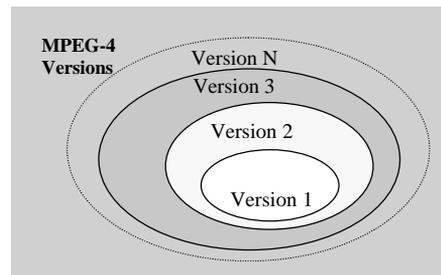
## MPEG-4

## MPEG-4

- **MPEG-4: Codificación de objetos audiovisuales**
- **Objetivos y funcionalidad:**
  - Representación de contenido visual, auditivo o audiovisual, mediante objetos (naturales o sintéticos)
  - Descripción de la composición de los objetos para crear escenas audiovisuales
  - Multiplexar y sincronizar los datos asociados con objetos multimedia, proporcionando calidad de servicio sobre canales de red
  - Interactuar con los objetos de la escena en el terminal

## MPEG-4

- Estándar MPEG-4: ISO/IEC 14496
- Estándar oficial en abril de 1999 (Versión 1)
- Sucesivas ampliaciones
  - Versión 2 en febrero 2000
  - Versiones adicionales



## Aplicaciones de MPEG-4 (I)

---

- Definido para aplicaciones en tres campos:
  - Televisión digital
  - Aplicaciones gráficas interactivas (contenido sintético)
  - Multimedia interactiva (WWW, distribución y acceso a contenido)
- Proporciona elementos tecnológicos estándar que permiten la integración de la producción, distribución y acceso a contenidos.

## Aplicaciones MPEG-4 (II)

---

- Aplicaciones multicast interactivas
  - IP Multicast sobre satélite
  - TV personalizada: seleccionando los objetos de interés (idioma, subtítulo, información sobreimpresionada, ...)
- Composiciones multimedia que permitan:
  - Mover texto y gráficos
  - Cambiar el punto de vista del usuario
  - Sincronizar audio, video, texto y gráficos
- Aplicaciones de trabajo colaborativo
  - 2D y 3D

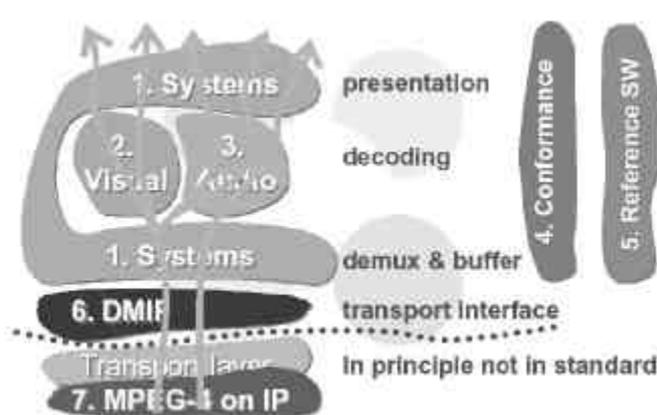
# Estándar MPEG-4

## Codificación de objetos audiovisuales

- Dividido en secciones:
  - Parte 1 – Sistemas
  - Parte 2 – Vídeo
  - Parte 3 – Audio
  - Parte 4 – Pruebas de conformidad
  - Parte 5 – Software de referencia
  - Parte 6 – DMIF (Delivery Multimedia Integration Framework)
  - Parte 7 – Software de codificación optimizado
  - Parte 8 – MPEG 4 sobre IP
  - Parte 9 – Hardware de referencia
  - Parte 10 – Codificador de vídeo avanzado

## Estándar MPEG-4 (II)

- Relación entre las partes del estándar:



Las flechas representan el flujo de bits entre los sistemas MPEG-4

## Características básicas MPEG-4

---

- La codificación de los objetos audiovisuales se basa en el concepto de escena.
- **Escena:** composición de objetos multimedia, organizados en forma de árbol.
- Las hojas del árbol son objetos multimedia primitivos.  
Ejemplos:
  - Imágenes fijas (fondo)
  - Objetos de vídeo (persona hablando)
  - Objetos de audio (voz asociada a una persona, música de fondo, ...)
- Los objetos estandarizados incluyen la representación de contenidos naturales y sintéticos en 2 y 3 dimensiones.

## Características básicas MPEG-4 (II)

---

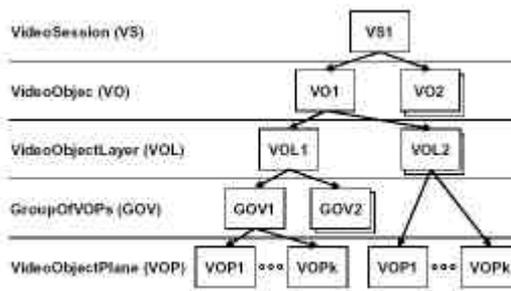
- Otros tipos de objetos definidos son:
  - Texto y gráficos
  - Bustos parlantes sintéticos con texto asociado para sintetizar el discurso y la animación. Cuerpos animados.
  - Sonido sintético
- Un objeto se codifica con elementos descriptivos (atributos) que permitan manejar el objeto en la escena de forma independiente a los elementos del contorno.

## Características básicas MPEG-4 (III)

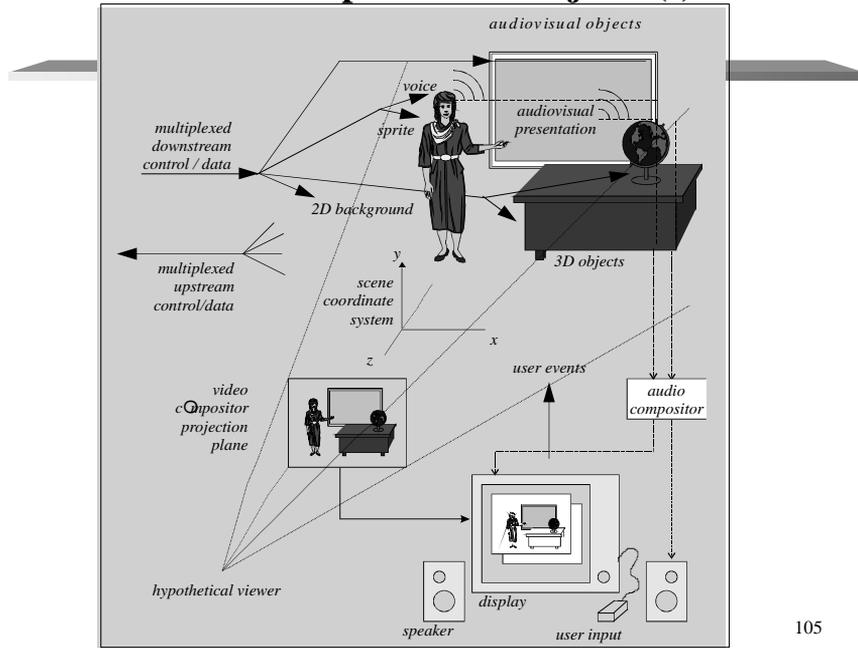
- Los elementos que describen las escenas permiten:
  - Situar objetos respecto a un sistema de coordenadas
  - Aplicar transformadas para cambiar la geometría o la acústica de un objeto
  - Agrupar objetos primitivos para formar objetos compuestos
  - Asociar streams de datos a los objetos para modificar sus atributos (sonido, parámetros de animación ...)
  - Cambiar, interactivamente, el punto de vista del usuario y el lugar de escucha dentro de la escena

## Características básicas MPEG-4 (IV)

- Jerarquía de clases

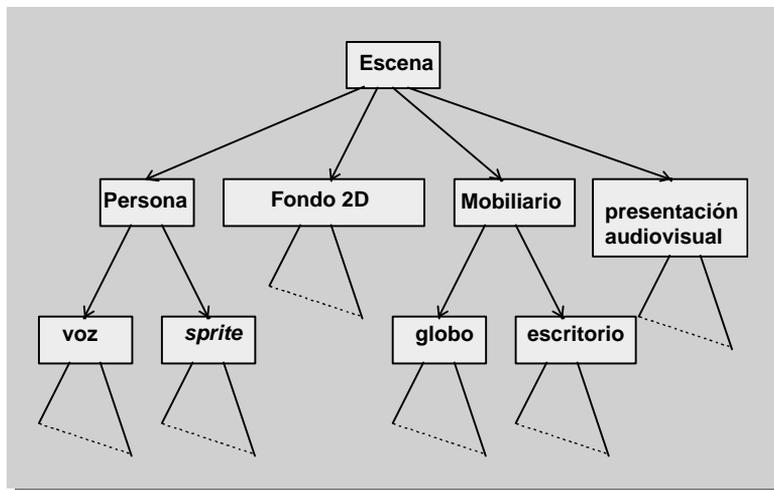


## Escena: composición de objetos (I)

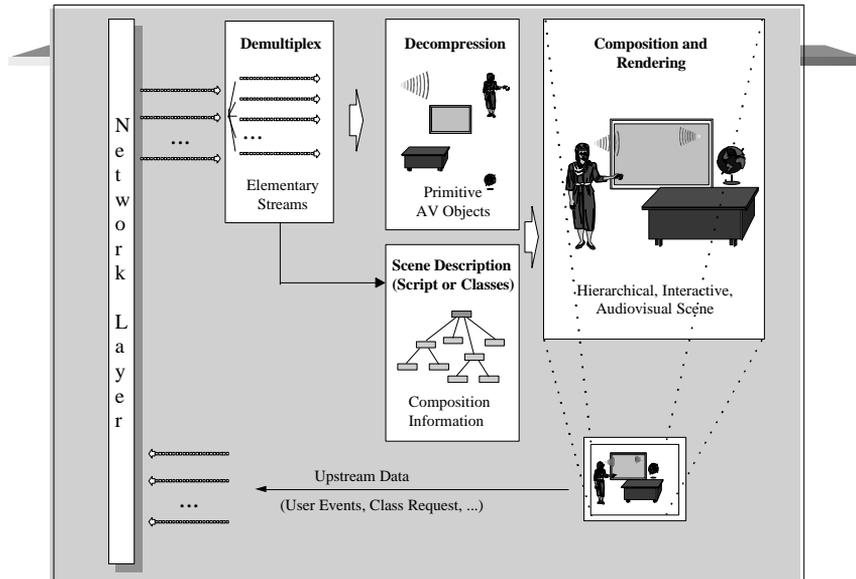


105

## Escena: composición de objetos (II)



## Modelo de terminal

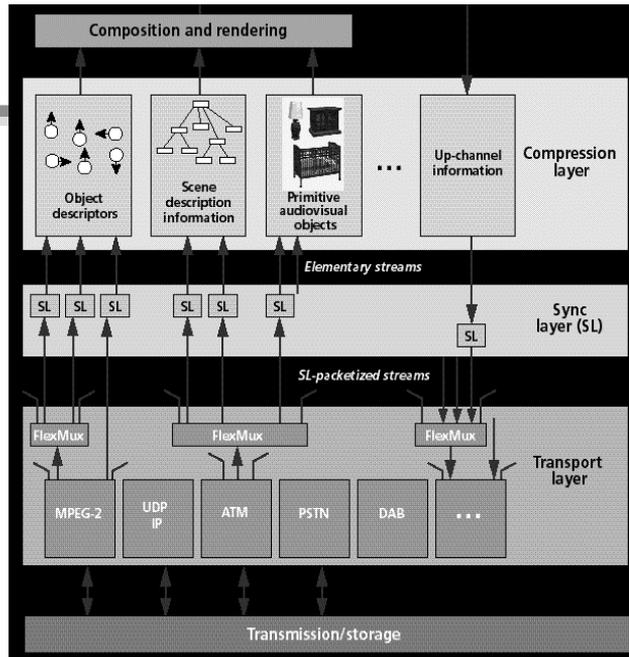


## Características básicas MPEG-4 (IV)

- Descripción y sincronización de los streams de datos de los objetos multimedia:
  - Descriptores de objeto:
    - Información de la configuración (parámetros necesarios en el receptor, precisión temporal de la codificación, ...)
    - Parámetros de calidad de servicio (QoS) de la transmisión:
      - Bitrate máximo
      - Tasa de error
      - Prioridad
  - Información para la gestión e identificación de los derechos de propiedad intelectual
  - Capa de sincronización
    - Marcas de tiempo (time stamping)
    - Independencia del tipo de medio

