

# El estándar de codificación de vídeo H.264/AVC

Claves para una codificación híbrida más potente (I)

---

Fernando Díaz de María y Manuel de Frutos López  
Dpto. de Teoría de la Señal y Comunicaciones  
Universidad Carlos III de Madrid

## Índice

---

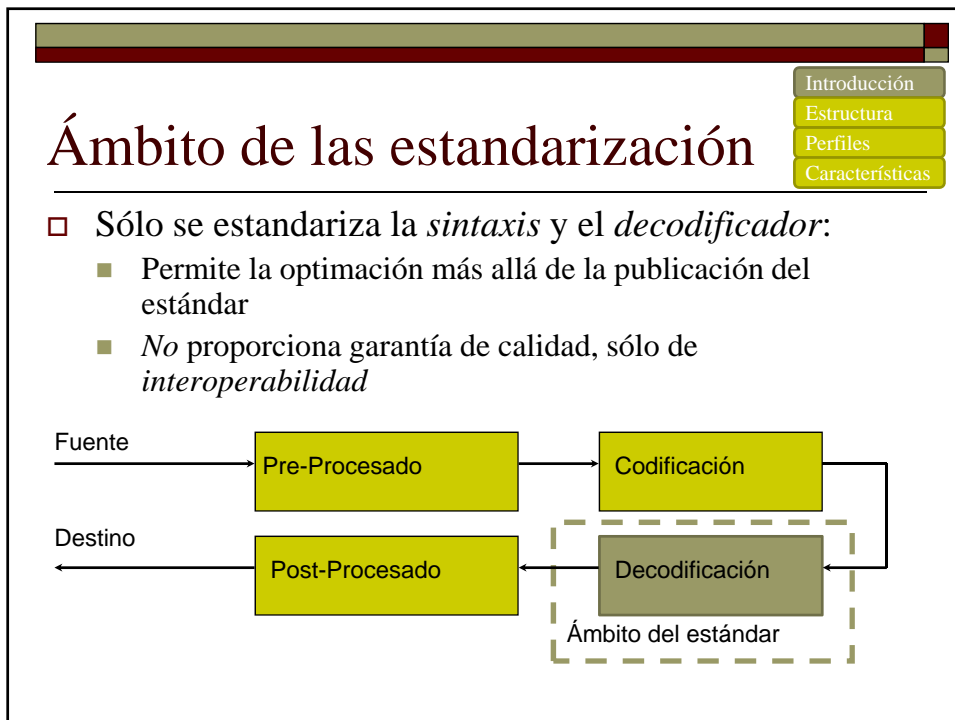
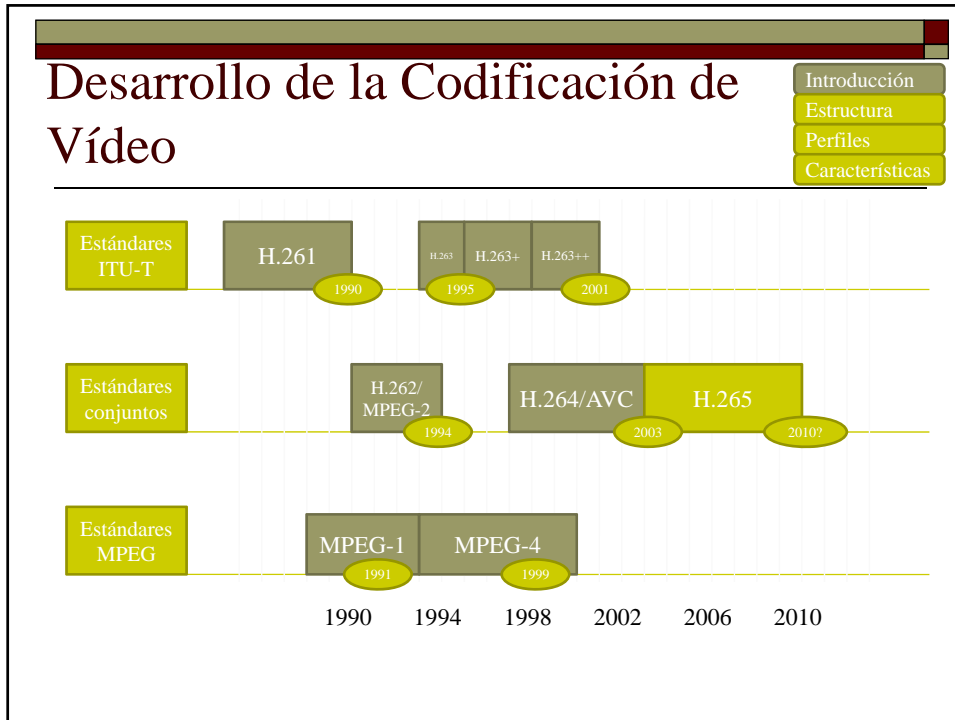
Introducción

