



Sistemas de compresión de vídeo

GRADO EN INGENIERÍA DE TECNOLOGÍAS Y SERVICIOS DE
TELECOMUNICACIÓN

Dpto. de Comunicaciones
Universidad Politécnica de Valencia



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Universitat Politècnica de València
www.upv.es



OBJETIVOS

Índice

Objetivos

Contenido

- Conocer la funcionalidad de los bloques básicos de un sistema de compresión genérico.
- Entender el efecto del perfil del tráfico de vídeo en las comunicaciones.

CONTENIDOS

Índice

Objetivos

Contenido

1. Introducción
2. Conceptos básicos de compresión de vídeo
3. Sistemas estándar de compresión de vídeo
4. Actividades y prácticas
5. Referencias bibliográficas

Índice

Objetivos

Contenido

1. Introducción

2. Conceptos básicos de compresión de vídeo

3. Sistemas estándar de compresión de vídeo

4. Actividades y prácticas

5. Referencias bibliográficas

1. Introducción

1. Introducción

- **Definición:**

Comunicaciones Multimedia [Kuo et al]:

“Tecnología para manipular, transmitir, y controlar las señales audiovisuales a través de redes comunicaciones”

Sistema Multimedia [Steinmetz]:

“Sistema o aplicación que soporta el procesamiento integrado de varios tipos de información, siendo al menos uno de ellos dependiente del tiempo”



IoT
(televigilancia, smart city,...)

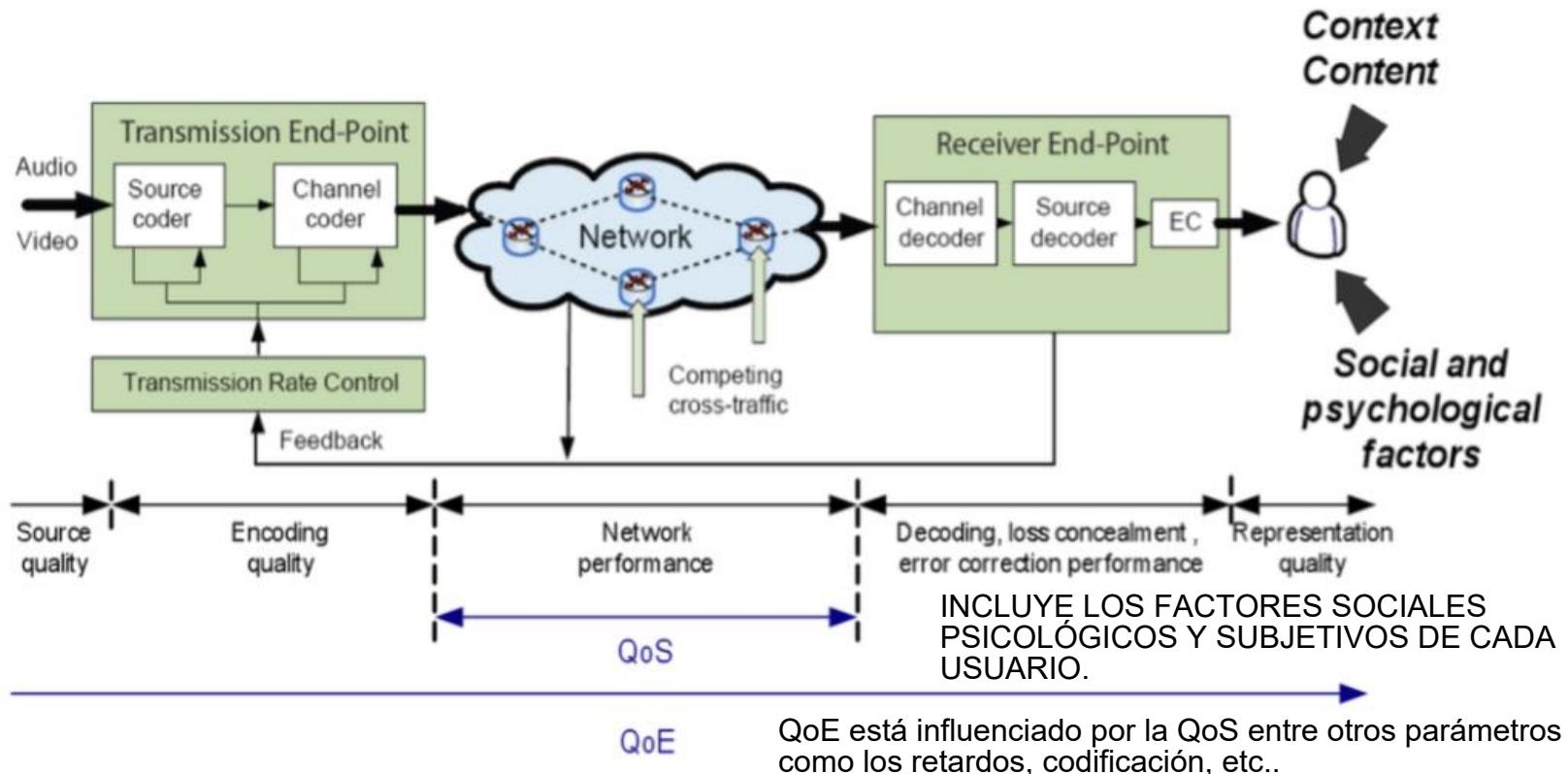
Inmersividad
(y efecto sensorial)

Calidad del Vídeo y 3D

Multipantalla
(y sincronización de contenidos)

1. Introducción

QoE in Networked Multimedia



Quality of Experience in Multimedia Systems and Services: A Journey Towards the Quality of Life, Christian Timmerer, Fernando Pereira, Touradj Ebrahimi, IEEE International Conference on Multimedia & Expo (ICME) 11th July 2016, Seattle, WA, USA.

1. Introducción

¿Qué **parámetros de codificación** se suelen utilizar en las aplicaciones de Internet?

1. Introducción



Contenedor: MP4

- Sin listas de edición (de lo contrario, el vídeo podría no procesarse correctamente)
- Atom moov al principio del archivo (inicio rápido)

Códec de audio: AAC-LC



Códec de vídeo: H.264



- Exploración progresiva (sin entrelazado)
- High Profile
- Dos fotogramas de B consecutivos
- GOP cerrado (grupo de imágenes equivalente a la mitad de la frecuencia de imagen)
- Codificación CABAC
- Frecuencia de bits variable. No se requiere ningún límite, pero puedes consultar los valores de frecuencia de bits recomendados a continuación como referencia
- Submuestreo de croma: 4:2:0

1. Introducción

HDR/SDR refiere al rango dinámico para codificar el color de imagen, y no a la definición.
HDR se realiza con 10 bits/muestra
SDR se realiza con 8 bits/muestra



Frecuencias de bits de vídeo recomendadas para las subidas SDR

Para ver vídeos nuevos en 4K, usa un navegador o dispositivo compatible con VP9.

Tipo	Frecuencia de bits, Frecuencia de imagen estándar (24, 25, 30)	Frecuencia de bits, Frecuencia de imagen alta (48, 50, 60)
2160p (4k)	35-45 Mbps	53-68 Mbps
1440p (2k)	16 Mbps	24 Mbps
1080p	8 Mbps	12 Mbps
720p	5 Mbps	7,5 Mbps
480p	2,5 Mbps	4 Mbps
360p	1 Mbps	1,5 Mbps

Tipos/formatos de vídeo en función de la resolución para cada uno de ellos.

Frecuencias de bits de vídeo recomendadas para las subidas HDR

Tipo	Frecuencia de bits, Frecuencia de imagen estándar (24, 25, 30)	Frecuencia de bits, Frecuencia de imagen alta (48, 50, 60)
2160p (4k)	44-56 Mbps	66-85 Mbps
1440p (2k)	20 Mbps	30 Mbps
1080p	10 Mbps	15 Mbps
720p	6,5 Mbps	9,5 Mbps

Es interesante conocer las características del ojo humano / visión para definir los parámetros de codificación de imagen. Así podemos hacerlo más eficiente.

<https://support.google.com/youtube/answer/1722171?hl=es>

1. Introducción



23.976 fps

Tasa binaria

Equivalente en almacenamiento

Resolution	Streaming Bitrate	Hourly Conversion
480p (720x480)	1750 kbps	~792 MB per hour
720p (1280x720)	3000 kbps	~1.3 GB per hour
1080p (1920x1080)	4300-5800 kbps	~1.9 GB to ~2.55 GB per hour
1440p (2560x1440)	6350 kbps	~2.8 GB per hour
4K (3840x2160)	8000-16000 kbps	~3.5 GB to ~7 GB per hour



8K VP9 vs H264 bitrate comparison = **21.2**Mbps VP9 vs **78.4** Mbps H264

4K VP9 vs H264 bitrate comparison = **17.3**Mbps VP9 vs **23.1** Mbps H264

2K VP9 vs H264 bitrate comparison = **8.5** Mbps VP9 vs **10.4** Mbps H264

1080p VP9 vs H264 bitrate comparison = **2.5** Mbps VP9 vs **2.3** Mbps H264

720p VP9 vs H264 bitrate comparison = **1.4** Mbps VP9 vs **1.3** Mbps H264

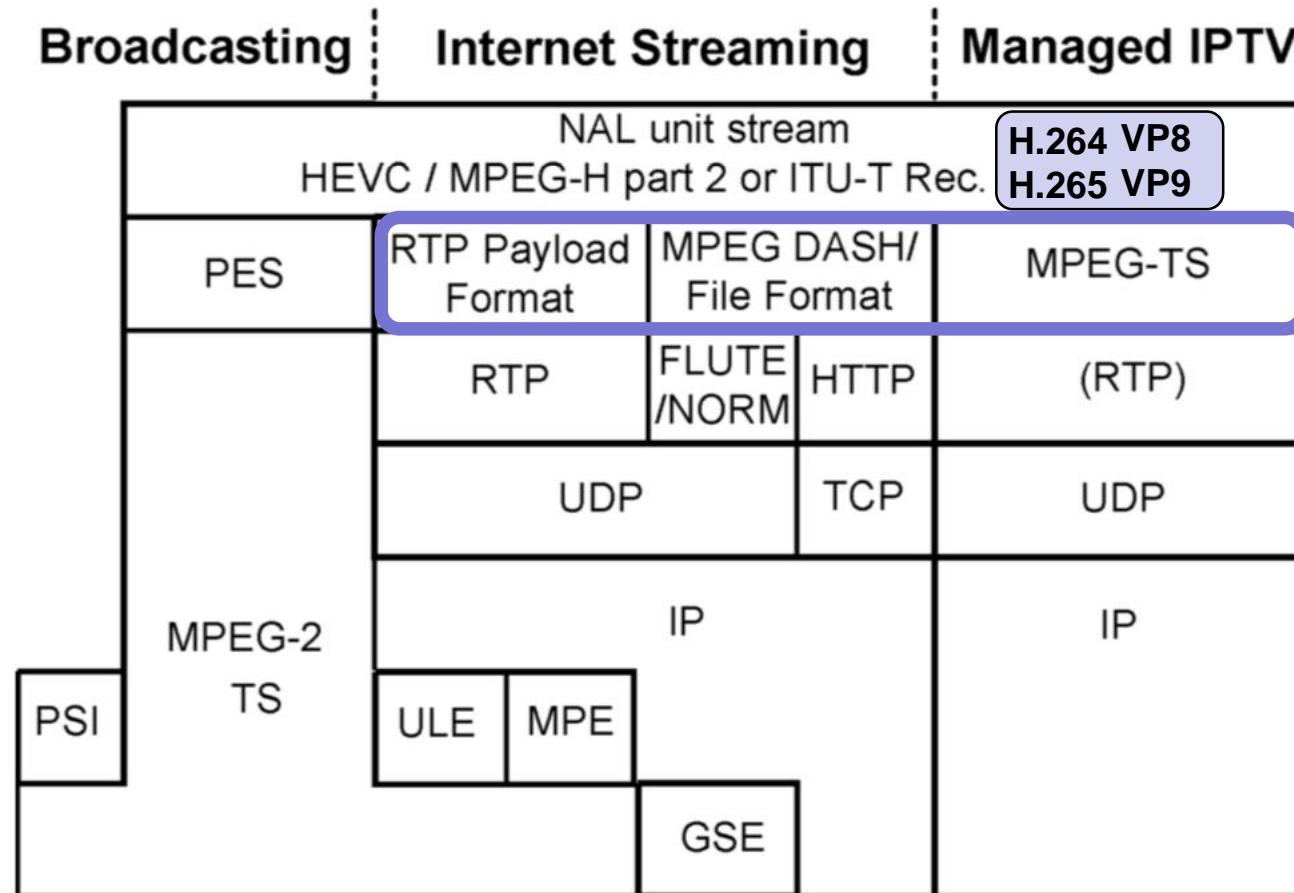
<https://www.howtogeek.com/338983/how-much-data-does-netflix-use/>

<https://www.tutorialguidacomefare.com/test-video-quality-720p-1080p-1440p-2160p-max-bitrate-which-compresses-youtube/>

1. Introducción

Protocolos para distribución de contenidos Multimedia:

Están por encima de los protocolos de transporte y aplicación que gestionan la comunicaciones multimedia.