- **Estructura de capas de MPEG-4.**

- Distribución del stream de datos



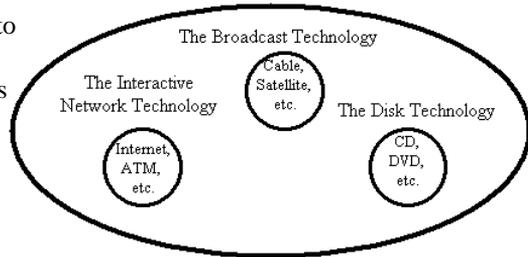
## Funcionalidades principales MPEG-4 (I)

- Transporte:
  - MPEG-4 no define capa de transporte
  - Se han definido adaptaciones a capas de transporte existentes:
    - Sobre MPEG-2 Transport Stream (extensión de MPEG-2 Sistema)
    - Sobre IP (en cooperación con la IETF)

## Funcionalidades principales MPEG-4 (II)

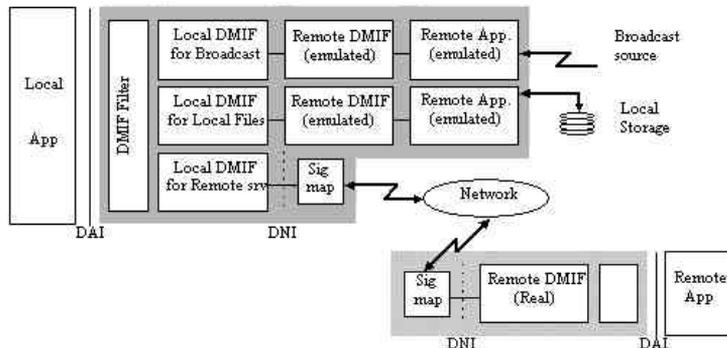
### Parte 6: DMIF (Delivery Multimedia Integration Framework)

- Interfaz entre aplicación y transporte. Independiza aplicación de transporte => una misma aplicación puede ejecutarse sobre diferentes capas de transporte
- Define:
  - Interfaz transparente para las aplicaciones independiente de la otra parte: parte interactiva, broadcast o almacenamiento.
  - Control del establecimiento de sesiones
  - Uso de redes heterogéneas entre las partes: IP, ATM, móvil, RTB, RDSI
  - Gestión de la capa de sincronización MPEG-4



## Funcionalidades principales MPEG-4 (III)

### • DMIF:

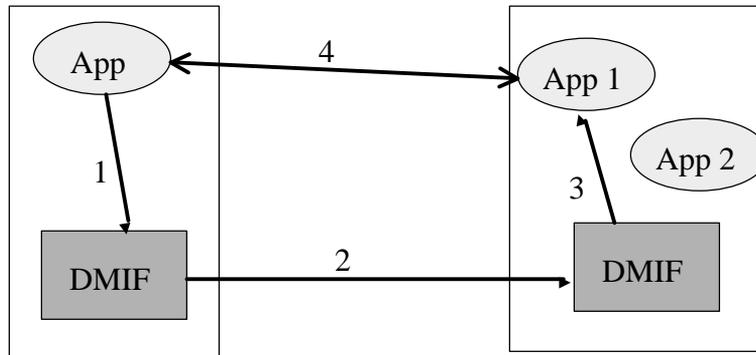


- Flows between independent systems, normative
- Flows internal to specific implementations, out of DMIF scope

DAI: DMIF Application Interface / DNI: DMIF Network Interface

## Funcionalidades principales MPEG-4 (IV)

- **DMIF:** Activación de servicio e inicio de intercambio de datos



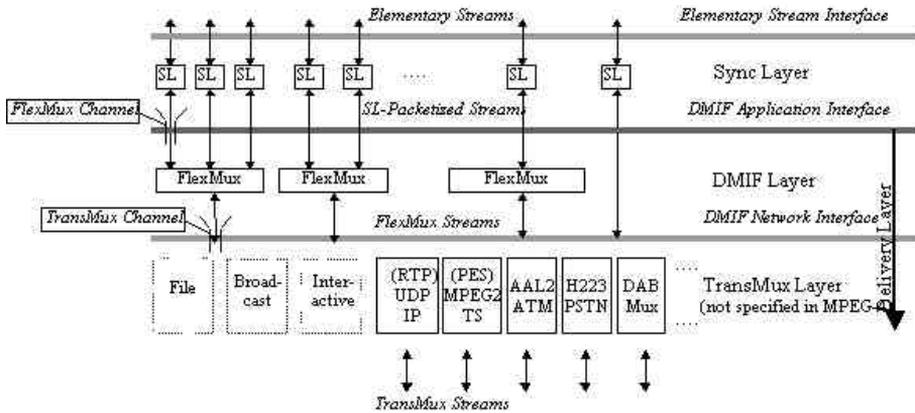
- 1.- Aplicación origen solicita activación del servicio a la capa DMIF local
- 2.- DMIF origen establece sesión con DMIF destino
- 3.- DMIF destino identifica a la aplicación destino y la reenvía la petición de activación
- 4.- La aplicación destino y la origen establecen el canal de datos

## Funcionalidades principales MPEG-4 (V)

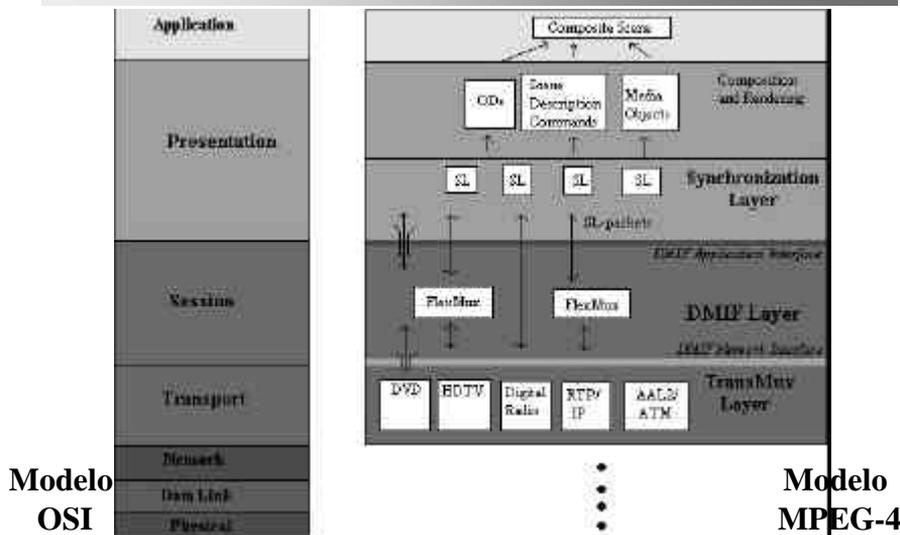
- **Parte 1: Sistemas**
  - Define la descripción de las relaciones entre los componentes audiovisuales de la escena. Incluye:
    - Sintaxis BIFS (Binary Format for Scenes): describe las relaciones espacio-temporales de los objetos en la escena.
    - Descriptores de objetos (Object Descriptors: OD): define las relaciones entre los streams elementales asociados a cada objeto.
    - Formato de fichero para el intercambio y creación de contenidos (mp4)
    - Interactividad: cliente/servidor, lanzamiento de eventos o seguimiento de las acciones del usuario, manejo de eventos entre objetos de la escena, ...
    - Multiplexación de streams, incluyendo información de tiempo (FlexMux)
    - MPEG-J: aplicaciones interactivas
    - Representación de texto multiidioma
    - Identificación de propiedad intelectual (IPR) de objetos multimedia

## Funcionalidades principales MPEG-4 (VI)

- **Sistemas:**

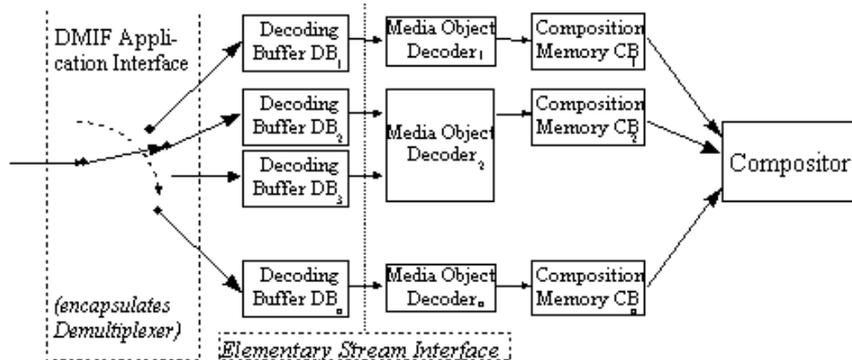


## Funcionalidades principales MPEG-4 (VII)



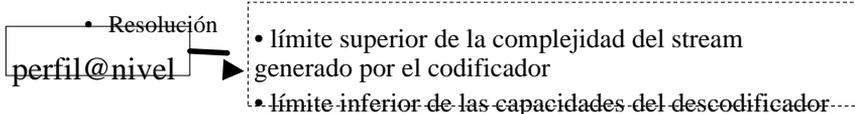
## Funcionalidades principales MPEG-4 (VIII)

- Sistemas: Modelo de Descodificador. Cada objeto de la escena se transmite de forma independiente.