Estructura

Perfiles

Características

- Introducción
- Estructura del H.264/AVC
- Perfiles
- Características relevantes
  - Filtro de suavizado
  - Codificación Entrópica
  - Transformación y Cuantificación
  - Codificación Intra
  - Estimación y Compensación de Movimiento



## Sobre los estándares previos

- Introducción
- Estructura
- Perfiles
- Características

---

- MPEG-2:
  - Éxito destacado (cientos de millones de decodificadores)
- MPEG-4:
  - Estándar enorme (capacidades muy específicas)
  - Retraso de licencias
  - MPEG-2:
    - Duplica prestaciones entre 2001 y 2006
    - MPEG-4: 15% de mejora respecto MPEG-2

## H.264/MPEG4-AVC

- Introducción
- Estructura
- Perfiles
- Características

---

- JVT: Joint Video Team (2001)
  - ITU: VCEG (Video Coding Expert Group)
  - ISO/IEC: MPEG (Motion Picture Expert Group)
- Estándar en 2003
  - H.264
  - MPEG-4 Part 10
  - Nombre oficial: AVC (Advanced Video Coding)

Introducción
Estructura
Perfiles
Características

## Consideraciones Generales

---

- Motivación:
  - Nuevos espacios de aplicación
    - “Mobile TV”
    - “HD TV”
  - “Network friendly” y mayor robustez frente a errores
  - Más amplio ámbito de aplicación
- Importantes ganancias en la relación calidad-régimen binario
- Considerable aumento de la carga computacional
- La estructura de codificación se mantiene inalterada respecto a otros estándares anteriores

Introducción
Estructura
Perfiles
Características

## AVC: Especificaciones

---

- “Back-to-Basis”
- No será compatible hacia atrás
- Pocas combinaciones de perfiles y niveles (“conformance points”)
- 50% de mejora (tasa) respecto a predecesores
- Flexibilidad frente a restricciones de retardo
  - “Low-Delay” para servicios conversacionales
  - Retardos más elevados para almacenamiento o “streaming”
- Asimetría codificador/decodificador
  - DVD, DTV
- Compromiso coste computacional codificador vs. calidad
- Especificación completa de la decodificación
- Altas prestaciones en aplicaciones de alta calidad

Introducción  
Estructura  
Perfiles  
Características

## Aspectos similares a estándares anteriores

---

- ❑ 4:2:0
- ❑ División de imágenes en tiras y MBs
- ❑ MBs de 16x16 para luma y 8x8 para croma
- ❑ “Block-based transform coding”
- ❑ Cuantificación y ordenamiento de los coeficientes de la transformación
- ❑ Codificación predictiva espacial de algunos parámetros
- ❑ Predicción temporal: EM y MVs
- ❑ Imágenes tipo I, B, P
- ❑ Codificación sin pérdidas de coefs., MVs, etc.