Índice

Objetivos

Contenido

1. Introducción

2. Conceptos básicos de compresión de vídeo

3. Sistemas estándar de compresión de vídeo

4. Actividades y prácticas

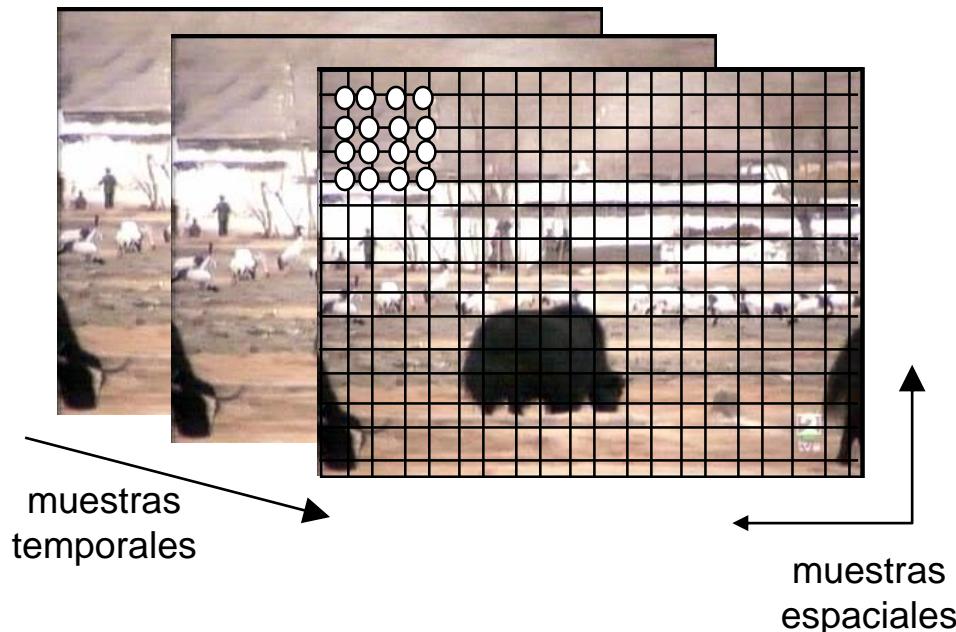
5. Referencias bibliográficas

2. CONCEPTOS BÁSICOS DE COMPRESIÓN DE VÍDEO

2. Conceptos básicos: digitalización

- **Fuente de información:** escena visual “natural” continua en espacio y tiempo
- **Señal digital:** muestrear (*picture element* o pixel) la escena real espacialmente (patrón rectangular de la imagen) y temporalmente (imágenes por unidad de tiempo)
- **Codificación de vídeo:** proceso de compresión y descompresión de una señal de vídeo digital
- **Pixel:** representado por valores que representan la luminancia y el color de la muestra

El ojo humano es más sensible a la luminancia.



Por debajo
de 20 frames/segundo
nuestro ojo ya puede
distinguir la traslación de
imágenes de un vídeo.
(se apreciaba en proyectores
antiguos)

Nuestra visión también da más resolución a una parte del espacio que nos rodea que a otras partes.
Focalizamos en una zona de nuestra visión directa.

2. Conceptos básicos: formatos de color

- **RGB:** cada pixel se codifica con ocho bits por cada color (24 bits)

Red / Green / Blue (el ojo humano es igual de sensible a los tres)

- **Luminancia y crominancia:** permite comprimir la señal aprovechando que el ojo humano es más sensible al brillo (luminancia, Y) que al color (crominancia, Cb, Cr)

Matriz de transformación / conversión (valores estándar)

$$\begin{bmatrix} Y' \\ Cb' \\ Cr' \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

En el procesado se trabaja con las componentes de luminancia (Y) y crominancia (Cb,Cr)

$$Y = Y' \times \frac{219}{255} + 16 \quad Cb = Cb' \times \frac{224}{255} + 128 \quad Cr = Cr' \times \frac{224}{255} + 128$$



YCbCr



Y



Cb

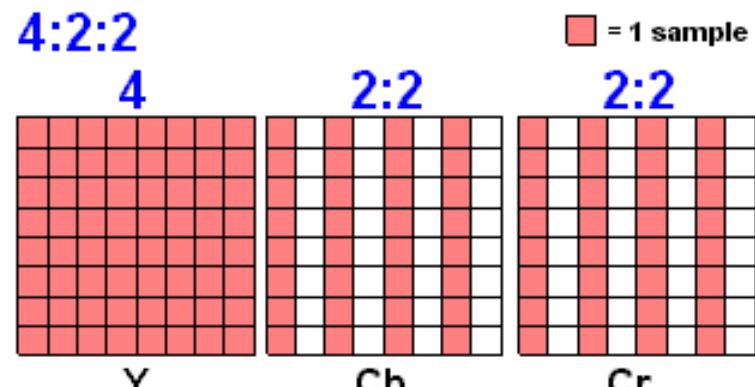
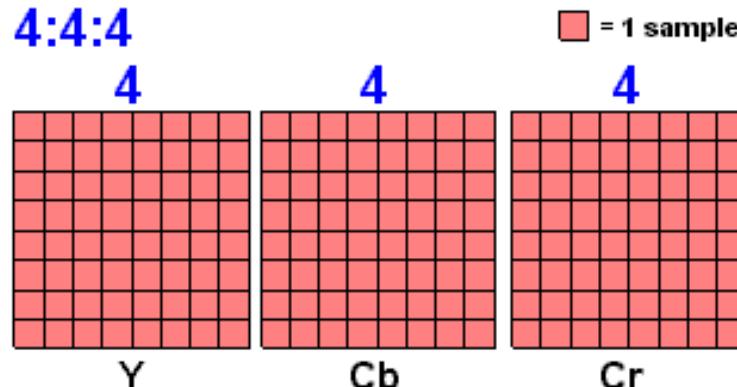


Cr

Cuando se transportan, se envían las matrices Y, Cb y Cr

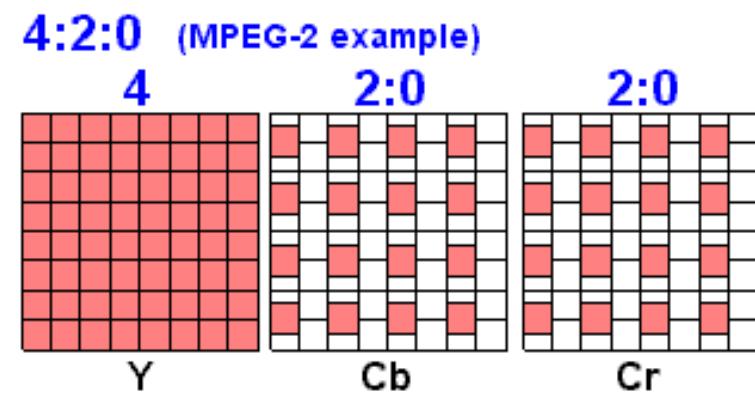
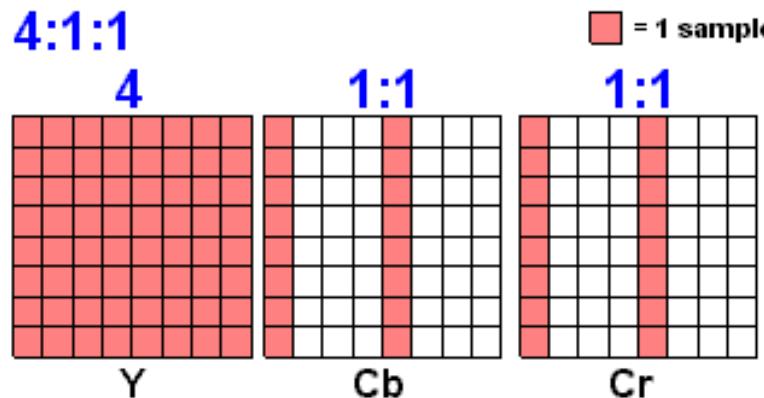
2. Conceptos básicos: formatos de muestreo

Estamos comprimiendo el vídeo mediante submuestreo con la menor pérdida posible de información.



El ojo es más sensible en horizontal podemos aprovechar para hacer un submuestreo 4:2:2 reduciendo el volumen de información sin apenas apreciarse en calidad de imagen.

Los espacios en blanco se llenan con el valor adyacente (izquierda). Esto es inapreciable)



En horizontal, nos quedamos con un píxel cada cuatro.

En 4:2:0 es como el 4:2:2 pero toma el valor medio de los píxel como se ve en la ilustración superior.

2. Conceptos básicos

¿Cuáles son las **tasas de transmisión** (o espacio ocupado) por los principales formatos de vídeo (SD, HD, 4K, ...) **sin comprimir**?

Es un dato importante para comparar otros formatos de compresión o uso de vídeo como medida de calidad de estos.

2. Conceptos básicos: formatos de vídeo

Formato	Resolución Y (horiz. x vert.)	
Sub-QCIF	128x96	
Quarter CIF (QCIF)	176x144	→
CIF	352x288	
4CIF	704x576	



CIF=Common Intermediate Format

2. Conceptos básicos: formatos de vídeo



180 x 120, e.g., Cell Phone



360 x 240, e.g., PDA, Video Phone



720 x 480, e.g., Standard TV



1920 x 1080, e.g., High Definition TV



2. Conceptos básicos: formatos de vídeo

8K 7680 x 4320

4K 4096 x 2160

4 veces la resolución 2K

2K 2048 x 1080

1080p 1920 x 1080

720p 1280 x 720

720p significa que son 720 píxeles en vertical y 1280 en horizontal
1080p con 1920 pixeles horizontales y 1080 verticales

2. Conceptos básicos: formatos de vídeo

EJEMPLO:

Formato	Resolución Y (horiz. x vert.)
---------	----------------------------------

Quarter CIF (QCIF) 176x144 →

Distintas formas de muestreo de una imagen:

Vídeo QCIF: 176x144 pixels y 15 fps

Resolución Y: 176x144 muestras (8bits/muestra)



4:4:4 Resolución Cb, Cr: 176x144 muestras (8bits/muestra)

Tasa de Bits: $176 \times 144 \times 8 \times 3 \times 15 = 9.123.840$ bps Muy elevado para una frecuencia baja (15fps)

4:2:0 Resolución Cb, Cr: 88x72 muestras (8bits/muestra)

Tasa de Bits: $(176 \times 144 \times 8 \times 15) + (88 \times 72 \times 8 \times 2 \times 15) = 4.561.920$ bps

Se ha reducido a la mitad aunque sigue siendo muy alta para este vídeo en función de la capacidad de transmisión

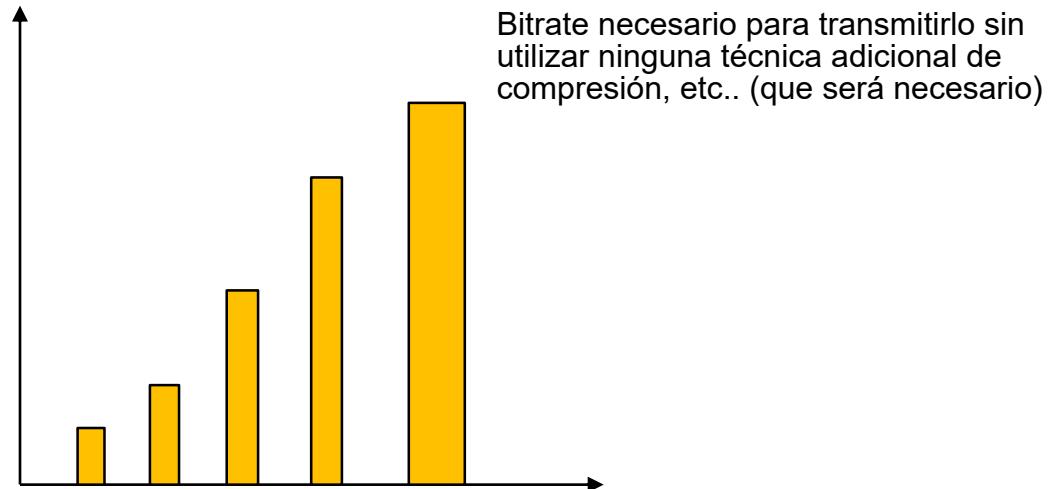
2. Conceptos básicos: formatos de vídeo

Ejercicio:

(Realizar
en casa)

Formato de vídeo	Bitrate
Vídeo 720p, 4:2:0, 24 fps, 8 bits/pixel	
Vídeo 1080p, 4:2:0, 24 fps, 8 bits/pixel	
Vídeo 2K, 4:2:2, 24 fps, 10 bits/pixel (HDR)	
Vídeo 4K, 4:2:2, 24 fps, 10 bits/pixel (HDR)	
Vídeo 8K, 4:4:4, 24 fps, 10 bits/pixel (HDR)	

Discusión de resultados:



2. Conceptos básicos

¿Cómo se mide la **calidad** de una **imagen** o de un **vídeo**?