## Perfiles/Niveles MPEG-4 (I)

- MPEG-4 formado por un gran número de herramientas (tools).
- Perfiles: agrupaciones de herramientas.
- Niveles: proporcionan una forma de limitar la complejidad computacional de los perfiles especificando parámetros de las herramientas.
  - Ejemplos:
    - Bitrate
    - Número máximo de objetos en la escena
    - frecuencia
    - Resolución

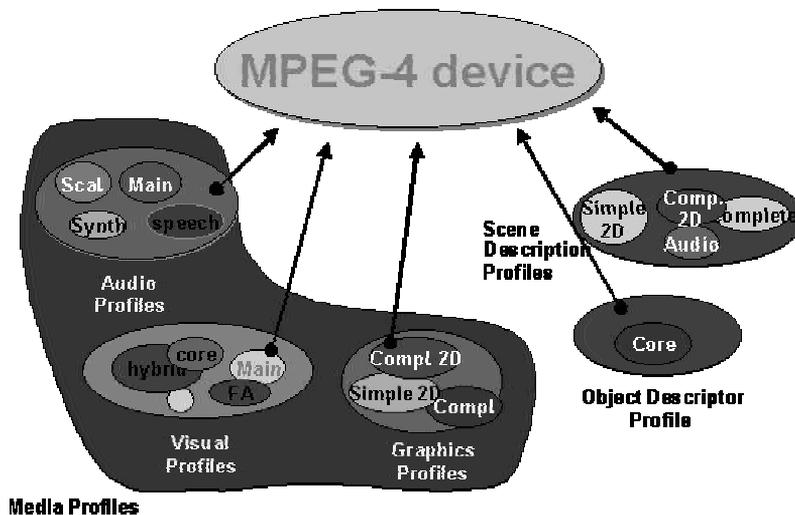


## Perfiles/Niveles MPEG-4 (II)

5 tipos de perfiles:

- Medios
  - **Audio**
  - **Vídeo**
  - **Gráficos** (nodos BIFS, elementos gráficos que se pueden utilizar en la escena)
- **Descripción de escena** (nodos BIFS, para construir la estructura de la escena y la relaciones entre objetos y acciones del usuario)
- **Descripción de objetos** (requisitos del terminal en término de descriptores de objetos y de la capa de sincronización)

## Perfiles/Niveles MPEG-4 (III)

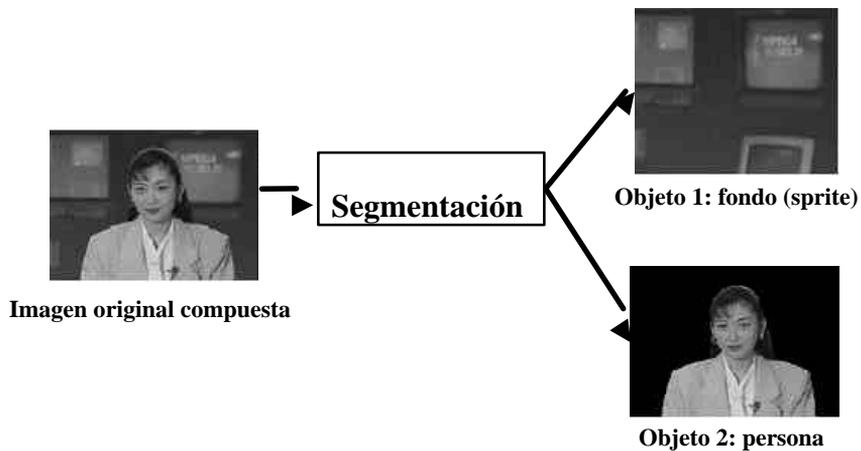


# MPEG-4

## MPEG-4: Vídeo y audio

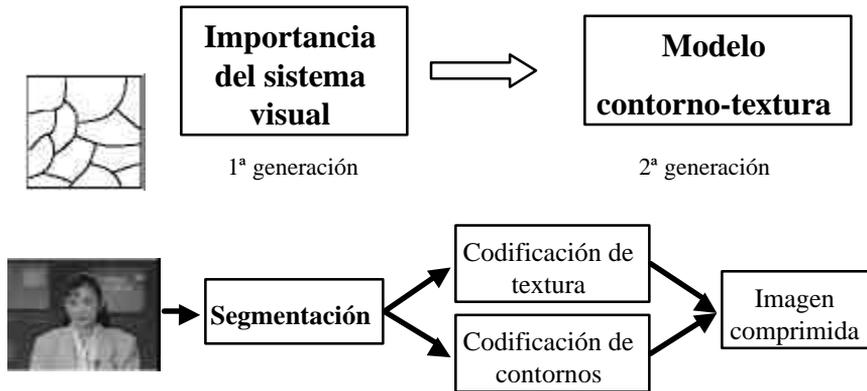
## MPEG-4: Vídeo. Bases (I)

### Segmentación de objetos



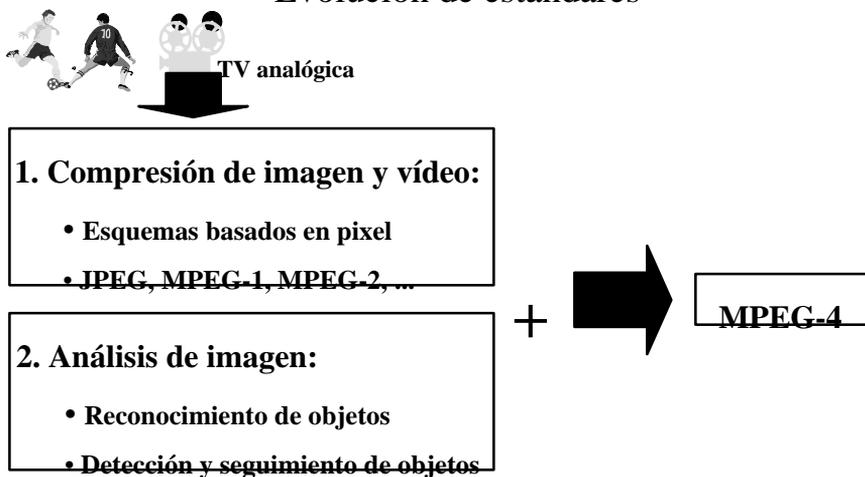
## MPEG-4: Vídeo. Bases (II)

Técnicas de nueva generación: codificación basada en segmentación



## MPEG-4: Vídeo. Bases (III)

Evolución de estándares



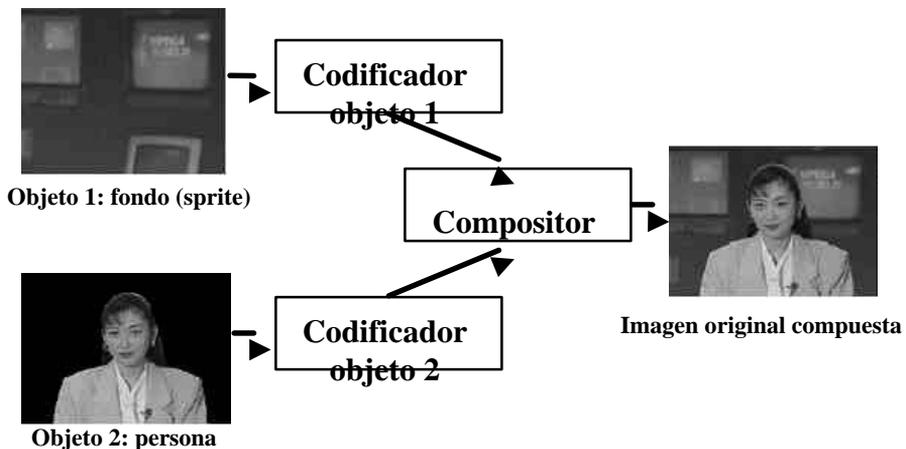
## MPEG-4: Vídeo. Bases (IV)

### Codificación de texturas

- Se utilizan técnicas clásicas:
  - DPCM
  - Métodos transformados (DCT, ...)
  - Cuantificación vectorial
  - Subbandas
  - Wavelets
- Fundamental:
  - Adaptación a la forma de la región u objeto
  - Independientes de la forma concreta

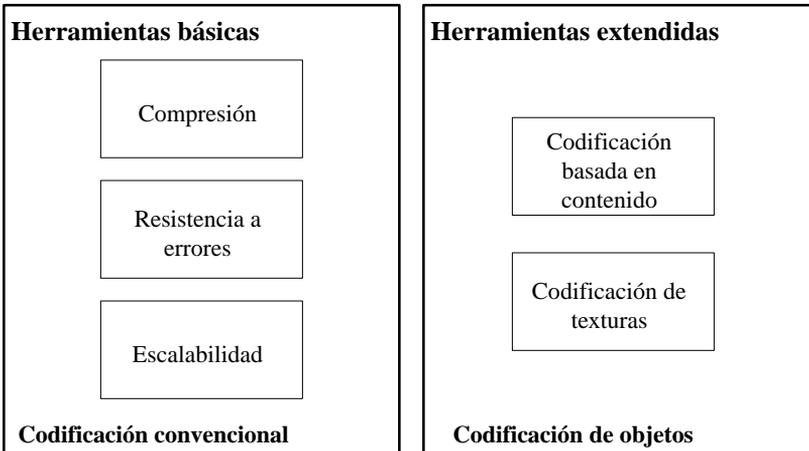
## MPEG-4: Vídeo. Bases (V)

### Codificación de objetos



## MPEG-4: Vídeo. Bases (VI)

### Herramientas de codificación

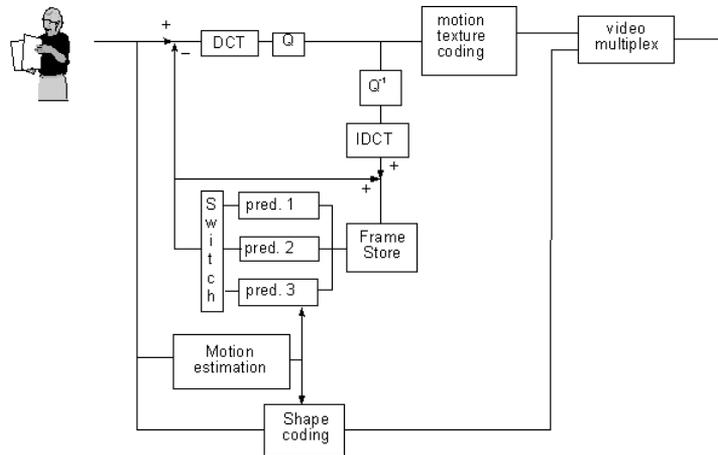


## MPEG-4: Vídeo. Bases (VII)

- **MPEG-4 vídeo** proporciona herramientas para un gran número de funcionalidades
- Aproximación integrada (base y extensiones)
- Basado en tecnología DCT (excepto para codificación de texturas)
- Compatible con la “baseline” de H.263
- Bastante compatible con MPEG-1
- Bastante compatible con MPEG-2