Introducción  
Estructura  
Perfiles  
Características

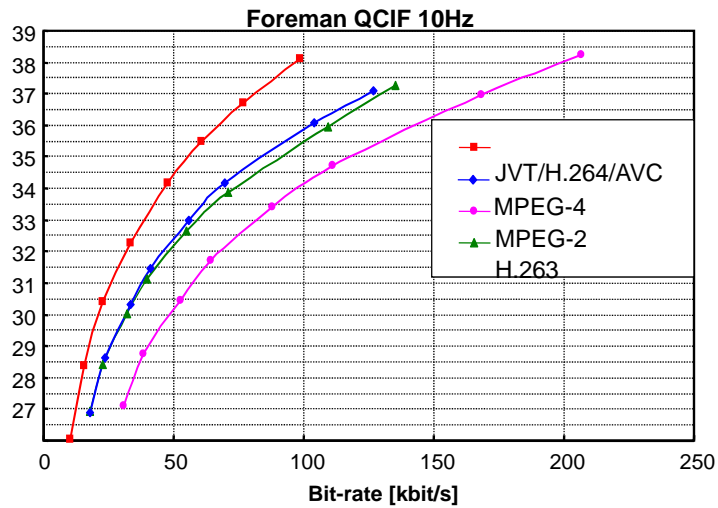
## Nuevos conceptos

---

- ❑ Transformada entera 4x4 (v1)
- ❑ Transformación adicional de los coeficientes DC de un MB
- ❑ Filtrado de suavizado de bordes dentro del bucle
- ❑ Versión extendida de la predicción espacial
- ❑ Distintos tamaños de bloque (estructura de árbol) para Estimación de Movimiento (EM)
- ❑ EM de ¼ pixel: importancia de la interpolación
- ❑ Multirreferencia
- ❑ Codificación sin pérdidas mejorada

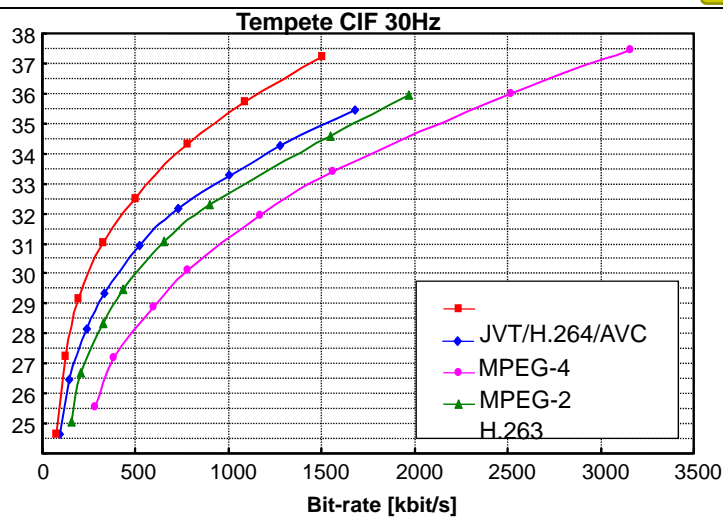
## Comparación con MPEG-2, H.263, MPEG-4