2. Conceptos básicos: calidad

- Métricas Objetivas de calidad visual:

Error cuadrático medio (MSE)

$$MSE = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_{org}(m, n) - I_{dec}(m, n)]^2$$

M: filas; N: columnas

Diferencia entre original
y decodificada

Relación Señal Ruido (SNR)

$$SNR = 10 \cdot \log_{10} \frac{\frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{org}^2(m, n)}{MSE}$$

Relación Señal Ruido de Pico (PSNR)

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{MSE}$$

```
function psnr_comm(I,J)
%Cálculo del PSNR
ImaOrigRGB = imread(I);
ImaCodRGB= imread(J);
ImaOrigY = rgb2gray(ImaOrigRGB);
ImaCodY = rgb2gray(ImaCodRGB);

[m,n]=size(ImaOrigY);

error=abs(ImaOrigY-ImaCodY);
error_double=double(error);
error2=error_double.^2;
errorsum=sum((sum(error2)));
error3=errorsum/(m*n);
psnr=10*log10((255*255)/(error3));
disp(sprintf('PSNR = +%5.2f dB',psnr))
figure, imshow(ImaOrigRGB),figure, imshow(ImaCodRGB);
figure, imshow(ImaOrigY), figure, imshow(ImaCodY);
```

2. Conceptos básicos: calidad



Imagen Original RGB



Imagen RGB 32 coef.

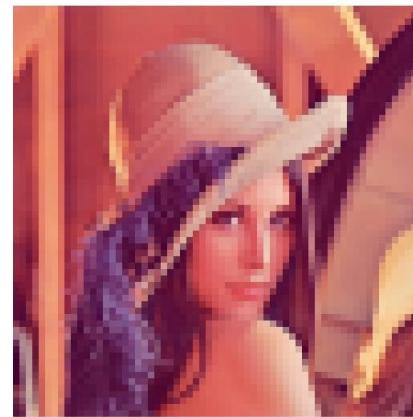


Imagen RGB 1 coef.



Imagen Original Y



Imagen Y 32 coef.
PSNR = 46.32 dB

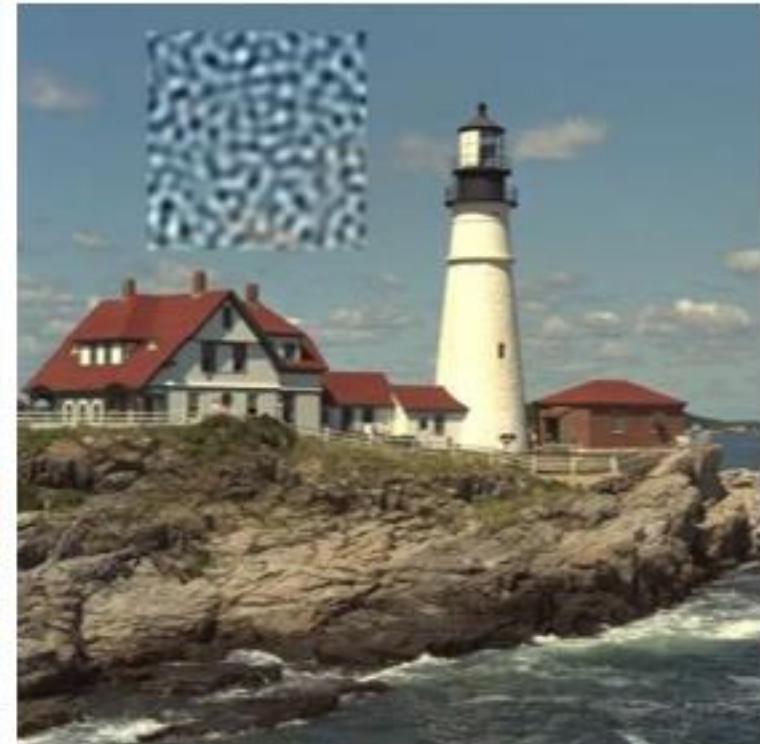


Imagen Y 1 coef.
PSNR = 25 dB

2. Conceptos básicos: calidad



(a)

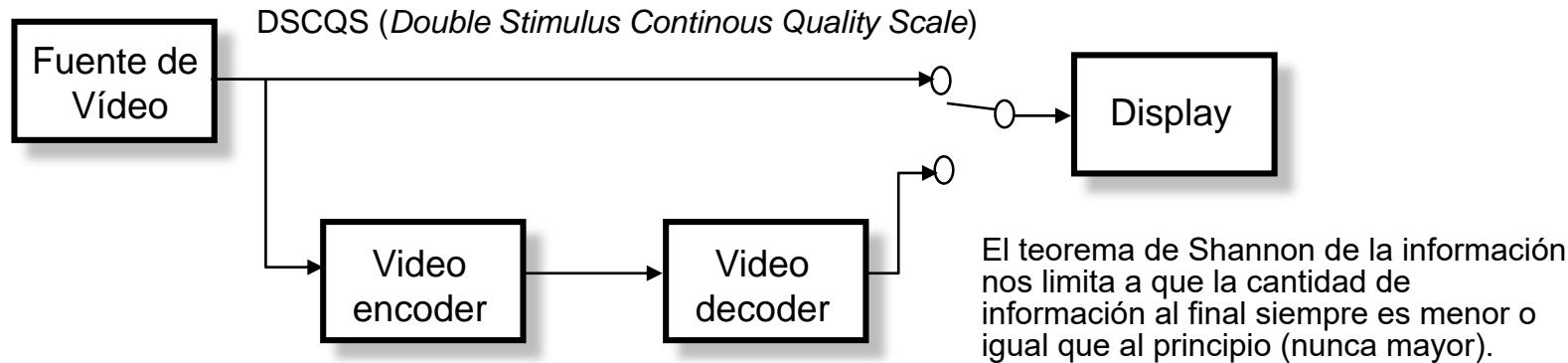


(b)

Objetivamente ambos tienen el mismo PSNR, sin embargo subjetivamente la imagen b) tiene menor calidad porque los errores se concentran en una zona. En la a) están distribuidos por toda y se aprecia con mayor calidad.

2. Conceptos básicos: calidad

- Métricas Subjetivas de calidad visual (**MOS**, *Mean Opinion Score*):



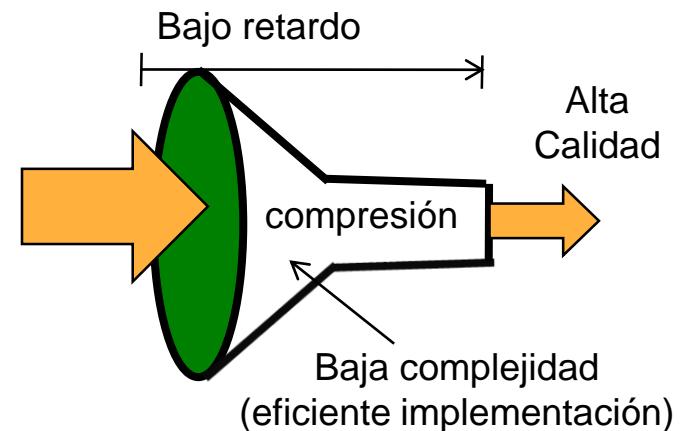
- Métricas de Prestaciones:

- **Ratio de compresión** (R_{comp})

$$R_{comp} = \text{tamaño_}I_{org}/\text{tamaño_}I_{comp}$$

- **Tiempo de Compresión** (t_{comp})

- **Tiempo de Descrompesión** (t_{decomp})



Índice

Objetivos

Contenido

1. Introducción

2. Conceptos básicos de compresión de vídeo

3. Sistemas estándar de compresión de vídeo

4. Actividades y prácticas

5. Referencias bibliográficas

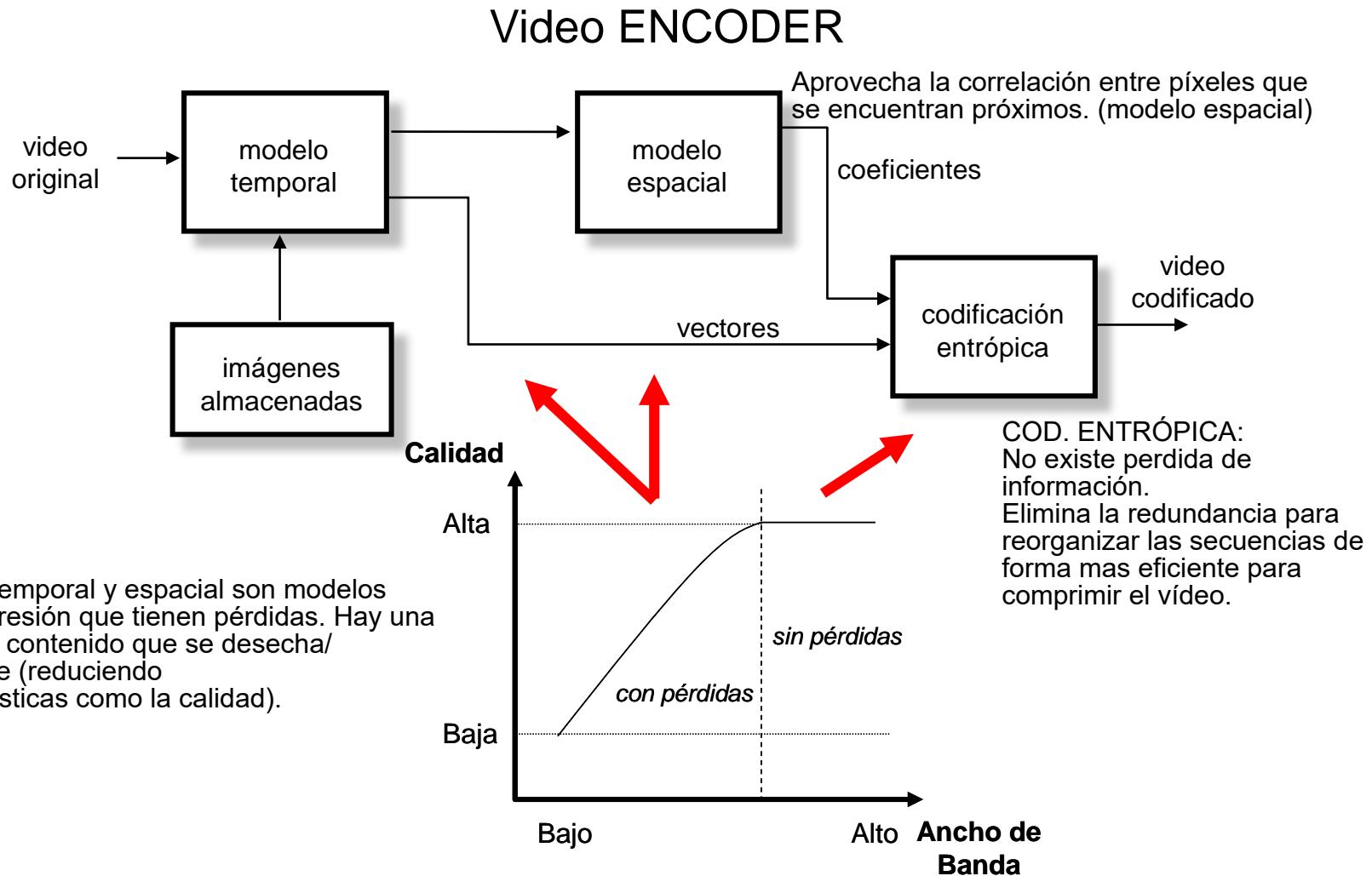
3. SISTEMAS ESTÁNDAR DE COMPRESIÓN DE VÍDEO

3. Sistemas estándar de compresión

¿Es posible **comprimir un vídeo** sin que se **degrade la calidad**?

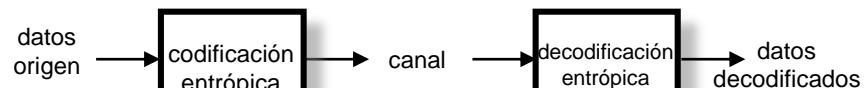
Se puede comprimir pero no lo vamos a hacer indefinidamente sin degradar la calidad.
Se verá afectado por una bajada de resolución, etc...