## MPEG-4: Vídeo. Bases (VIII)

### Modelo de codificador de vídeo



## MPEG-4: Vídeo. Funcionalidades (I)

### Básicas

- Eficiencia de codificación
  - 5 kbps <-> 50 Mbps
  - Resolución: pequeña-TV
  - Vídeo progresivo y entrelazado
- Resistencia a errores
  - Entornos móviles
- Escalabilidad
  - Espacial
  - Temporal

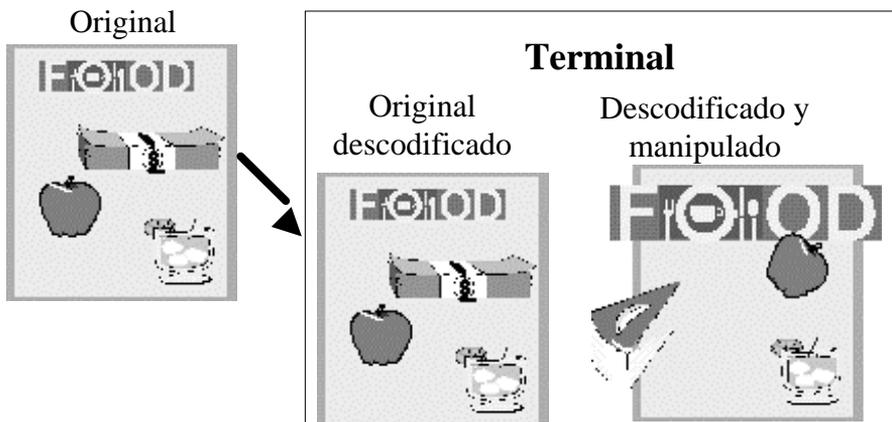
## MPEG-4: Vídeo. Funcionalidades (II)

### Extendidas: codificación basada en contenido

- Permite al usuario el acceso a objetos de formas arbitrarias codificados en la escena
- Permite alta interacción con el contenidos de la escena
- Manipulación del contenido de la escena sobre el stream

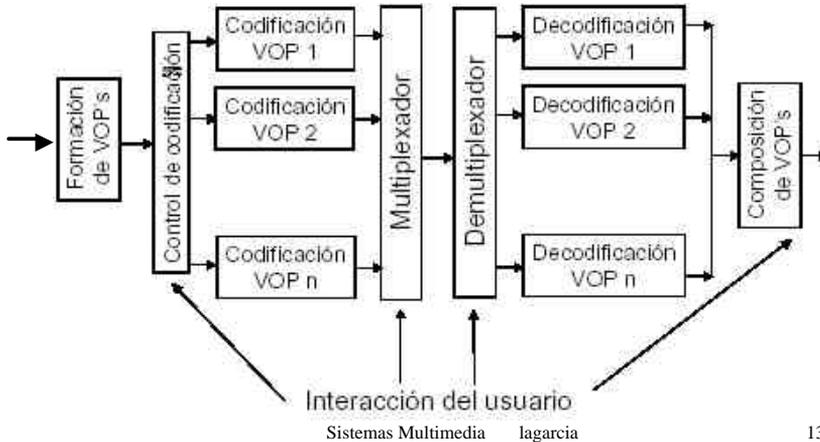
## MPEG-4: Vídeo. Funcionalidades (III)

### Manipulación de objetos



## MPEG-4: Vídeo. Funcionalidades (IV)

Capas de vídeo: cada objeto de vídeo en la escena se codifica y transmite de forma separada

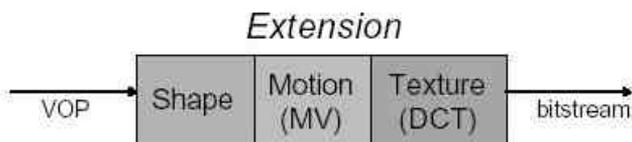
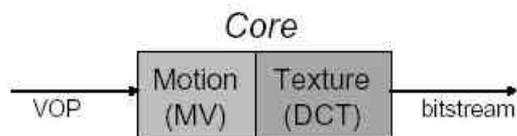


Sistemas Multimedia lagarcia

133

## MPEG-4: Vídeo. Funcionalidades (V)

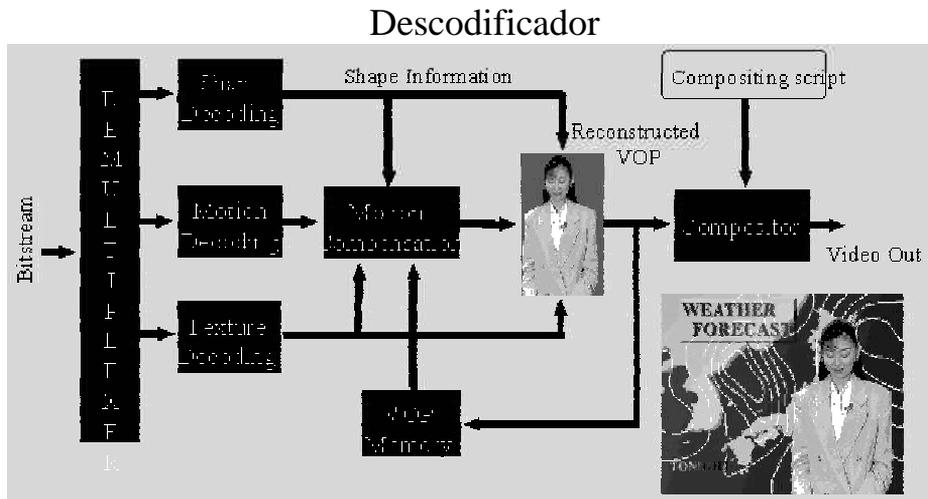
Codificación básica y extensiones



Sistemas Multimedia lagarcia

134

## MPEG-4: Vídeo. Funcionalidades (VI)



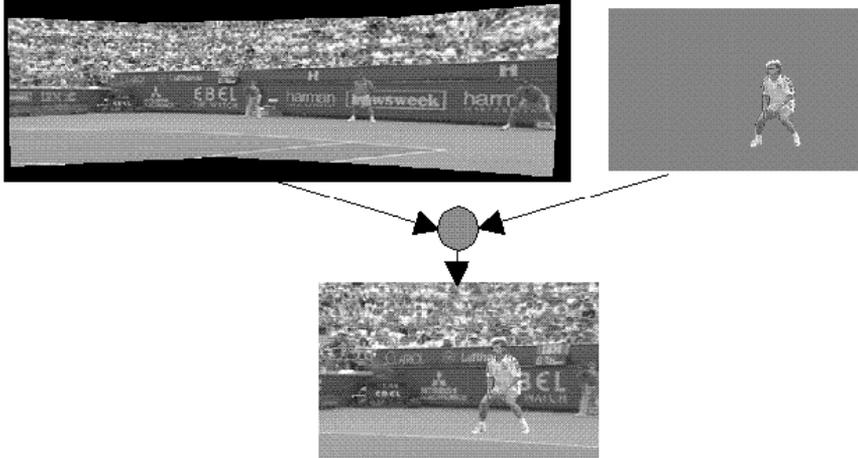
## MPEG-4: Vídeo. Funcionalidades (VII)

Funcionalidades dependientes del contenido:

- Codificación de formas
- Sprites
- Escalabilidad (basada en contenido)
- Codificación de texturas escalable (Wavelets)
- Resistencia a errores

## MPEG-4: Vídeo. Funcionalidades (VIII)

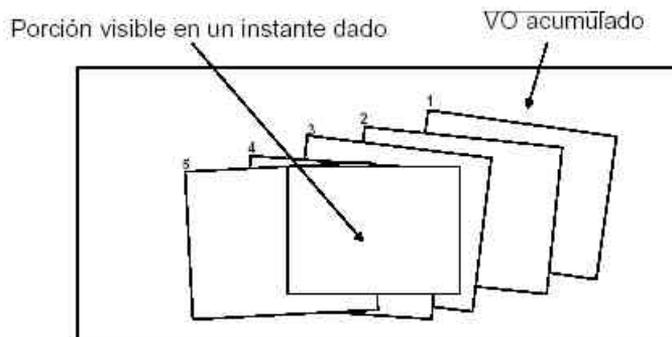
### Codificación de sprites



## MPEG-4: Vídeo. Funcionalidades (IX)

### Codificación de sprites

- Sprite: Objeto visual acumulado



## MPEG-4: Vídeo. Funcionalidades (X)

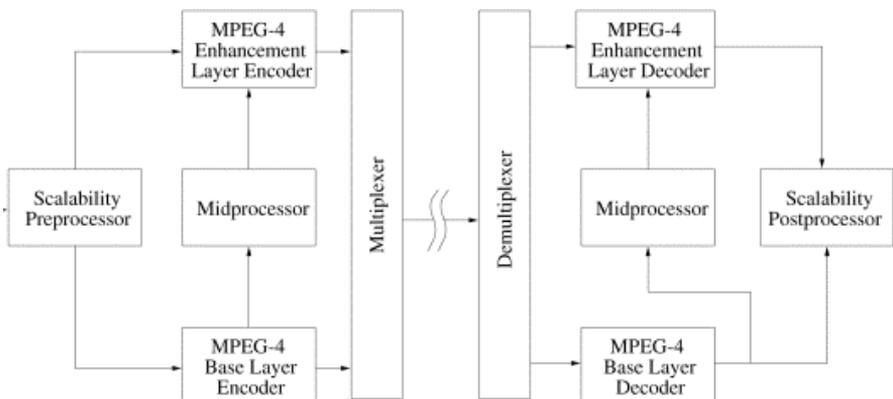
### Codificación de sprites

- La idea es análoga a la utilización de fondos en dibujos animados



## MPEG-4: Vídeo. Funcionalidades (XI)

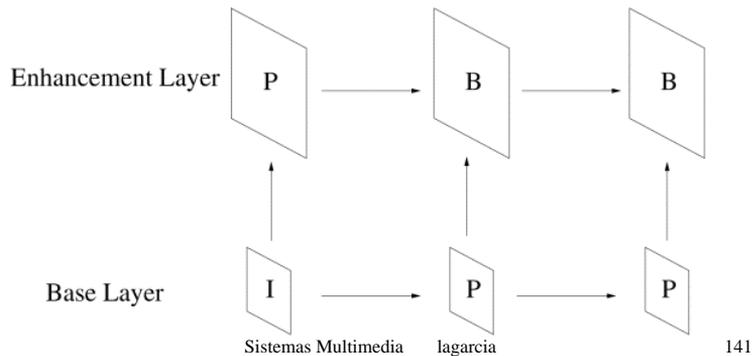
### Codificación escalable



## MPEG-4: Vídeo. Funcionalidades (XII)

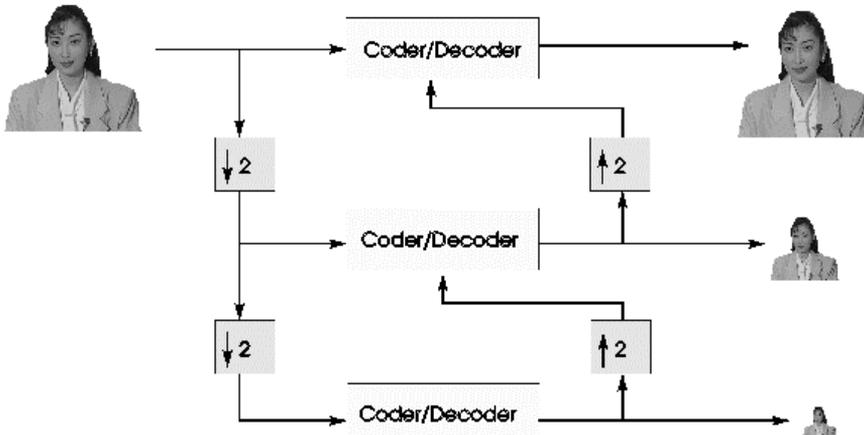
### Codificación escalable espacial (I)