- Introducción
- Estructura
- Perfiles
- Características



## Comparación con MPEG-2, H.263, MPEG-4

- Introducción
- Estructura
- Perfiles
- Características



Introducción
Estructura
Perfiles
Características

## BitRate respecto a otros estándares

---

Reducción de tasa binaria

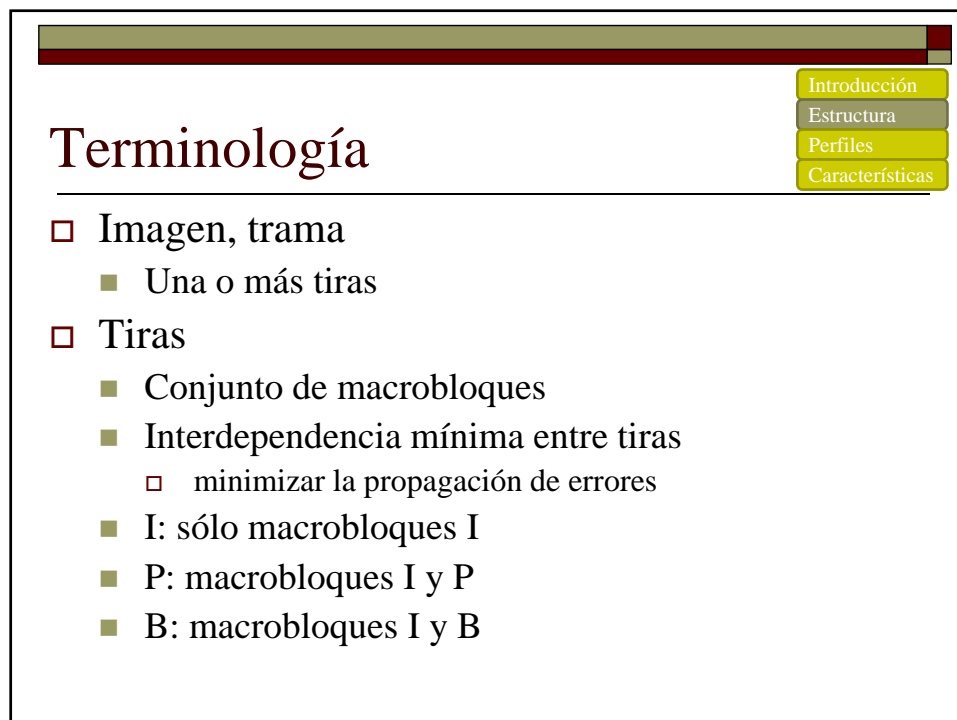
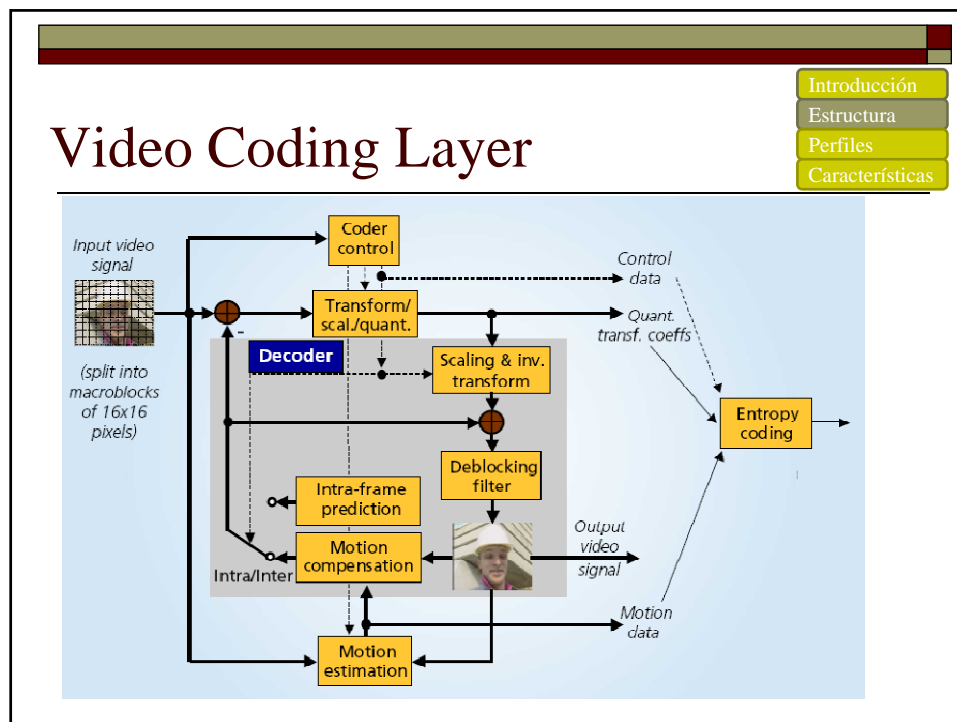
Codec	MPEG-4	H.263	MPEG-2
AVC	39%	49%	64%

Introducción
Estructura
Perfiles
Características

## Estructura del H.264/AVC

---

- Video Coding Layer (VCL)
  - Codificación de fuente
- Network Abstraction Layer (NAL)
  - Formateo de la representación generada en el VCL para una diversos niveles de transporte y medios de almacenamiento





Introducción  
Estructura  
Perfiles  
Características

## Terminología

---

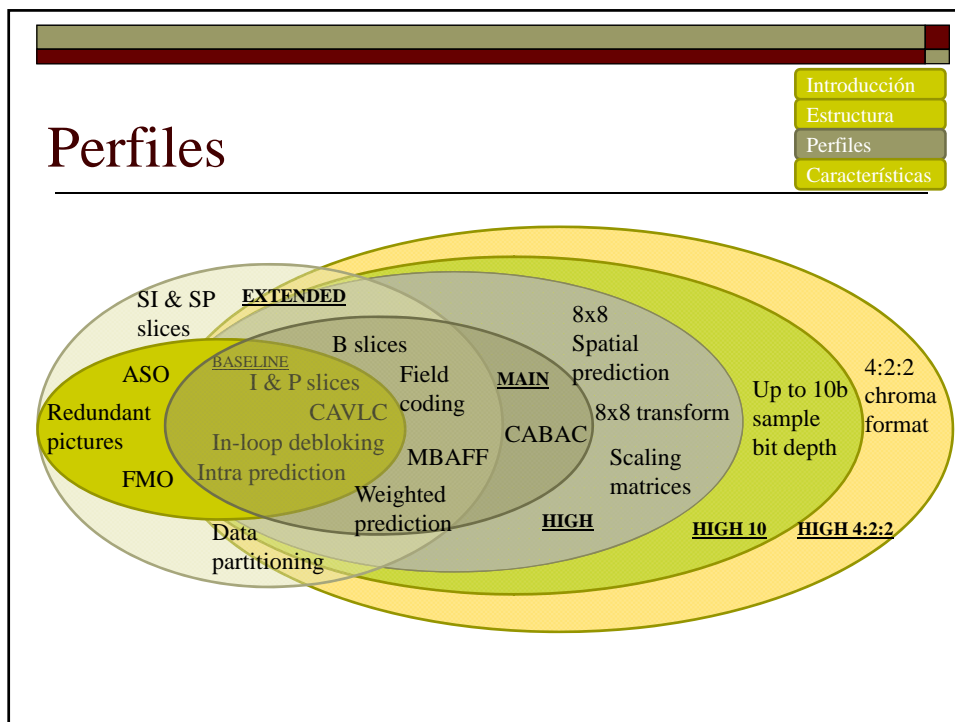
- **Macrobloques**
  - I: predicción *intra* a partir de píxeles decodificados de la misma tira
  - P: predicción *inter* a partir de una referencia
  - B: predicción *inter* a partir de una o dos referencias
- **Imágenes de referencia organizadas en dos listas:**
  - *lista 0*: primero referencias pasadas
  - *lista 1*: primero referencias futuras