3. Sistemas estándar: modelo genérico



3. Sistemas estándar: codificación entrópica

- Sistema de compresión de propósito general
 - Elimina *redundancia estadística* de los datos
 - Se consigue cambiando secuencias de bits por secuencias codificadas que ocupan menos espacio
 - No se producen pérdidas de información
 - Útil para ficheros de texto (*zip, arj,*)
- No se puede comprimir un archivo ya comprimido con codificación entrópica pues ya se ha alcanzado el estado óptimo.
- La compresión con pérdidas si lo permite.



• RLE (*Run-Level Encoding*)

- Representa la información más compacta, representando la serie de símbolos por valores codificados 3D (*last, run, level*)

Es práctico cuando hay una cantidad considerable de 0s.

Esto implica una reducción de datos considerable.
Es la misma información de manera más eficiente.

Serie de datos: 16,0,0,-3,5,6,0,0,0,0,-7,0,0.....0

Datos codificados: (0,0,16), (0,2,-3),(0,0,5),(0,0,6),(1,4,-7)

El número de valores de la secuencia debe ser un valor conocido de antemano.

• VLC (*Variable-Length Coding*)

- Mapea símbolos origen a palabras código variables (VLC's)
- Las palabras código pueden tener tamaños variable (número entero de bits)
- Símbolos más frecuentes se representan con VLC's de menor tamaño y símbolos menos frecuentes con VLC's de mayor tamaño

La idea es que el valor medio mejore con esta técnica del último punto.

Dada una secuencia, la representamos como ternas
last: es un 0 siempre excepto en la última terna.

run: numero de 0 previos al valor

level: el valor de la secuencia original.

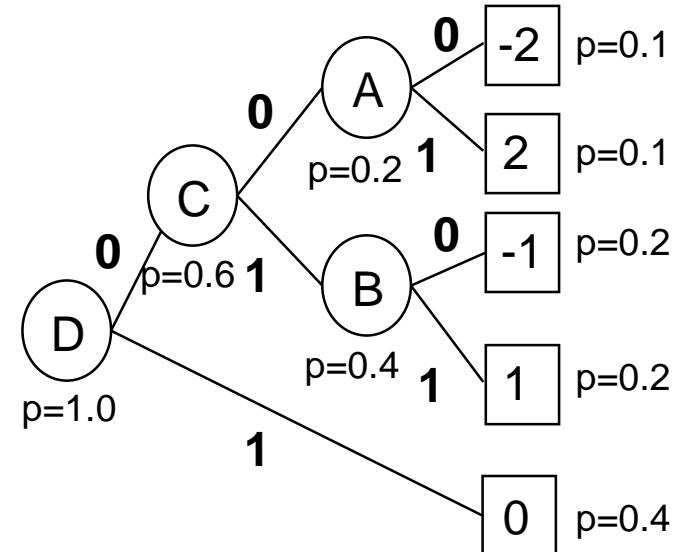
3. Sistemas estándar: codificación entrópica

Ejemplo: **Codificación Huffman**

- La asignación de símbolo --> VLC en función de la probabilidad de aparición de cada símbolo
- Fases: calcular la probabilidad y asignar las palabras código

Símbolo	Probabilidad (p)	$\log_2(1/p)$
-2	0.1	3.32
-1	0.2	2.32
0	0.4	1.32
1	0.2	2.32
2	0.1	3.32

Símbolo	VLC	Bits (actual)	$\log_2(1/p)$
-2	000	3	3.32
-1	010	3	2.32
0	1	1	1.32
1	011	3	2.32
2	001	3	3.32



Bits recibidos 0111000
Símbolos 1 0 -2

3. Sistemas estándar: codificación entrópica

Ejemplo: **Codificación Huffman aplicada a la transmisión de vídeo**

- El decodificador debe utilizar la misma tabla que el codificador
- La transmisión de la tabla añade *overhead* (reduce la eficiencia de compresión especialmente para vídeos de pequeño tamaño)
- La tabla no puede calcularse hasta que el vídeo entero ha sido codificado (retardo en el proceso de codificación)
- Estándares de compresión de imágenes y vídeo utilizan tablas basadas en la probabilidad de vídeos “genéricos”

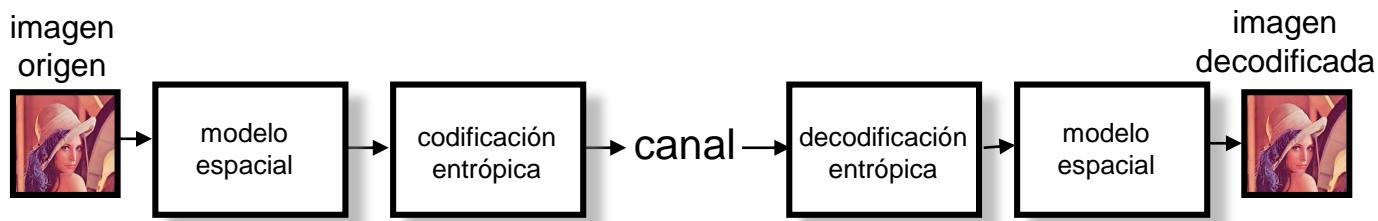
Símbolo	VLC		Last	Run	Level	VLC
0	1	MPEG-4 Motion	0	0	1	10s
+0.5	010	Vector Difference	0	1	1	110s
- 0.5	011	(MVD) VLC's	0	2	1	1110s
+1	0010		0	0	2	1111s
- 1	0011		1	0	1	0111s
+2	0000110		0	3	1	01101s
- 2	0000111		0	4	1	01100s
.....		MPEG-4 Visual	0	5	1	01011s
		Transform	0	0	3	010101s
		Coefficient	0	1	2	010100s
		(TCOEF) VLC's	0	6	1	010011s
					

3. Sistemas estándar de compresión

¿Cómo funciona la **compresión de imágenes** (tipo JPEG)?

3. Sistemas estándar: modelo espacial

- Codificación entrópica: poco eficiente para imágenes y vídeo
- Necesario incluir un *modelo espacial* para representar la señal de imagen/vídeo apropiada para la codificación entrópica



- Modelo espacial aprovecha la *redundancia subjetiva*:
 - Zonas de la imagen cuyos píxeles con alta correlación
 - Menor sensibilidad a las “altas frecuencias”
 - Eliminar elementos de la imagen o el vídeo sin afectar a la percepción visual
 - Fases: *Transformación, Cuantificación y Reordenamiento*

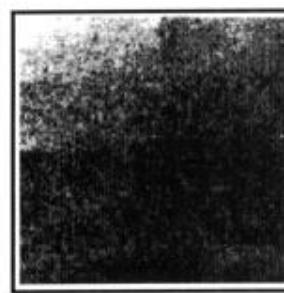


3. Sistemas estándar: modelo espacial

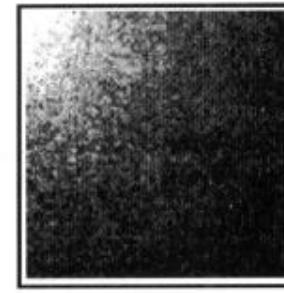
- **Transformación: *DCT (Discrete Cosine Transform)***
 - i. Datos decorrelados en el dominio de la transformada y compactos (concentración de la energía en pocos valores)
 - ii. Reversible
 - iii. Computacionalmente tratable



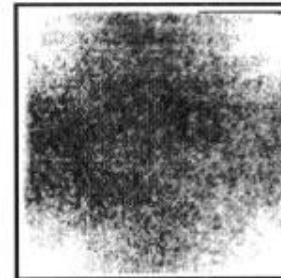
(a) Test Image



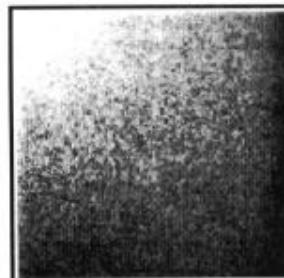
(c) Discrete Cosine transform



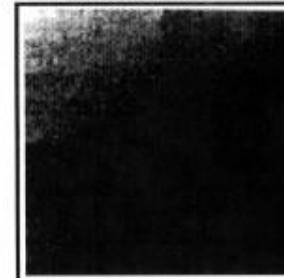
(d) Discrete Sine transform



(b) Discrete Fourier transform

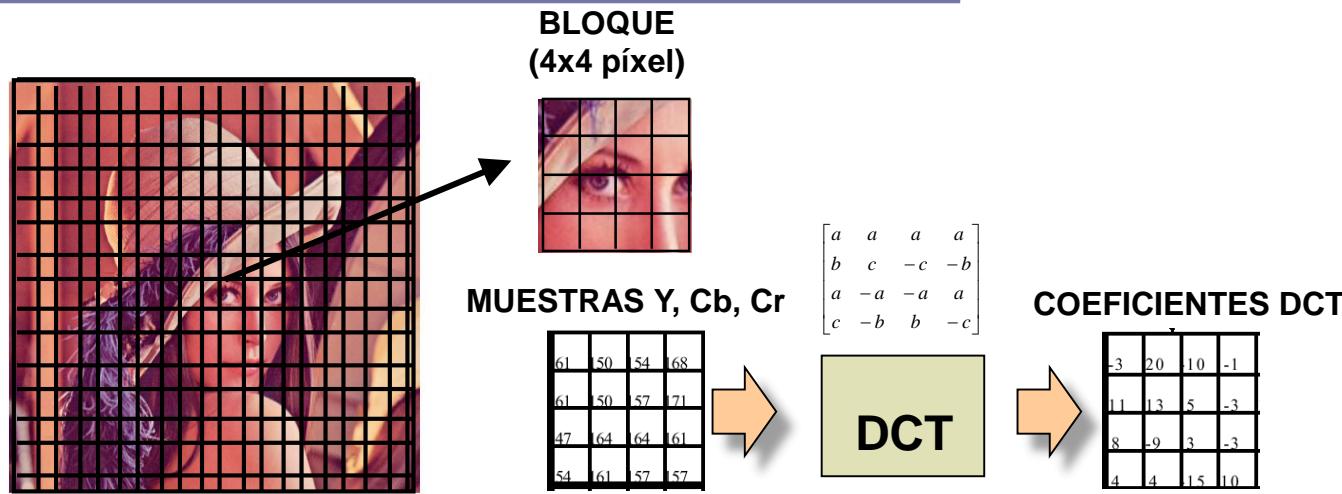


(e) Discrete Hadamard transform



(f) Karhunen-Loeve transform

3. Sistemas estándar: modelo espacial



$$X \quad Y = AXA^T$$

X Matriz $N \times N$ de muestras

Y Matriz $N \times N$ de coeficientes DCT

A Matriz $N \times N$ de transformación

$$A_{ij} = C_i \cos \frac{(2j+1)i\pi}{2N}$$

$$C_i = \sqrt{\frac{1}{N}} (i=0)$$

$$C_i = \sqrt{\frac{2}{N}} (i>0)$$

$$Y_{xy} = C_x C_y \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{ij} \cos \frac{(2j+1)y\pi}{2N} \cos \frac{(2i+1)x\pi}{2N}$$

$$X_{ij} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C_x C_y Y_{xy} \cos \frac{(2j+1)y\pi}{2N} \cos \frac{(2i+1)x\pi}{2N}$$

3. Sistemas estándar: modelo espacial

Ejemplo: Cálculo de la DCT de un bloque 4x4

$$X = \begin{bmatrix} 5 & 11 & 8 & 10 \\ 9 & 8 & 4 & 12 \\ 1 & 10 & 11 & 4 \\ 19 & 6 & 15 & 7 \end{bmatrix} \quad A = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.653 & 0.271 & -0.271 & -0.653 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.271 & -0.653 & 0.653 & -0.271 \end{bmatrix}$$

$$Y = AXA^T = \begin{bmatrix} 35.0 & -0.079 & -1.5 & 1.115 \\ -3.299 & -4.768 & 0.443 & -9.010 \\ 5.5 & 3.029 & 2 & 4.699 \\ -4.045 & -3.010 & -9.384 & -1.232 \end{bmatrix}$$

3. Sistemas estándar: modelo espacial

Ejemplo: Reconstrucción (IDCT) de un bloque 4x4

126	159	178	181
98	151	181	181
80	137	176	156
75	114	88	68

Bloque original

537.25	-76.00	-54.75	-7.75
106.05	-34.99	12.71	6.14
-42.75	46.53	10.25	-9.80
20.20	-12.85	-3.91	8.49

Coeficientes DCT

134	134	134	134
134	134	134	134
134	134	134	134
134	134	134	134

Reconstrucción (IDCT) 1 coef.

169	169	169	169
149	149	149	149
120	120	120	120
100	100	100	100

Reconstrucción (IDCT) 2 coef.