- La capa base se codifica igual que si no fuera escalable
- VOPs de la capa ampliada se codifican como P o B
- Los VOP de la capa base son la referencia de la ampliada



## MPEG-4: Vídeo. Funcionalidades (XIII)

### Codificación escalable espacial (II)

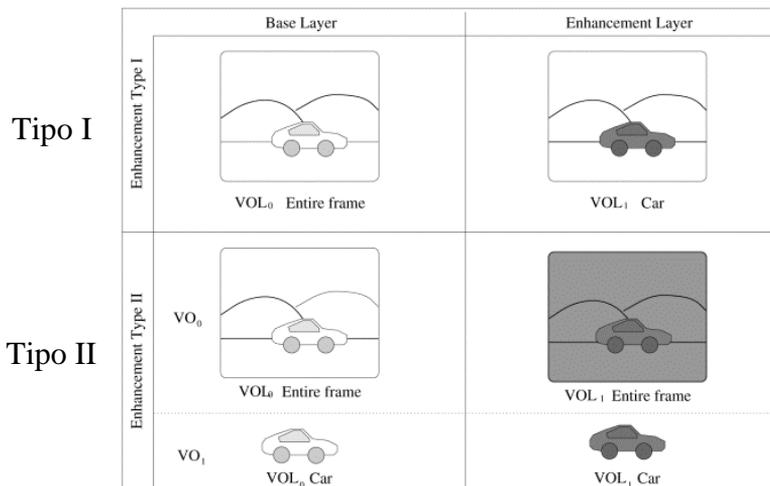


## MPEG-4: Vídeo. Funcionalidades (XIV)

### Codificación escalable temporal (I)

- La capa ampliada aumenta el frame rate de la capa base, llevando información a visualizar entre los frames de la capa base
- La capa ampliada puede generarse según dos tipos:
  - Tipo I: la ampliada amplía la resolución de sólo una parte de la capa base
  - Tipo II: la ampliada amplía la resolución de toda la capa base

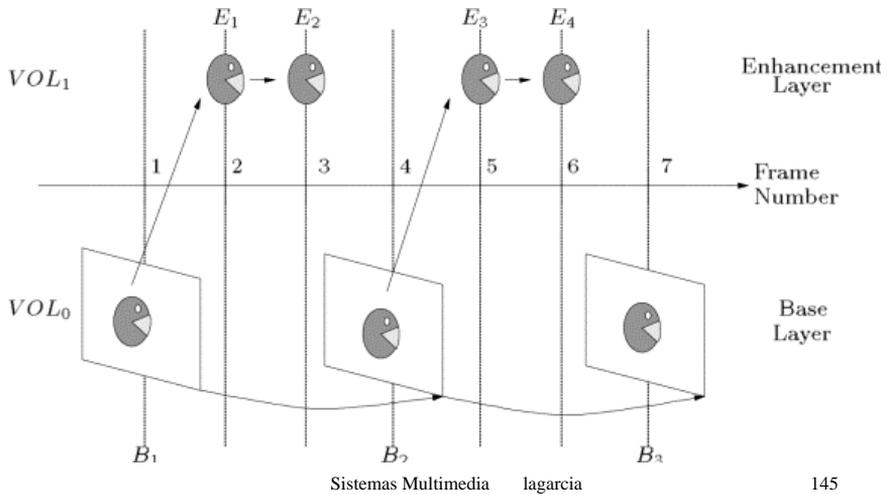
## MPEG-4: Vídeo. Funcionalidades (XV)



 Region to be enhanced by layer

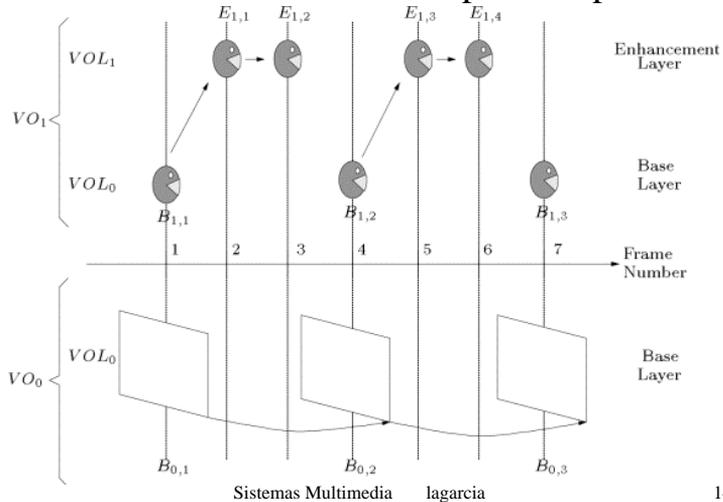
## MPEG-4: Vídeo. Funcionalidades (XVI)

### Codificación escalable temporal: Tipo I



## MPEG-4: Vídeo. Funcionalidades (XVII)

### Codificación escalable temporal: Tipo II



## MPEG-4: Vídeo. Funcionalidades (XVIII)

### Codificación escalable de texturas (**Wavelets**) (I)

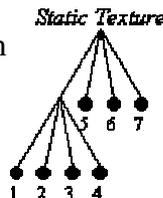
- La información de textura se separa en subbandas aplicando la DWT (Discrete Wavelet Transform)
- La DWT se aplica de forma recursiva sobre las subbandas generando un “árbol” de subbandas con la información de la imagen.

## MPEG-4: Vídeo. Funcionalidades (XIX)

### Codificación escalable de texturas (**Wavelets**) (II)

- Ejemplo de DWT con 2 niveles de descomposición:

- Textura original dividida en 4 subbandas
- La banda de menor  $f$  se subdivide en otras 4.



1	2	5
3	4	
6		7

- Aquí la subbanda 1 representa la banda de baja frecuencia: componente DC
- Las otras subbandas son las componentes AC
- Las componentes DC y AC se procesan diferente

## MPEG-4: Vídeo. Funcionalidades (XX)

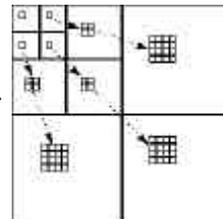
### Codificación escalable de texturas (**Wavelets**) (III)

- Codificación subbanda DC:
  - Los coeficientes de DC se codifican con un esquema predictivo
  - Cada coeficiente puede ser “predecido” en función de su vecino de la izquierda o el superior.
  - A continuación se cuantifican

## MPEG-4: Vídeo. Funcionalidades (XXI)

### Codificación escalable de texturas (**Wavelets**) (IV)

- Codificación subbandas AC:
  - Habitualmente los coeficientes de estas subbandas son muy próximos a 0 tras la cuantificación => la eficiencia de la codificación depende de la codificación del valor y la posición de los coeficientes no nulos.
  - La técnica de codificación explota la fuerte correlación entre los valores de la DWT a través de las escalas de una misma localización y orientación. Los coeficientes del padre están fuertemente correlacionados con los de los hijos.



## MPEG-4: Vídeo. Funcionalidades (XXII)

### Codificación escalable de texturas (**Wavelets**) (V)

- Codificación subbandas AC:
  - El algoritmo “zero-tree” explota esta relación entre los coeficientes: si un coeficiente es cero en una escala superior, es muy probable que los descendientes también sean cero => forman un árbol de ceros.
  - “zero-tree” existe en cualquier nodo cuyo coeficiente sea cero y todos sus hijos también lo sean.
  - Con este principio, los coeficientes se codifican aritméticamente, utilizando un símbolo para indicar si existe un “árbol de ceros” y el valor de los coeficientes.

## MPEG-4: Vídeo. Funcionalidades (XXIII)

### Escalabilidad wavelet: SNR



## MPEG-4: Vídeo. Funcionalidades (XXIV)

### Escalabilidad wavelet: Espacial



## MPEG-4: Vídeo. Funcionalidades (XXV)

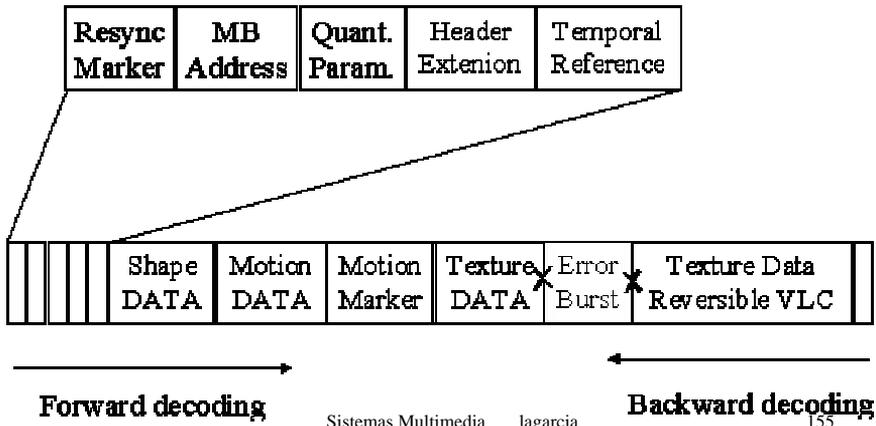
### Resistencia a errores

- Funcionalidad muy importante para el acceso universal en entornos propensos a errores, como las comunicaciones móviles.
- Herramientas MPEG-4:
  - **Resincronización:** inserción de marcas en el stream que permitan al receptor saltar a la siguiente marca ante la aparición de errores y reanudar la decodificación. Permite incluir marcas tras un número constante de bits codificados (MPEG-2 y H.263 lo permiten tras un número constante de macrobloques)
  - **Partición de datos:** separa los bits de codificación de movimiento de los de información de textura.
  - **Código de extensión de cabecera:** códigos binarios que permiten la inclusión de redundancia en la cabecera, fundamental para la correcta decodificación de vídeo.
  - **VLC's reversibles (RVLC):** para reducir la influencia de errores en los datos decodificados. Los RVLC son códigos que pueden decodificarse en las dos direcciones (atrás y adelante)