Introducción  
Estructura  
Perfiles  
Características

## Perfiles

---

- **Baseline**
  - Mínimo coste computacional, máxima robustez frente a errores
  - Teléfonos móviles, “mobile TV”, videotelefonía, videoconferencia y comunicaciones inalámbricas
- **Main**
  - Máxima tasa de compresión, menor énfasis en robustez
  - Difusión de TV y almacenamiento 8DVD, Blu-ray)
- **Extended (contiene a baseline)**
  - Compromiso entre los dos anteriores con atención especial a las aplicaciones de video streaming
- **High**
  - Moderado incremento en complejidad respecto al Main
  - Vídeo profesional



# El perfil “High”

Introducción

Estructura

Perfiles

Características

Reducción de la tasa binaria			
Codificador	AVC HP con CAVLC	AVC MP con CABAC	MPEG-2 MP@HL
AVC HP con CABAC	17.9%	9.9%	58.8%

## Filtro de suavizado de bloques

- Introducción
- Estructura
- Perfiles
- Características
- "Deblocking"

---

- ❑ Suaviza el "efecto bloque"
- ❑ Dentro del bucle de compensación de movimiento
  - Permite que la estimación de movimiento se realice sobre planos decodificados que tienen un menor "efecto bloque"
- ❑ Actúa típicamente (es adaptativo según contenido) sobre 2 píxeles a ambos lados de la frontera
- ❑ Muy útil para factores de compresión altos
  - Incluso para factores bajos, cuando una secuencia es difícil, el efecto bloque sigue siendo llamativo
- ❑ Típicamente, reduce un 5-10% la tasa binaria

## Filtro de suavizado

- Introducción
- Estructura
- Perfiles
- Características
- "Deblocking"

---



1) Sin Filtro



2) Con el filtro de AVC

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Transformación](#)

## Transformación y Cuantificación

- Transformación
  - ¿Para qué hacer una transformación precisa de algo que va a cuantificarse de forma muy gruesa?
  - Transformada entera 4x4 (v1): mejora la interoperabilidad
  - Reduce el impacto visual del “efecto bloque”
  - Para alta fidelidad, una transformada más grande se adapta mejor a las texturas homogéneas
    - v2: 8x8
  - 16 transformaciones por MB
    - Transformada Hadamard sobre los 16 coeficientes DC

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Transformación](#)

## Transformación

- DCT 4x4:
 
$$Y = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} \cdot X \cdot \begin{bmatrix} a & b & a & c \\ a & c & -a & -b \\ a & -c & -a & -b \\ a & -b & a & -c \end{bmatrix}$$

con :

$$a = \frac{1}{2}; \quad b = \sqrt{\frac{1}{2}} \cos\left(\frac{\pi}{8}\right); \quad c = \sqrt{\frac{1}{2}} \cos\left(\frac{3\pi}{8}\right)$$