144	159	179	194
124	138	159	173
95	110	130	145
75	89	110	124

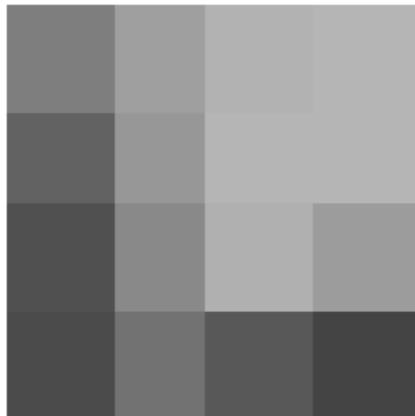
Reconstrucción (IDCT) 3 coef.

146	179	187	165
95	146	179	175
66	117	150	146
76	109	117	96

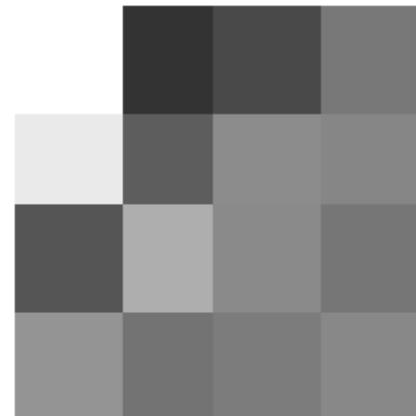
Reconstrucción (IDCT) 5 coef.

3. Sistemas estándar: modelo espacial

Ejemplo: Reconstrucción (IDCT) de un bloque 4x4



Bloque original



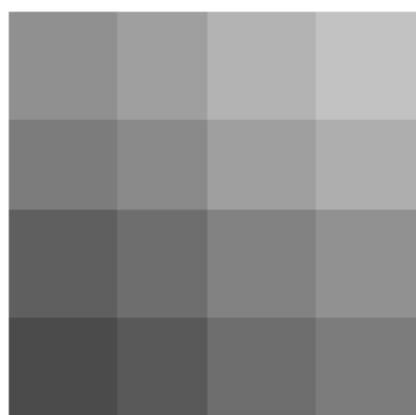
Coeficientes DCT



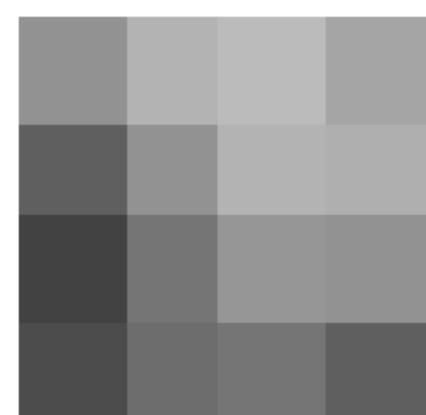
Reconstrucción (IDCT) 1 coef.



Reconstrucción (IDCT) 2 coef.



Reconstrucción (IDCT) 3 coef.



Reconstrucción (IDCT) 5 coef.

3. Sistemas estándar: modelo espacial

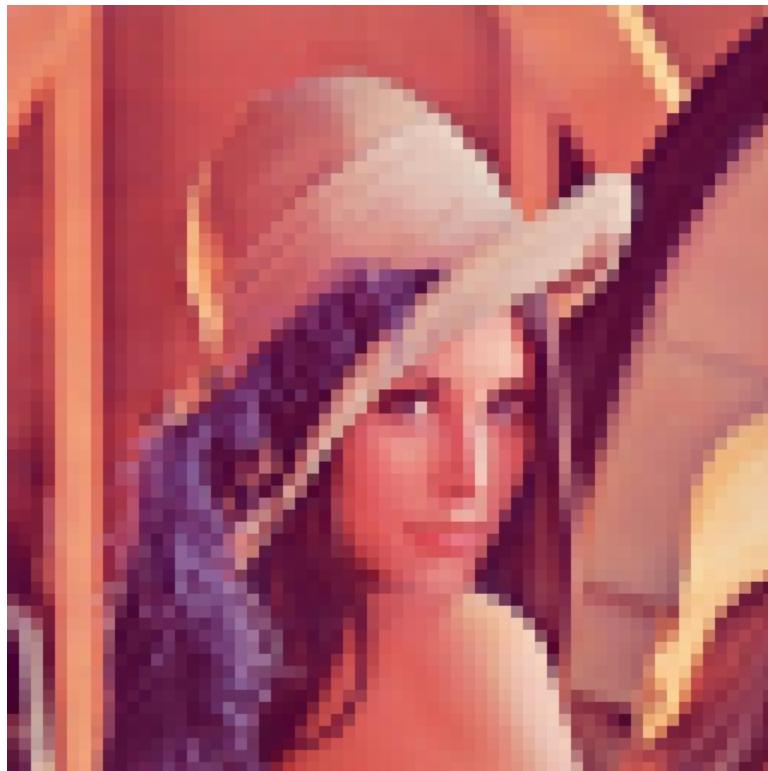


Imagen reconstruida con
1 coef/bloque

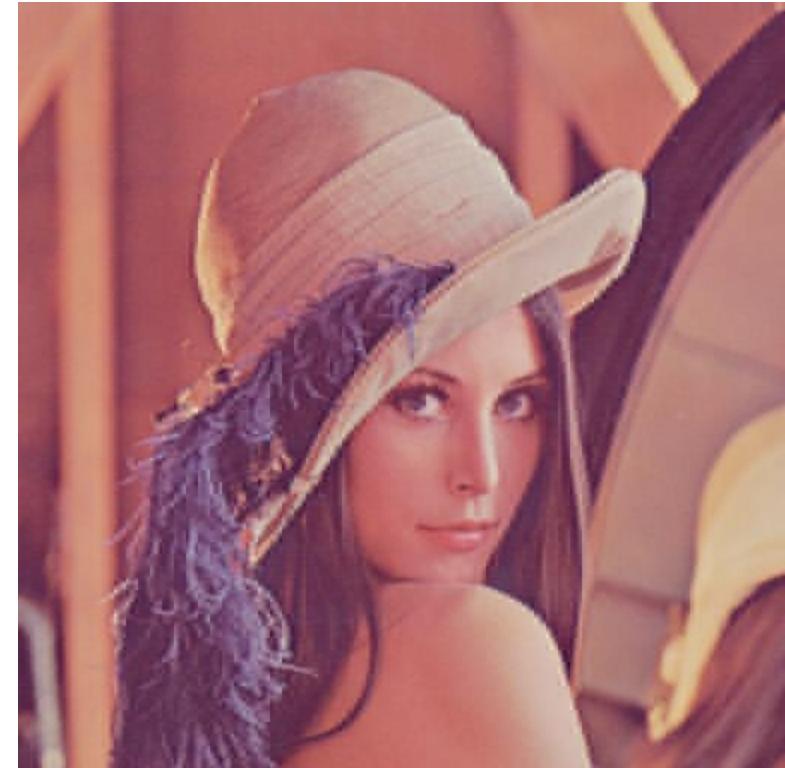


Imagen reconstruida con
8 coef/bloque

3. Sistemas estándar de compresión

¿Qué significa el **parámetro QP**?

3. Sistemas estándar: modelo espacial

- **Cuantificación: *Cuantificación escalar***

- Establecer una cuantificación de los coeficientes DCT (Q_{coef})
- Parámetro: QP o “*quantizer*”
- Proceso con pérdidas (irreversible) (R_{coef})
- Valores altos de QP obtiene alta compresión (baja calidad)
- Valores bajos de QP obtiene baja compresión (alta calidad)

$$Y = AXA^T = \begin{bmatrix} 35.0 & -0.079 & -1.5 & 1.115 \\ -3.299 & -4.768 & 0.443 & -9.010 \\ 5.5 & 3.029 & 2 & 4.699 \\ -4.045 & -3.010 & -9.384 & -1.232 \end{bmatrix}$$

$$Q_{coef} = round(coef / QP) \quad QP=[1 \dots 51] \quad R_{coef} = QP * Q_{coef}$$

$$\text{Ejem.: } QP = 12$$

$$Q_{coef} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad R_{coef} = \begin{bmatrix} 36 & 0 & 0 & 0 \\ 0 & 0 & 0 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -12 & 0 \end{bmatrix}$$

3. Sistemas estándar: modelo espacial

- **Reordenamiento: Zigzag**

- Los coeficientes de mayor valor están típicamente en las posiciones de “baja frecuencia, DC” (alrededor de la posición (0,0))
- El reordenamiento más eficiente para codificaciones “típicas” de vídeo es en zigzag empezando por coeficiente DC

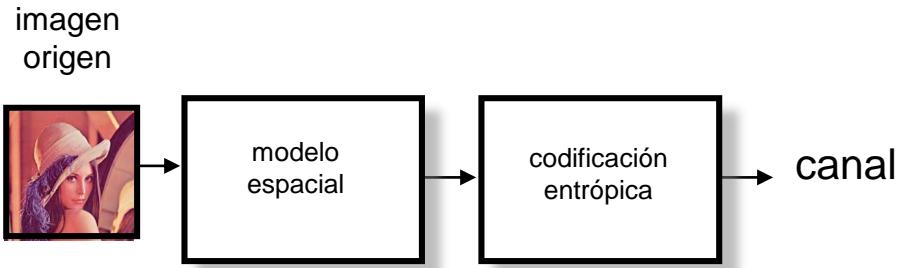
3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

COEFICIENTES DCT CUANTIFICADOS

Serie de datos: 3,0,0,0,0,0,0,0,0,0,-1,0,-1,0

Datos codificados: (0,0,3), (0,11,-1),(1,1,-1)
MPEG-4 Visual Transform Coefficient (TCOEF) VLC's

$$Q_{coef} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

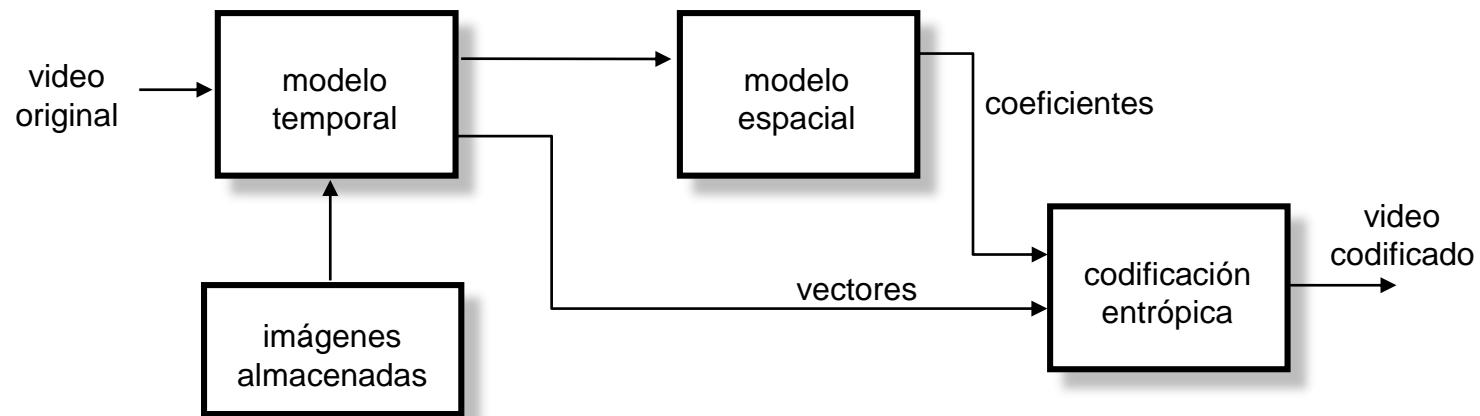


3. Sistemas estándar de compresión

¿Qué es y cómo se aprovecha la
redundancia temporal?