## MPEG-4: Vídeo. Funcionalidades (XXVI)

### Resistencia a errores

- Resumen de la influencia de las herramientas de resistencia a errores en la sintaxis del stream MPEG-4



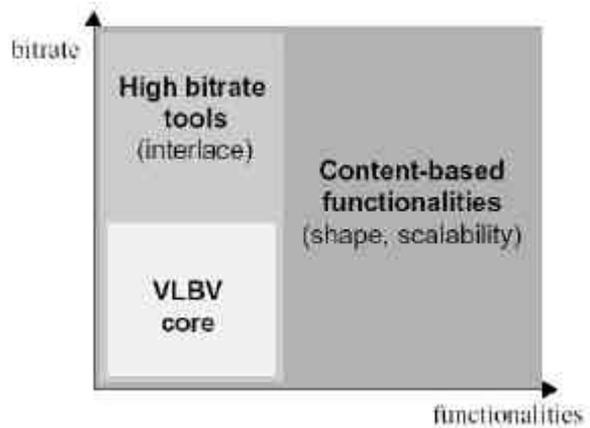
## MPEG-4: Vídeo. Funcionalidades (XXVII)

### Resumen:

- Permite codificación híbrida de imágenes naturales y video con escenas sintéticas.
  - Ejemplo: presencia virtual de los participantes en una videoconferencia.
- Funcionalidades:
  - Amplio espectro de tasas binarias (desde 5 kbps hasta Gbps)
  - Soporte de vídeo progresivo y entrelazado
  - Variedad de resoluciones, desde pequeñas ventanas (sub-QCIF) hasta cine digital (4k x 4k)
  - Codificación de texturas (imagen sintética)
  - Acceso aleatorio al vídeo (pausa, FF, RW, ...)
  - Codificación escalable, con granularidad fina
  - Codificación de formas y planos alfa
  - Protección frente a errores
  - Animación facial y corporal (define parámetros, no modelos)
  - Codificación de mallas poligonales (2D y 3D)

## MPEG-4: Vídeo

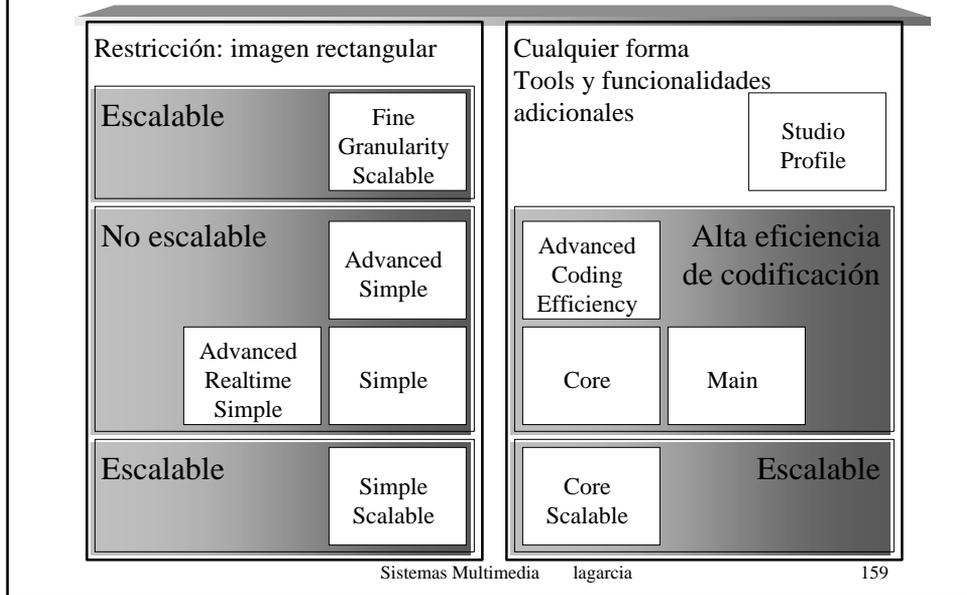
- Herramientas MPEG-4 Vídeo Natural



## Perfiles y niveles MPEG-4: Vídeo (I)

- Perfiles de vídeo natural:
  - **Simple:** vídeo natural rectangular de aspecto arbitrario (ancho/alto). Bajo bitrate y resistente a errores. Sólo planos I/P (plano  $\leftrightarrow$  frame VOP  $\leftrightarrow$  GOP) (VOP: Video Object Planes)
  - **Simple Scalable:** versión escalable del Simple, con escalabilidad temporal y espacial utilizando el Simple como capa base.
  - **Core:** Aumento de calidad respecto al Simple al utilizar interpolación bidireccional (B-VOP) y soportar objetos de forma arbitraria.
  - **Main:** objeto de vídeo de mayor calidad. Soporta formas en escala de grises, sprites y contenido entrelazado.
  - **N-bit:** igual que el core pero la precisión de la codificación del pixel puede variar de 4 a 12 bits para luminancia y crominancia.

## Perfiles y niveles MPEG-4: Vídeo (II)



## Perfiles y niveles MPEG-4: Vídeo (III)

### Perfil visual Simple

Simple

- Base de todos los perfiles de vídeo
- Procede de la especificación H.263
- Sólo soporta objetos de vídeo de forma rectangular (frames)
- Precisión de  $\frac{1}{2}$  pixel para compensación de movimiento
- Sólo utiliza movimiento de bloques
- Frames tipo I y P
- Implementación pequeña para dispositivos limitados

## Perfiles y niveles MPEG-4: Vídeo (IV)

Perfil visual Advanced Simple 

Advanced Simple
--------------------

 = 

Simple
--------

 + 

Coding Tools
-----------------

- Sólo soporta objetos de vídeo de forma rectangular (frames)
- Precisión de 1/4 pixel para compensación de movimiento
- Soporta frames B
- Soporta vídeo entrelazado (TV)

## Perfiles y niveles MPEG-4: Vídeo (V)

Perfil visual Core 

Core
------

 = 

Simple
--------

 + 

Shape Coding Tools
--------------------------

- Orientado a objetos: codificación de objetos de vídeo de formas arbitrarias
- Formas definidas por máscaras binarias
- Soporta frames B

## Perfiles y niveles MPEG-4: Vídeo (VI)

Perfil visual Advanced  
Coding Efficiency

Advanced  
Coding  
Efficiency

= Advanced  
Simple

+ Shape  
Coding  
Tools

- Orientado a objetos: codificación de objetos de vídeo de formas arbitrarias
- Formas definidas por máscaras en escala de grises
- Codificación de texturas “Shape adaptive DCT”

## Perfiles y niveles MPEG-4: Vídeo (VII)

Perfil visual Main

Main

= Core

+ Sprite  
Coding

- Orientado a objetos: codificación de objetos de vídeo de formas arbitrarias
- Formas definidas por máscaras en escala de grises
- Soporte adicional para la codificación de “sprites”

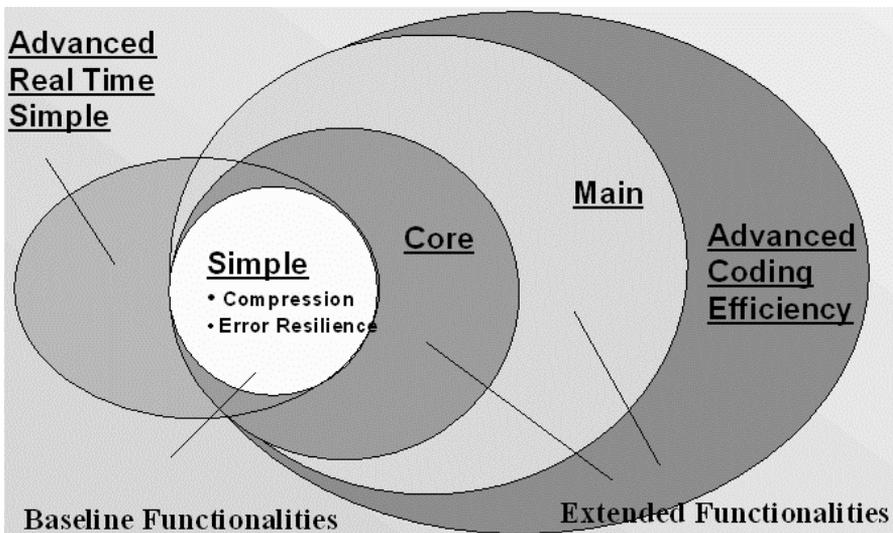
## Perfiles y niveles MPEG-4: Vídeo (VIII)

### Perfil visual Studio

Studio  
Profile

- Soporta aplicaciones de estudio
- Sólo frames I
- Soporta bitrates y calidades muy elevadas (hasta 600 Mbps para HDTV)
- Soporta 4:0:0, 4:2:2 y 4:4:4
- Formas definidas por máscaras binarias y en escala de grises

## Perfiles y niveles MPEG-4: Vídeo (IX)



## Perfiles y niveles MPEG-4: Vídeo (X)

Visual Tools	Visual Object Types						PGS
	Simple	Core	Main	Real Time Simple	Advanced Coding Efficiency	Advanced Simple	
Basic	✓	✓	✓	✓	✓	✓	✓
Error Resilience	✓	✓	✓		✓	✓	✓
Short Header	✓	✓	✓	✓	✓	✓	✓
B-VOP		✓	✓		✓	✓	✓
Method 1/2 Quantization		✓	✓		✓	✓	✓
P-VOP based Temporal Scalability		✓	✓		✓		
Binary Shape		✓	✓		✓		
Grey Shape			✓		✓		
Interlace			✓		✓	✓	✓
Sprite			✓				
Dynamic Resolution Conversion				✓			
NewPred				✓			
Global MC					✓	✓	
Quarter-pel MC					✓	✓	
SA-DCT					✓		
FGS							✓
FGST							✓