## Transformación

Introducción

Estructura

Perfiles

Características

Transformación

### □ Factorizando:

$$Y = \begin{pmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & -1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & -d \end{bmatrix} \cdot X \cdot \begin{bmatrix} 1 & 1 & 1 & d \\ 1 & d & -1 & -1 \\ 1 & -d & -1 & 1 \\ 1 & -1 & 1 & -d \end{bmatrix} \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix} \end{pmatrix}$$

con:

$$d = \frac{c}{b} = \sqrt{2} - 1 \approx 0.414$$

## Transformación

Introducción

Estructura

Perfiles

Características

Transformación

- Aproximando  $d$  por  $0.5$  y reescalando las matrices
  - Sólo hay multiplicaciones por 2 (desplazamientos)
  - El escalado puede precalcularse e integrarse con la etapa de cuantificación

$$Y = \begin{pmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \cdot X \cdot \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix} \end{pmatrix}$$

ahora (para garantizar la ortogonalidad de la transformación):

$$a = \frac{1}{2}; \quad b = \sqrt{\frac{2}{5}}$$

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Codificación](#)

## Codificación entrópica

- El número óptimo de bits para representar un símbolo no tiene por qué ser entero
  - Información:  $\log_2(1/P)$
  - Los métodos que asignan a una palabra código un número entero de bits no pueden alcanzar el óptimo
- Codificación aritmética
  - Alternativa a Huffman más cercana al óptimo
  - Secuencia de símbolos → número fraccionario

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Codificación](#)

## Codificación entrópica adaptativa

- Adaptativa al contexto
  - El éxito de la codificación entrópica depende de la precisión de las probabilidades de los símbolos
  - Se emplean características locales (espacio-temporales) para actualizar estas probabilidades.
- CAVLC (Context-based Adaptive Variable Length Coding) codificación de longitud variable adaptativa según contexto
- CABAC (Context-based Adaptive Binary Arithmetic Coding) : codificación binaria aritmética adaptativa según contexto
  - Mejoras sustanciales (10-20%) para altos regímenes binarios, tanto en SDTV como en HDTV

- Introducción
- Estructura
- Perfiles
- Características
- Intra

## Codificación INTRA en H.264

- LUMA
  - INTRA 16x16 (4 modos):
    - 1 DC
    - 2 direccionales
    - 1 “plano”
  - INTRA 4x4 (9 modos):
    - 1 DC
    - 8 direccionales
  - (v2): INTRA 8x8 (9 modos)
    - (filtro de suavizado)
    - 1 DC
    - 8 direccionales

- CROMA: Intra “8x8” (4 modos)
  - 1 DC
  - 2 direccionales
  - 1 “plano”

- Introducción
- Estructura
- Perfiles
- Características
- Intra

## Predicción INTRA

- 9 Modos de predicción luminancia (4x4):

0 (vertical)

1 (horizontal)

2 (DC)

3 (diagonal down-left)

4 (diagonal down-right)

5 (vertical-left)

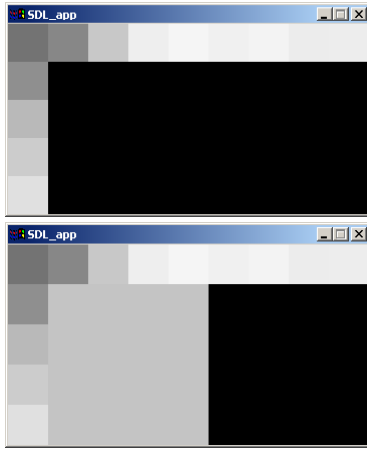
6 (horizontal-down)

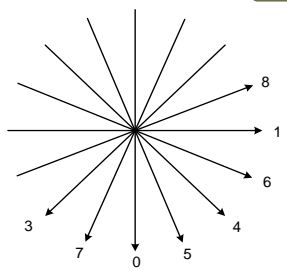
7 (vertical-right)

8 (horizontal-up)

## Codificación INTRA 4x4

- Introducción
- Estructura
- Perfiles
- Características
- Intra





## Codificación INTRA 4x4

- Introducción
- Estructura
- Perfiles
- Características
- Intra

- Modo más probable
 

	B
A	MB actual

Modo más probable =  $\min(\text{modo A}, \text{modo B})$
  
- Codificación del modo para cada bloque 4x4:
 

1	Modo más probable (1 bit)			
0	X	X	X	Modo no probable + modo de codificación (1+3 = 4 bits)



## Codificación INTRA 16x16

- Introducción
- Estructura
- Perfiles
- Características
- Intra

---

- 4 modos:
  - Vertical
  - Horizontal
  - DC
  - Plano
    - Zonas de la imagen donde la variación de la luminancia es suave