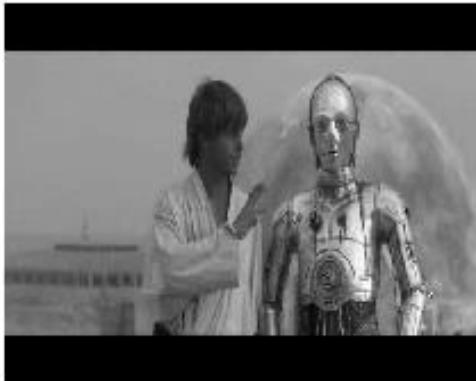
3. Sistemas estándar: modelo temporal

- Imágenes consecutivas del vídeo son muy parecidas
- Necesario incluir un *modelo temporal* para aprovechar la *redundancia temporal*



3. Sistemas estándar: modelo temporal

- **Redundancia temporal: energía residual**



F_{n-1} reconstruida



F_n

- Reducción de los datos a codificar si obtenemos la diferencia entre F_n y F_{n-1} reconstruida (*imagen residual*)
- Mayor reducción utilizando estimación y compensación de movimiento

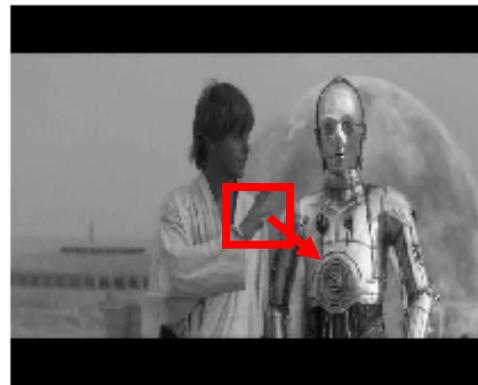
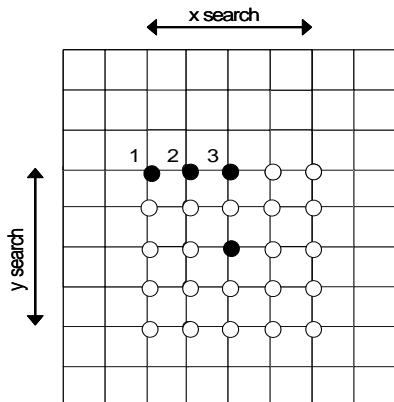


Imagen residual

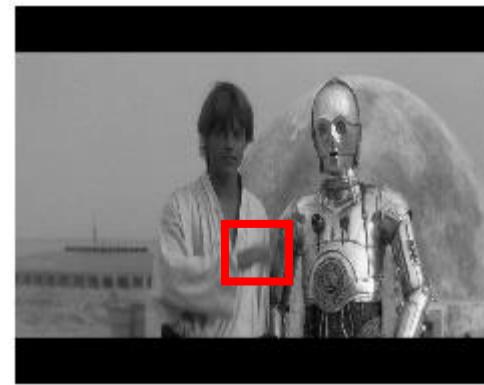
3. Sistemas estándar: modelo temporal

- **Estimación del movimiento**

- Utilizar como unidad el “macrobloque” (bloques de 4x4, 8x8 o 16x16 pixeles)
- Comparar el “macrobloque” a codificar del frame actual (F_n) con macrobloques vecinos del frame anterior (F_{n-1}) y obtener el “macrobloque de referencia”
→ **algoritmo de búsqueda y comparación de energía residual**
- Calcular el “offset” entre el “macrobloque” a codificar y el de referencia
→ **vector de movimiento**



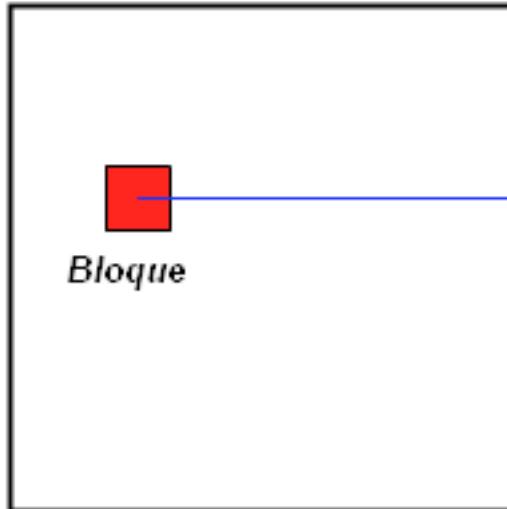
F_{n-1} reconstruída



F_n

3. Sistemas estándar: modelo temporal

Fotograma k (actual)



Fotograma k-1 (referencia)

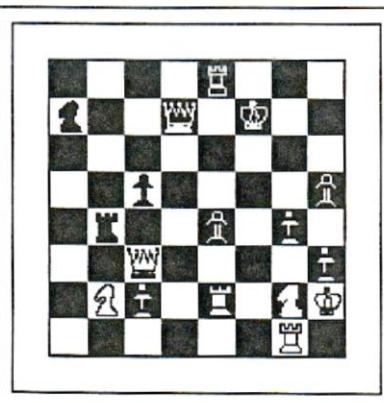
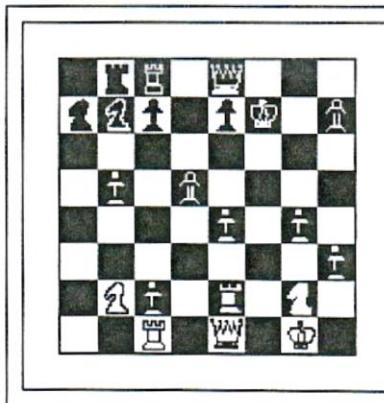
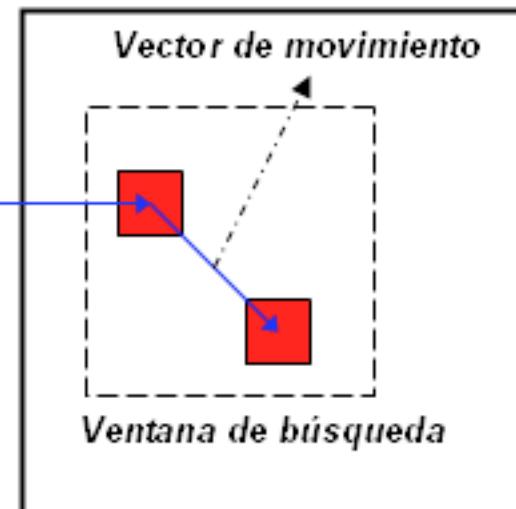


Fig. 3.a. Fotograma anterior / Fig 3.b. Fotograma actual

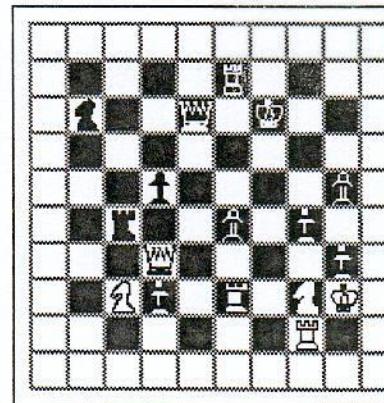


Fig. 3.c.

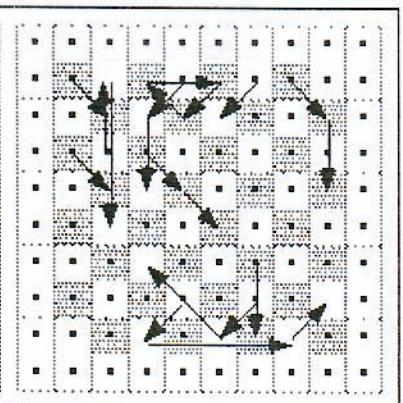


Fig. 3.d.

3. Sistemas estándar: modelo temporal

Frame 1 $s[x,y,t-1]$ (previous)



Frame 2 $s[x,y,t]$ (current)

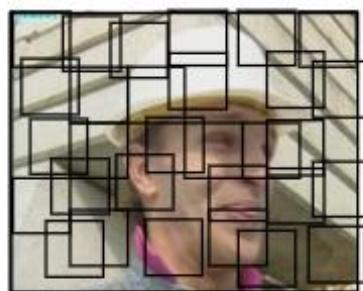


Partition of frame 2 into blocks
(schematic)



Size of Blocks

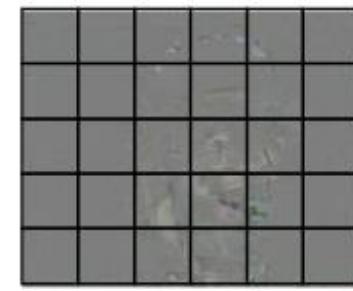
Accuracy of Motion Vectors



Referenced blocks in frame 1



Frame 2 with
displacement vectors



Difference between motion-compensated prediction and current frame $u[x,y,t]$

3. Sistemas estándar: modelo temporal

Error Measures for Block-matching

- Mean squared error (sum of squared errors)

$$SSD(d_x, d_y) = \sum_{y=1}^{By} \sum_{x=1}^{Bx} [s(x, y, t) - s'(x - d_x, y - d_y, t - \Delta t)]^2$$

- Sum of absolute differences

$$SAD(d_x, d_y) = \sum_{y=1}^{By} \sum_{x=1}^{Bx} |s(x, y, t) - s'(x - d_x, y - d_y, t - \Delta t)|$$

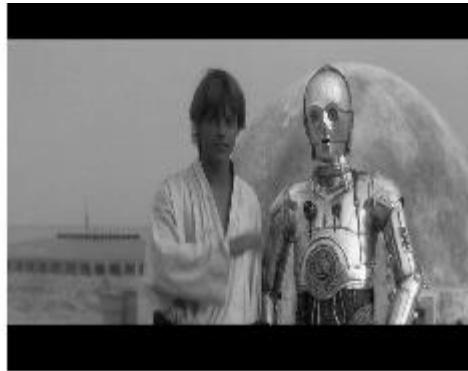
- Approximately same performance
→ SAD less complex for some architectures

3. Sistemas estándar: modelo temporal

- **Compensación del movimiento**



F_{n-1} reconstruida con compensación de movimiento



F_n

- Obtener un frame de referencia a partir de F_{n-1} reconstruida y la información de la estimación de movimiento (desplazamiento de macrobloques)
- Obtener la diferencia entre el frame actual (F_n) y el frame de referencia F_{n-1} con estimación de movimiento
- Codificar la diferencia y transmitir los vectores de movimiento
- Ventaja: mayor eficiencia de compresión



Imagen residual
(con compensación de movimiento)

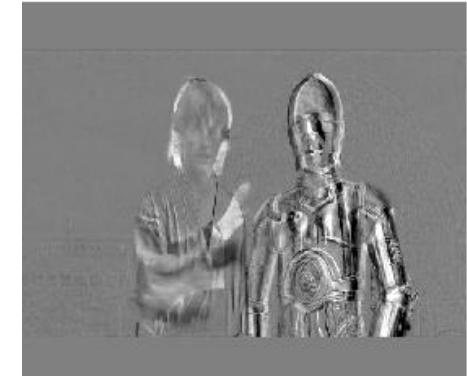


Imagen residual
(sin compensación de movimiento)

3. Sistemas estándar: modelo temporal



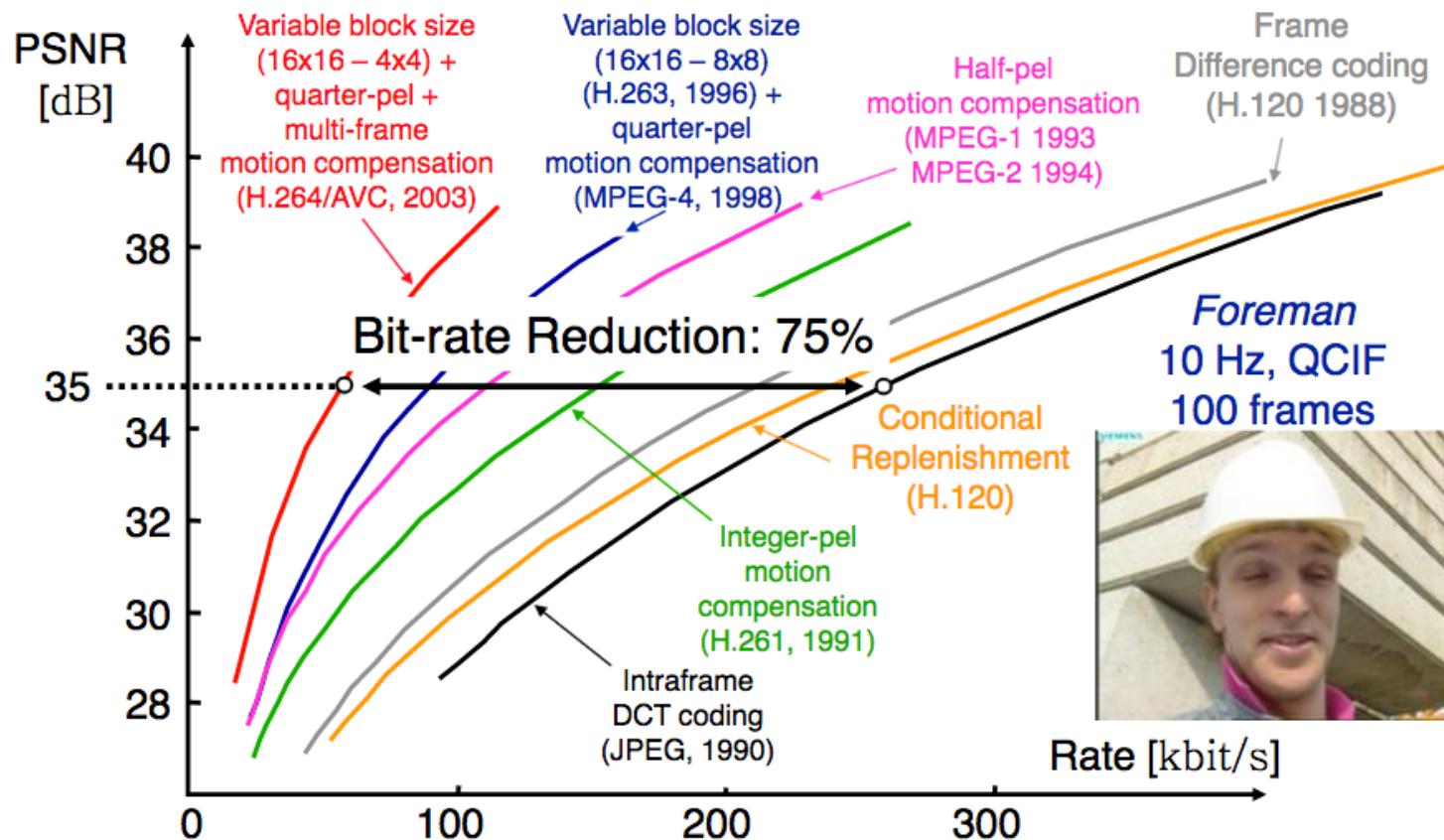
<http://people.xiph.org/~xiphmont/demo/index.html>