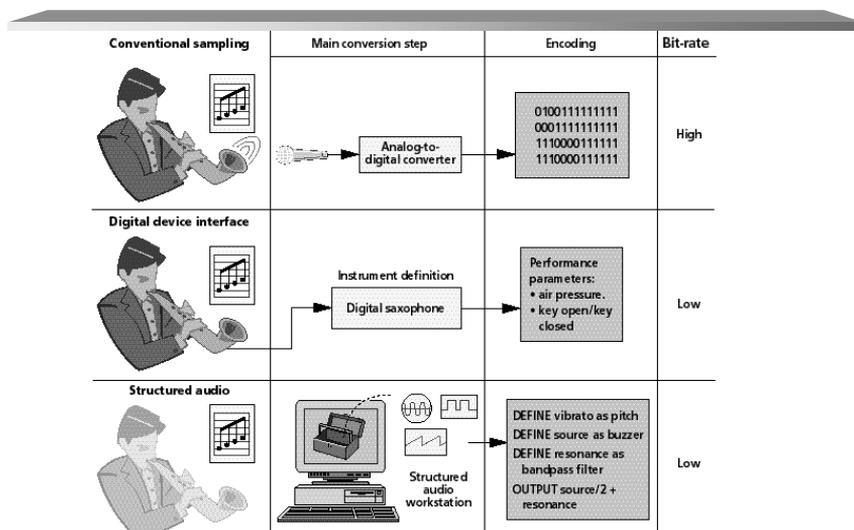
## Perfiles y niveles MPEG-4: Vídeo (XI)

- Very Low Bit rate Video (VLBV Core):
  - Formato imagen YUV 4:2:0. 8 bits/pixel
  - Resolución espacial y temporal dependiente de aplicación
  - Soporta frames I, P y B
  - Bloques de 8x8 y macrobloques de 16x16 para la estimación y compensación de movimiento
  - Codificación con DCT + cuantificación + entropía
  - Predicción eficiente de coeficientes DC y AC en la codificación intraframe
  - Slices de macrobloques adaptativos, marcadores de movimiento y mecanismos de resincronización mejorados

## Perfiles y niveles MPEG-4: Vídeo (XII)

- Perfiles para audio, vídeo, gráficos y descripciones de escenas
  - Vídeo
    - Simple
    - Simple Scalable
    - Core
    - Main
    - N-Bit
    - Advanced Real-Time Simple Profile (v. 2)
    - Core Scalable (v. 2)
    - Advanced Coding Efficiency (v.2)
    - Advanced Simple (v. 3)
    - Simple Studio (v. 3)
    - Core Studio (v. 3)
    - Fine Granularity Scalability

## MPEG-4: Audio. Funcionalidades (I)



## MPEG-4: Audio. Funcionalidades (II)

---

- **Audio:**
  - **Audio genérico**
    - Desde bajos bitrates hasta alta calidad (desde 6 kbps hasta multicanal)
      - PAC: Parametric Audio Coding, permite manipulación de sonido a baja velocidad
      - FGS: Fine Granularity Scalability
  - **Códecs para voz**
    - Bitrates de 2 a 24 kbps.
      - HVXC (Harmonic Vector eXcitation Coding) permite que el usuario pueda modificar la velocidad durante la reproducción
      - CELP:
        - » Muestreo 8 KHz => 6-12 kbps
        - » Muestreo 16 KHz => 18 kbps

## MPEG-4: Audio. Funcionalidades (III)

---

- **Audio:**
  - **Audio sintético**
    - Lenguaje para describir instrumentos (no sólo musicales)
    - Fórmulas matemáticas para la generación de sonidos
    - No estandariza el método para sintetizar
  - **Voz sintética**
    - Conversión texto voz (TTS: Text To Speech) permite codificar voz desde 200 bps hasta 1.2 kbps
    - Incluye la codificación de la voz y parámetros asociados:
      - Duración fonemas
      - Pitch contour

## **Perfiles y niveles MPEG-4: Audio**

---

- Audio
  - Speech
  - Synthesis
  - Scalable
  - Main
  - High Quality Audio (v. 2)
  - Low Delay Audio (v. 2)
  - Natural Audio (v. 2)
  - Mobile Audio Internetworking (v. 2)

## **Perfiles y niveles MPEG-4: Gráficos**

---

- Gráficos
  - Simple Facial Animation
  - Scalable Texture
  - Basic Animated 2-D Texture
  - Hybrid Visual
  - Advanced Scaleable Texture
  - Advanced Core
  - Simple Face and Body Animation

# MPEG



## MPEG: familia de estándares de compresión de vídeo

### MPEG-7

### MPEG-7



- “Multimedia Content Description Interface”
  - Estándar internacional en Septiembre de 2001
- Objetivo básico: permitir una búsqueda eficiente de contenido multimedia utilizando descripciones estándar.
- Objetivos del estándar:
  - Búsqueda de información multimedia
  - Filtrado
  - Gestión
  - Procesado

## MPEG-7 (II)

---

- MPEG 1, MPEG 2, MPEG 4 -> representación del contenido en sí (los bits)
- MPEG 7: representación de información sobre el contenido (los bits de los bits), *metadatos*
- Estándar de descripción, no de representación

## MPEG-7: Aplicaciones

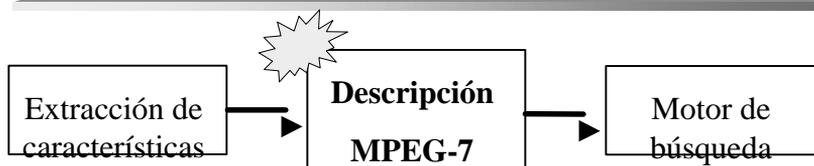
---

- Aplicaciones de navegación y búsqueda
- Bibliotecas Digitales.
  - Ej: catálogos de imágenes, diccionario musical, etc.
- Servicios de Directorios Multimedia.
  - Ej: páginas amarillas.
- Selección de Canales en TV: canal de radio, canal de TV.
- Edición multimedia
  - Servicio de noticias personalizado en función del perfil de usuario

## MPEG-7

- Métodos estandarizados para descripción de contenidos audiovisuales:
  - Imágenes
  - Gráficos
  - Modelos 3-D
  - Audio
  - Voz
  - Vídeo
  - Composición

## MPEG-7: Ámbito



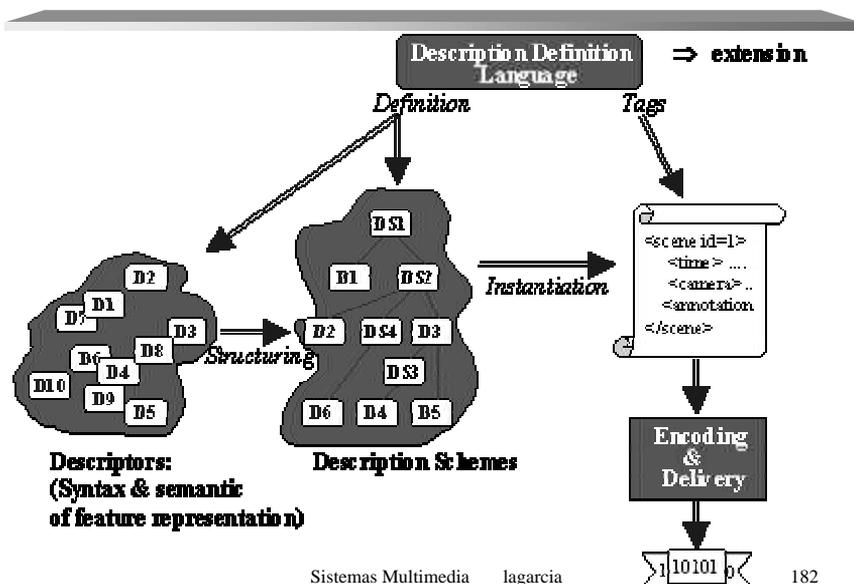
- |  |   |   |
|--|---|---|
| <ul style="list-style-type: none"><li>• Análisis de contenidos (D, DS)</li><li>• Herramientas de anotación (DS)</li><li>• Propiedad intelectual (DS)</li></ul> | <ul style="list-style-type: none"><li>• Esquemas de Descripción (DS)</li><li>• Descriptores (D)</li><li>• Lenguaje de descripción (DDL)</li></ul> | <ul style="list-style-type: none"><li>• Búsqueda &amp; filtrado</li><li>• Clasificación</li><li>• Manipulación</li><li>• Resumen e indexado</li></ul> |
|--|---|---|

# MPEG-7

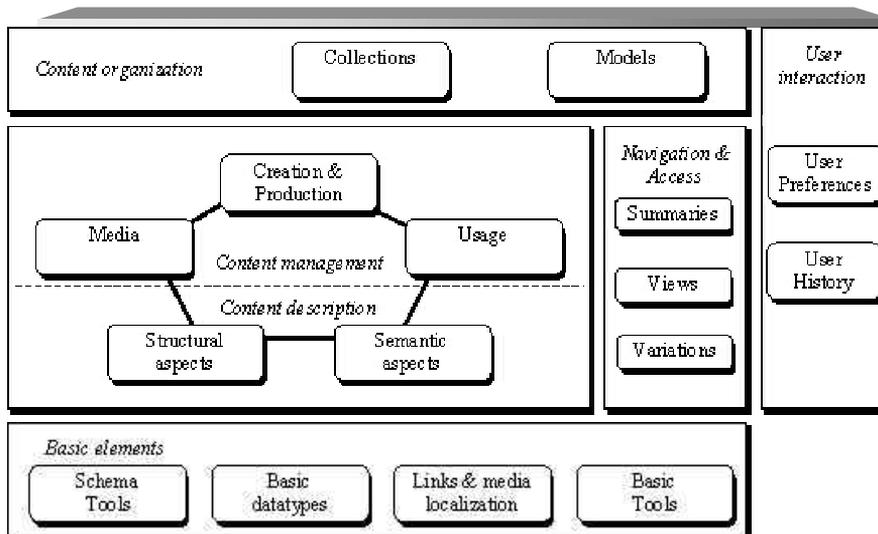
- Descriptor (D)
  - Representación escrita de un atributo o característica de un contenido multimedia
  - Puede ser:
    - Estructural (Color: negra y naranja)
    - Semántico (¿Qué es?: mariposa)
    - Catálogo (Autor: fotógrafo)
- Esquema de Descripción (DS)
  - Agrupación de descriptores
- Descripción:
  - Conjunto completo de D's y DS's que describen un objeto
- Lenguaje de descripción (DDL):
  - XML



## MPEG-7: agrupación de elementos



## MPEG-7: Contenido de MDS



**MDS: Multimedia Description Schemes**

Sistemas Multimedia lagarcia

183

## MPEG-7: ISO/IEC 15398

- **Descripción de contenidos audiovisuales**
  - Parte 1 – Sistemas
    - Transporte de descripciones, compresión
  - Parte 2 – DDL
    - Lenguaje de definición de descriptores. Ampliación de XML-Schema
  - Parte 3 – Vídeo
    - Descriptores de vídeo, imágenes estáticas y modelos 3D: color, textura, movimiento de cámara,...
  - Parte 4 – Audio
    - Descriptores de audio: tono, timbre, reconocimiento de habla, ...
  - Parte 5 – MDS Esquema de Descripción Multimedia
    - Descriptores que no son visuales ni audio puros: semánticos, catálogo, ... (autor, situación, actores, ...)
  - Parte 6 – XM
    - Modelo de experimentación: sw de referencia. Produce descripciones MPEG-7 de ejemplo