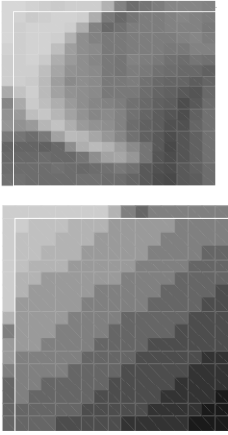


## Estimación de Movimiento

- Introducción
- Estructura
- Perfiles
- Características
- Inter

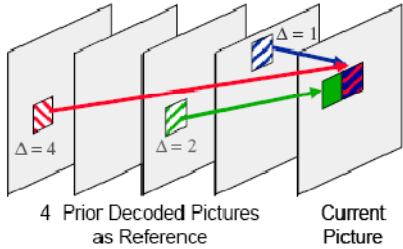
---

- Múltiples referencias almacenadas
  - Poco impacto computacional en el decodificador (más requerimientos de memoria)
- Predicción con una o dos referencias
- Precisión del VM de 1, ½ ó ¼ píxel
- Múltiples tamaños de bloque
  - Representación más precisa del campo de VMs vs. información lateral necesaria para representar los VMs
  - Útil cuando un MB (16x16) no pertenece a un mismo objeto
- Predicción ponderada
- Modos *Skip* y *Directo*



## Múltiples referencias

Introducción  
Estructura  
Perfiles  
Características  
Inter



4 Prior Decoded Pictures as Reference

Current Picture

## Tiras B

Introducción  
Estructura  
Perfiles  
Características  
Inter

- Generalización de las Bs de MPEG-4
  - Bipredicción vs. Bidireccional
    - Lista 0, lista 1, bi-predicción, predicción directa
  - Pueden usarse como referencia
- Predicción ponderada
  - Explícita:
    - Pesos como información en el *bitstream*
  - Implícita:
    - Cálculo de los pesos según distancia a las referencias

42

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Inter](#)


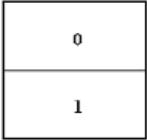
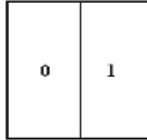
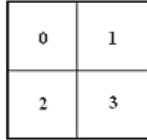


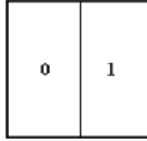
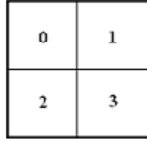
## VMs

- Interpolación de  $\frac{1}{2}$  y  $\frac{1}{4}$  de píxel
  - Se mejora la interpolación respecto a MPEG-2 y MPEG-4: primer filtro ( $\frac{1}{2}$ ) de orden 6, segundo ( $\frac{1}{4}$ ) de orden 2
- Vectores fuera de marco
- Vector de croma a partir del de luma
- Codificación predictiva. Se envía MVD

Estimación de movimiento en H.264 43

[Introducción](#)  
[Estructura](#)  
[Perfiles](#)  
[Características](#)  
[Inter](#)

## Partición del MB

MB	 16x16	 16x8	 8x16	 8x8
Sub-MB	 8x8	 8x4	 4x8	 4x4

Estimación de movimiento en H.264 44

## Modos *Skip* y *Directo*

- Introducción
- Estructura
- Perfiles
- Características
- Inter

---

- P\_SKIP (tiras P) o B\_SKIP (tiras B)
  - No se transmiten MVs, ni referencias, ni residuo
- Directo16x16 y 8x8 (tiras B)
  - No se transmiten MVs ni referencias
  - Espacial
    - Predicción del VM a partir del de los vecinos
  - Temporal
    - Predicción a partir de bloque cosituado

46

## Referencias

---

- T. Wiegand, G.J. Sullivan, G. Bjøntegaard and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 13, no. 7, pp. 560-573, Jul. 2003.
- G.J. Sullivan and T. Wiegand, " Video Compression – From Concepts to the H.264/AVC Standar", *Proceedings of the IEEE*, vol. 30, no. 1, pp. 18-31, Jan. 2005.
- T. Wiegand and G.J. Sullivan, "The H.264/AVC Video Coding Standard", *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 148-153, Mar. 2007
- D. Marpe, T. Wiegand and G.J. Sullivan, " The H.264/MPEG4 Advance Video Coding Standard and its Applications, *IEEE Communications Magazine*, vol. 44, no. 8, pp. 134-143, Aug. 2006.