<http://people.xiph.org/~xiphmont/demo/theora/matrix-19-theora-telemetry.ogg>

3. Sistemas estándar de compresión

¿Cuál es el **modelo** (entrópico, espacial, temporal) qué **más aporta a la compresión?**

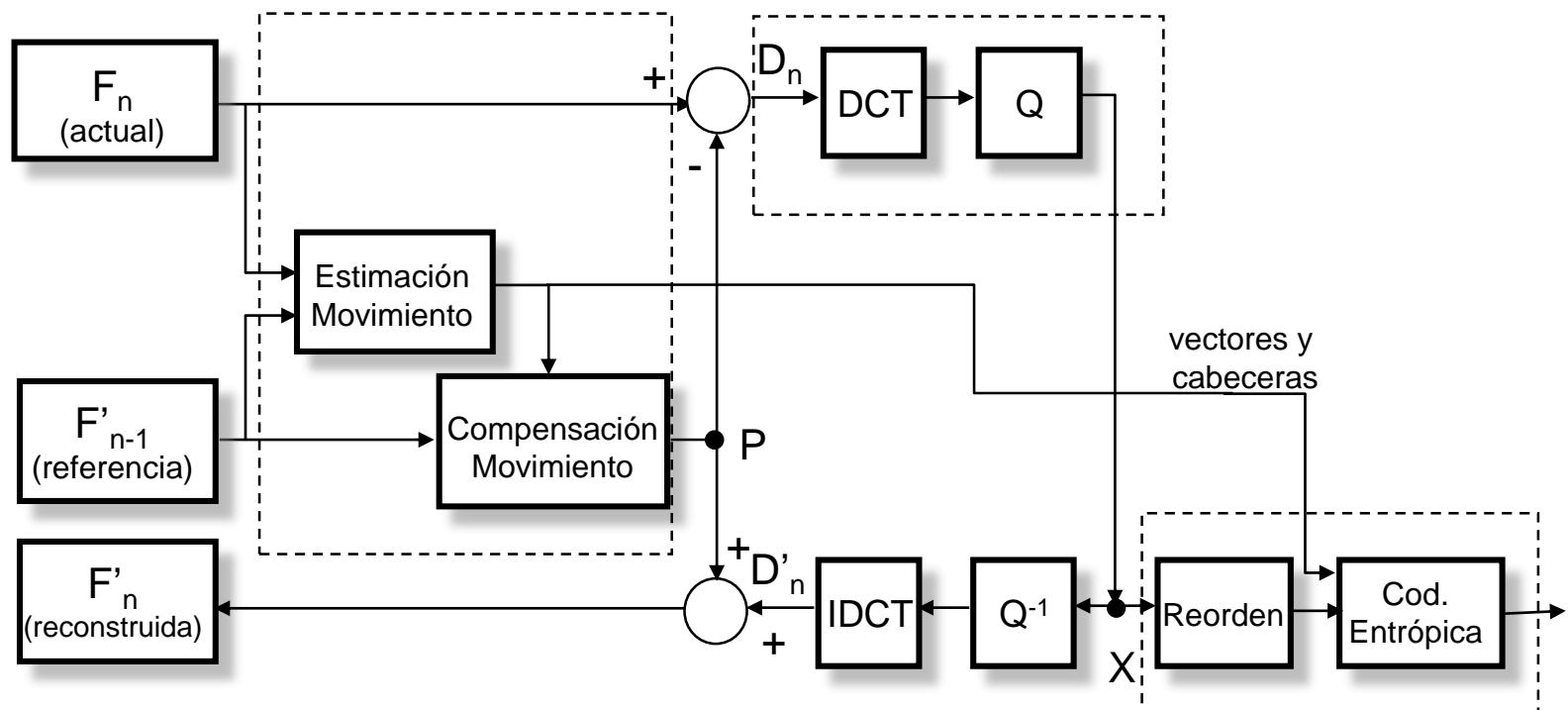
3. Sistemas estándar: modelo temporal



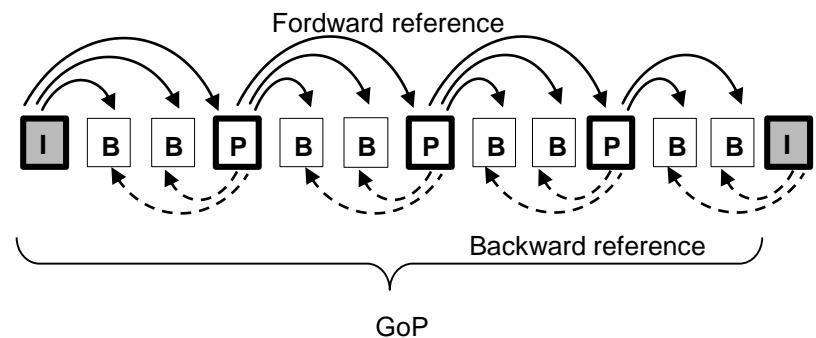
3. Sistemas estándar de compresión

¿Todos los **frames** generados por el sistema de compresión son del mismo **tipo**?

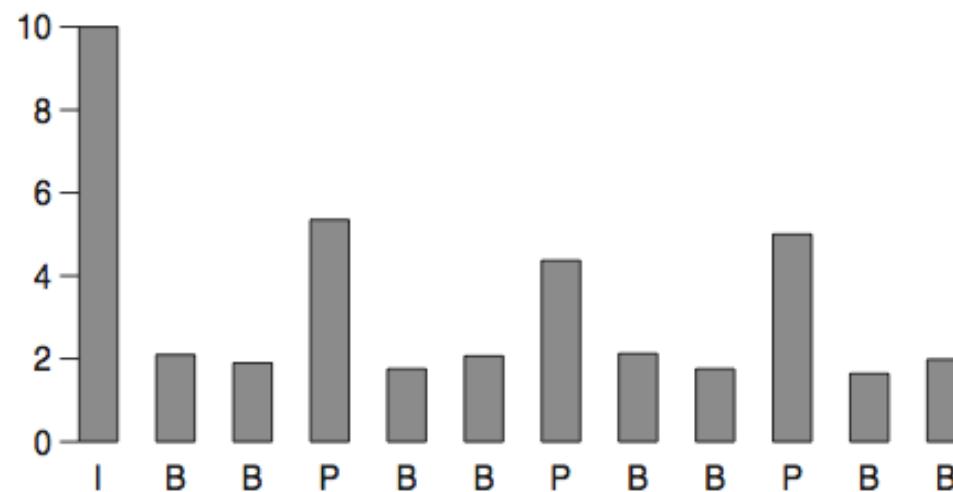
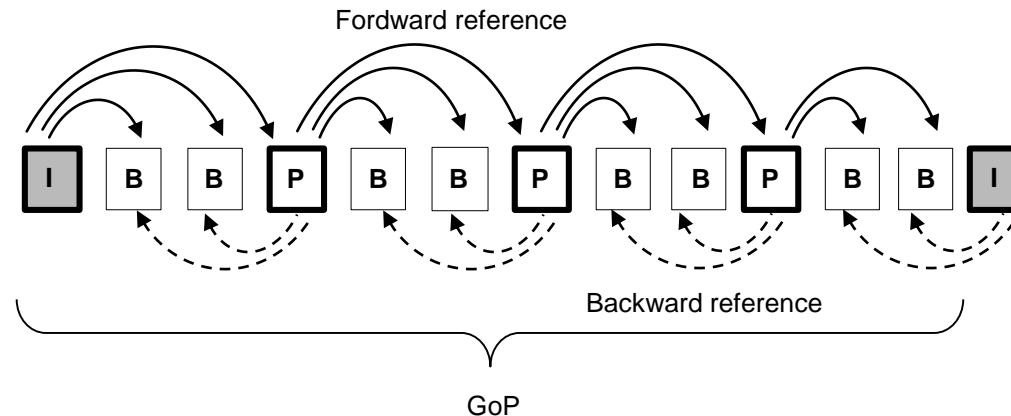
3. Sistemas estándar: diagrama de bloques