Sistemas Multimedia lagarcia

184

# MPEG

---

## MPEG: familia de estándares de compresión de vídeo

### Conclusiones

## Estándares MPEG: situación actual

---

- **MPEG-1:** Codificación Cerrado
- **MPEG-2:** Transmisión Cerrado
- **MPEG-4:** Codificación objetos En desarrollo (\*)
- **MPEG-7:** Metainformación Reciente



## Familia H.26x

---

### Otros estándares de codificación de vídeo:

## H.26x

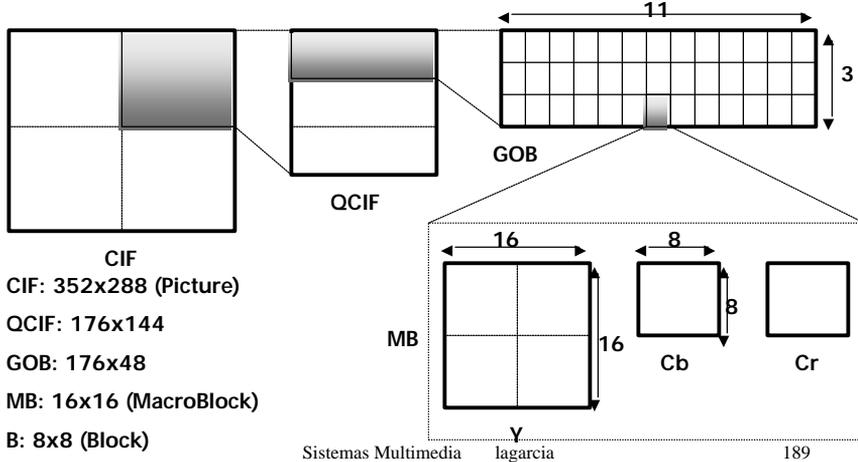
## H.261 (I)

---

- Estándar del CCITT para servicios de videoconferencia y videoteléfono sobre redes digitales, tipo RDSI,  
 $p \times 64 \text{ kbps}$ ,  $p = 1, \dots, 30$ .
- **Formato señal entrada:**
  - Interoperabilidad con NTSC y PAL
  - Resolución: QCIF (176x144) a 30 fps
  - Vídeo progresivo (no soporta entrelazado)
  - Formato Y Cb Cr (4:2:0)
  - También soporta QCIF 352x288
- **Codificación**
  - Utiliza redundancia temporal (inter) y espacial (intra)
  - Cuadros I y P (no B)
  - Estimación de movimiento con precisión “full pixel”
  - DCT, VLC

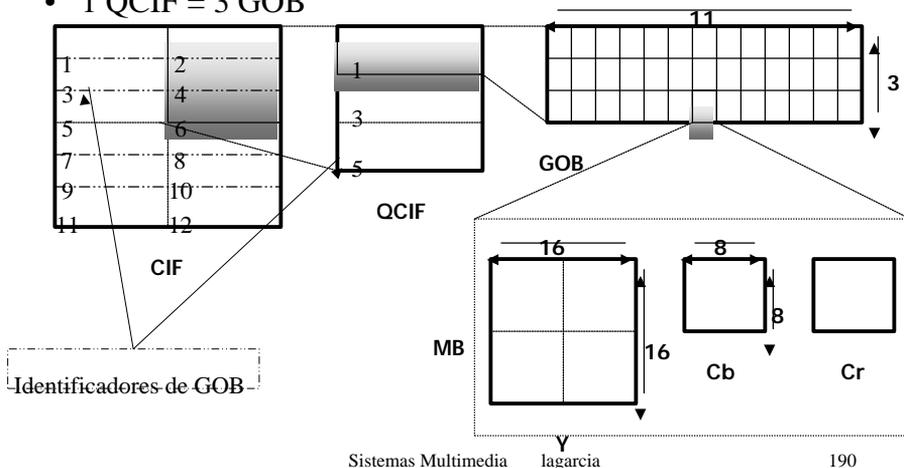
## H.261 (II)

- Stream organizado en 4 capas de compresión:
  - Picture, GOB (Group Of Block), MB (MacroBlock) y Block



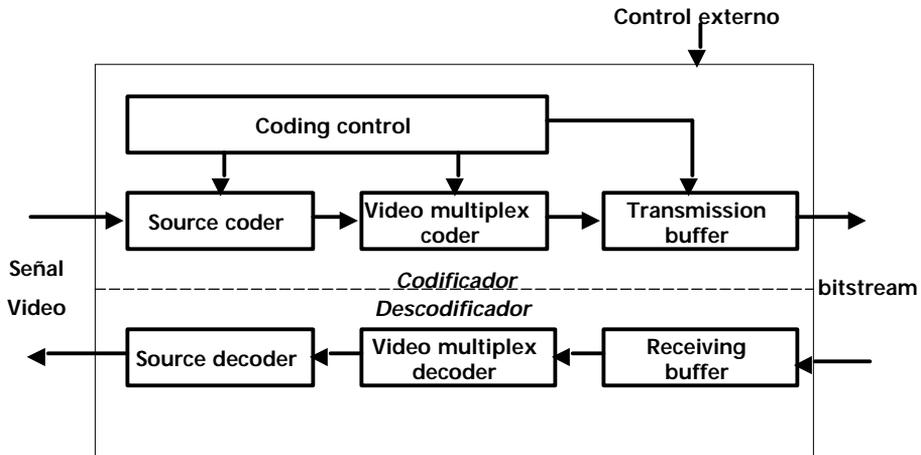
## H.261 (III)

- 1 CIF = 4 QCIF                      1 GOB = 11 x 3 MB
- 1 QCIF = 3 GOB



## H.261 (IV)

- Esquema genérico de codificador/descodificador:



## H.263

- Estándar para transmisión de videoconferencia a bajo ancho de banda.
- Evolución del H.261 para disminuir el ancho de banda
- Opera a tasas menores de 64 kbps.
- Formatos de imagen:
  - CIF (352x288), 30 fps, 4:2:0
  - QCIF (176x144), 30 fps, 4:2:0
  - SQCIF (128x96), 30 fps, 4:2:0
- Vídeo no entrelazado

## H.263 (II)

- Codificación similar a H.261:
  - Submuestreo temporal
  - Cuadros I, P y B
  - Mejoras de rendimiento y recuperación frente a errores.
  - VLC mejorada
  - Estimación movimiento con precisión “half pixel”
  - Vectores de movimiento no restringidos a la imagen (opcional)
  - Codificación aritmética (opcional)
  - Compensación de movimiento por bloques solapados (opcional)
  - Modo P-B (opcional)

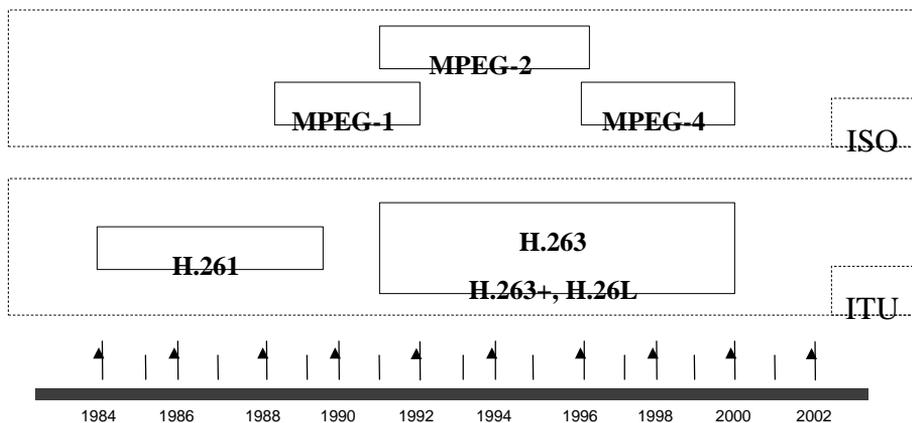
## Formatos de vídeo en H.261 y H.263

<b>Formato imagen</b>	<b>Resolución (HxV)</b>	<b>H.261</b>	<b>H.263</b>
SQCIF	128x96	Opcional	<b>Obligatoria</b>
QCIF	176x144	<b>Obligatoria</b>	<b>Obligatoria</b>
CIF	352x288	Opcional	Opcional
4CIF	704x576	Indefinido	Opcional
16CIF	1408x1152	Indefinido	Opcional

# Resumen

## Resumen y comparación de estándares

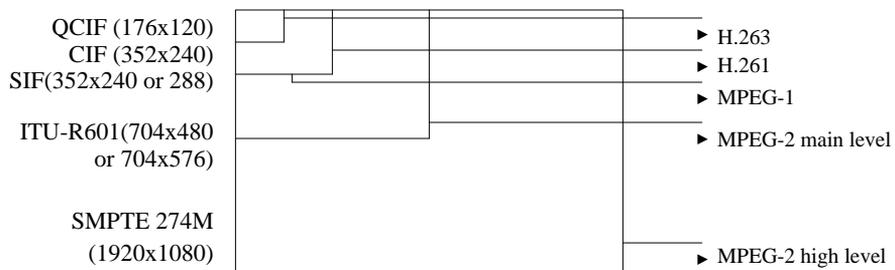
# Estándares de codificación



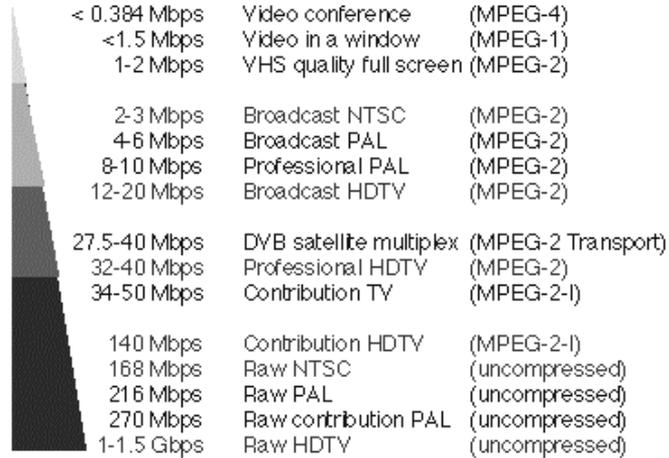
## Comparación estándares compresión

Parámetro	JPEG	MPEG	H.261
Compresión simétrica	SI	NO	NO
Compresión intraframe	SI	SI	SI
Compresión interframe	NO	SI	SI
Calidad imagen	Alta	Baja/Media/Alta	Baja
Requisitos ancho de banda	Alta	Baja/Media/Alta	Baja

## Formatos de vídeo



## Estándares de codificación: bitrates (I)



## Estándares de codificación: bitrates (II)