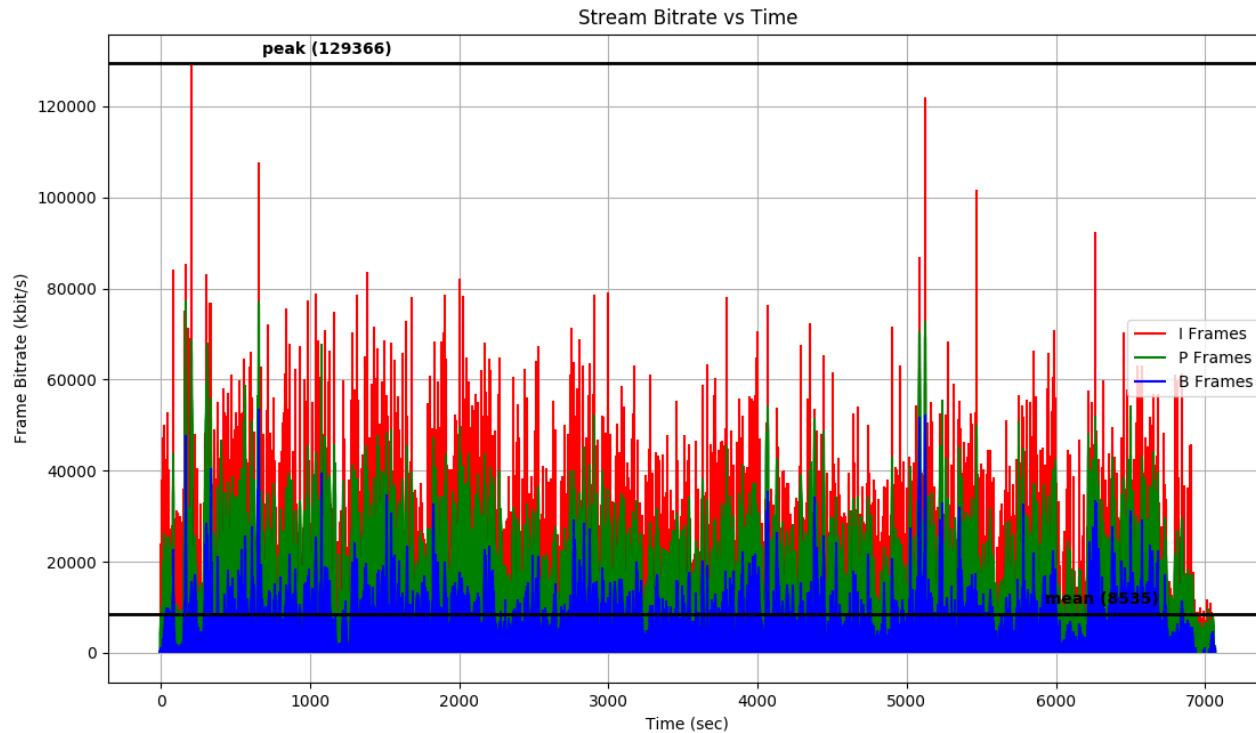
Vídeo ENCODER



3. Sistemas estándar: diagrama de bloques



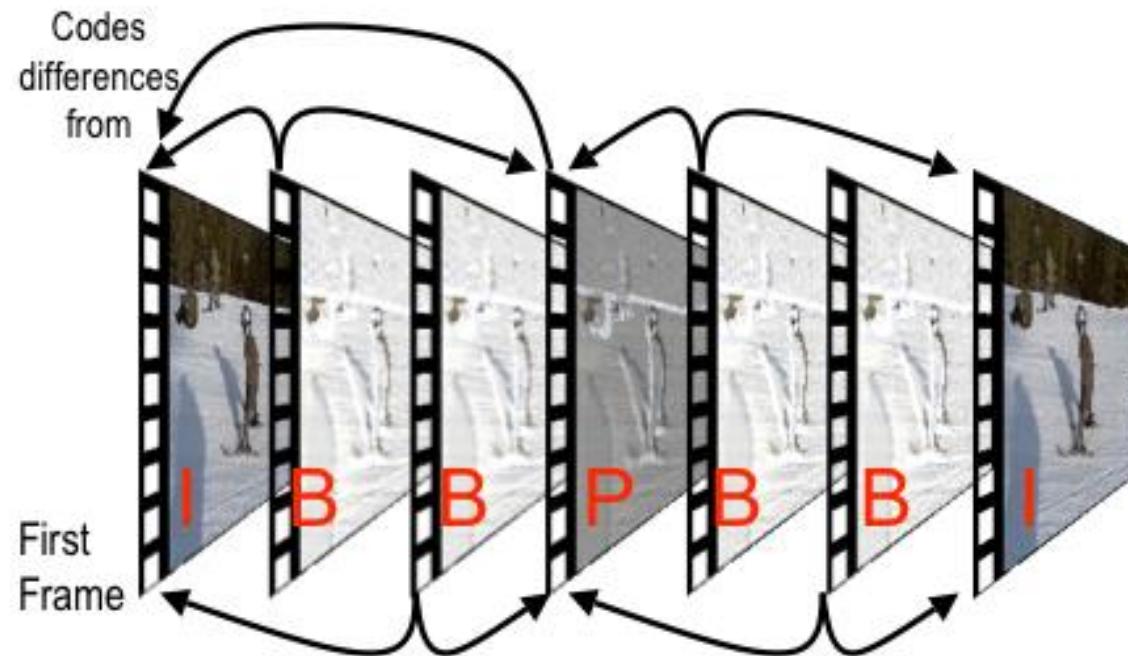
3. Sistemas estándar: diagrama de bloques



3. Sistemas estándar: diagrama de bloques

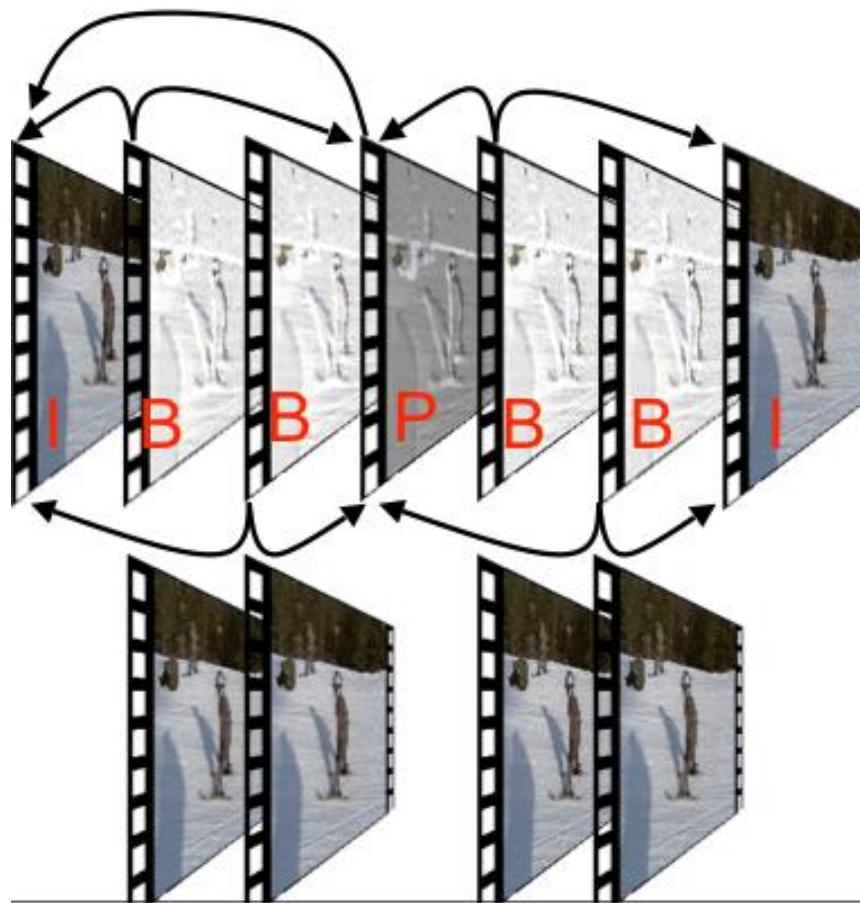
Frame Ordering

- Up to encoder to choose I, P, B frame ordering.
- Eg IBBPBBIBBPBBPI...



3. Sistemas estándar: diagrama de bloques

Encoding Order



1. Encode I-frame 1
2. Store frame 2
3. Store frame 3
4. Encode P frame 4
5. Encode B frame 2
6. Encode B frame 3
7. Store frame 5
8. Store frame 6
9. Encode I frame 7
10. Encode B frame 5
11. Encode B frame 6

3. Sistemas estándar: diagrama de bloques

Transmission Order

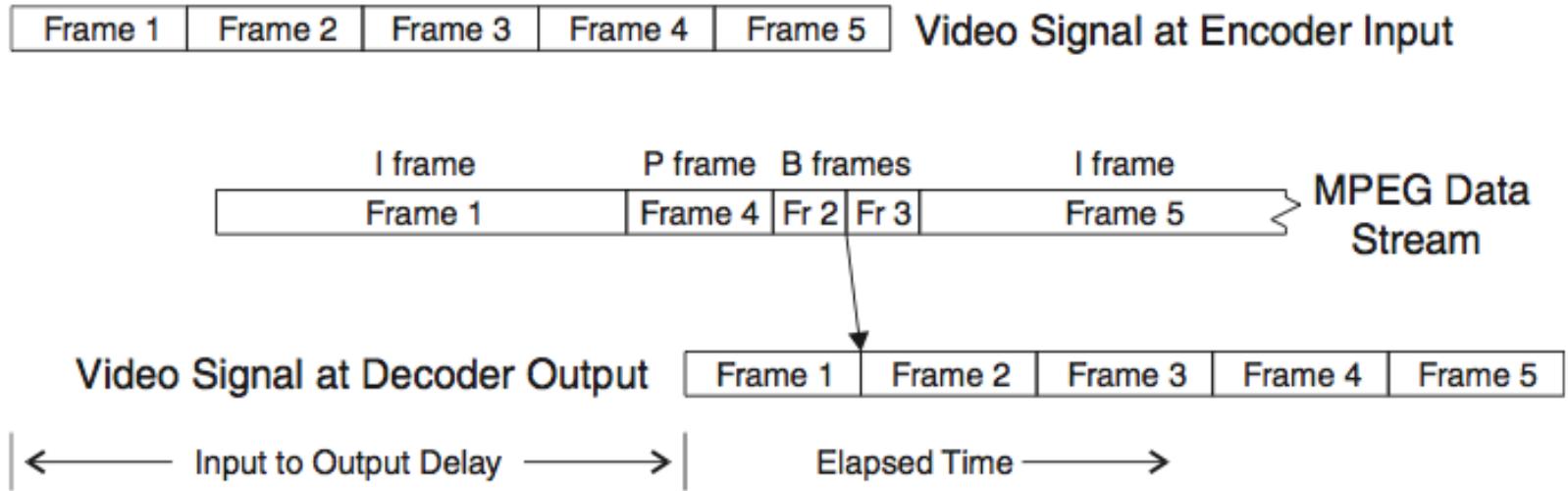
- Frames are encoded out of order
- Need to be decoded in the order they're encoded.
 - Common to send out of order.

Eg: I₁B₂B₃P₄B₅B₆I₇B₈B₉P₁₀B₁₁B₁₂I₁₄
sent in the order

I₁P₄B₂B₃I₇B₅B₆P₁₀B₈B₉I₁₄B₁₁B₁₂

- Allows decoder to decode as data arrives, although it still has to hold decoded frames until it has decoded prior B frames before playing them out.

3. Sistemas estándar: diagrama de bloques



3. Sistemas estándar: diagrama de bloques

B-frame disadvantages

- Computational complexity.
 - More motion search, need to decide whether or not to average.
- Increase in memory bandwidth.
 - Extra picture buffer needed.
 - Need to store frames and encode or playback out of order.
- Delay
 - Adds several frames delay at encoder waiting for need later frame.
 - Adds several frames delay at decoder holding decoded I/P frame, while decoding and playing prior B-frames that depend on it.

3. Sistemas estándar: diagrama de bloques

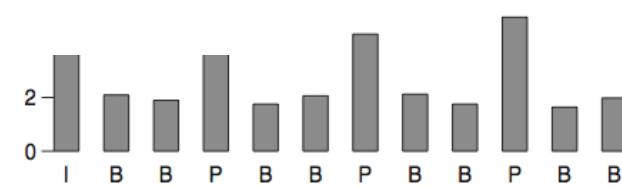
B-frame advantage

- B-frames increase compression.
- Typically use twice as many B frames as I+P frames.

Type	Size	Compression
I	18KB	7:1
P	6KB	20:1
B	2.5KB	50:1
Average	4.8KB	27:1

Typical values.

Really depends on video content.



3. Sistemas estándar: diagrama de bloques

Comparison of Long and Short GOPs

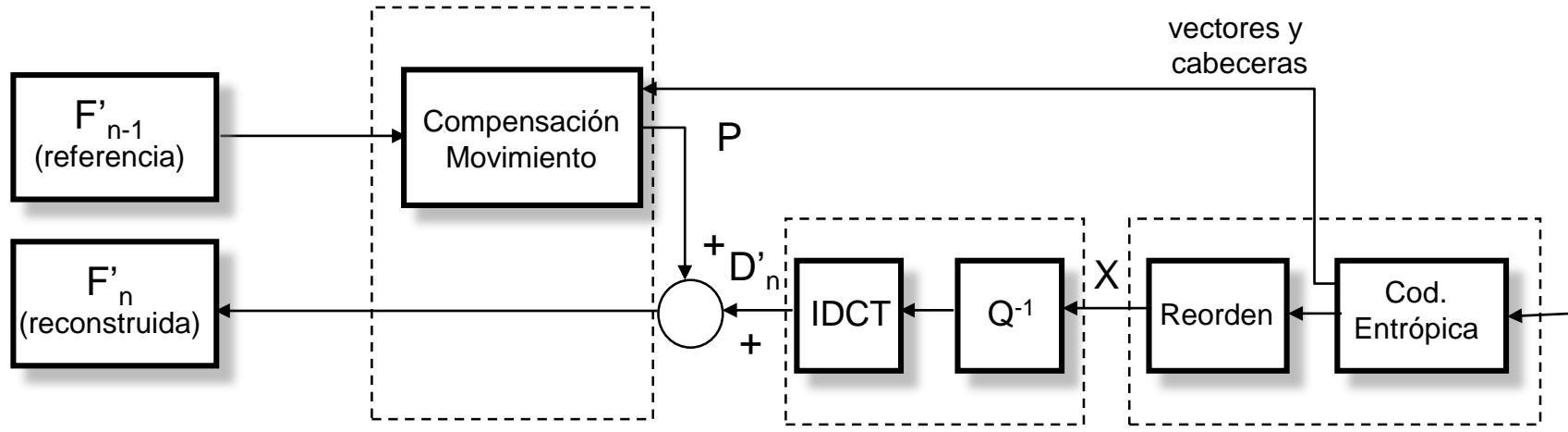
Long GOP Benefits

- As we have seen, I frames take more bits to encode accurately than P or B frames. When a long GOP is used, fewer I frames are in the stream, so the overall bit rate is reduced.
- For a network with a fixed bandwidth, more video signals can be carried when each uses a long GOP to reduce overall bandwidth demands. This can benefit both terrestrial and satellite-based networks.

Short GOP Benefits

- When video images have lots of rapid scene changes (such as in action movies and music videos), a short GOP structure may create a better video image.
 - In compressed video delivery systems (such as DSL and digital cable), short GOPs allow quicker channel-change times, since it is easier for the decoder to find an I frame to begin the decoding process for a new channel.
 - With a short GOP, I frames occur more often, so any accumulated errors are cleared out more rapidly.
 - When streams are edited, it is easier to make the cuts only at an I frame. With short GOP signals, I frames occur more often, so editing is easier and more precise.
 - With noisy video sources, prediction errors can rapidly accumulate. Shorter GOP signals are less sensitive to noise, because few consecutive frames are based on predicted values
-

3. Sistemas estándar: diagrama de bloques

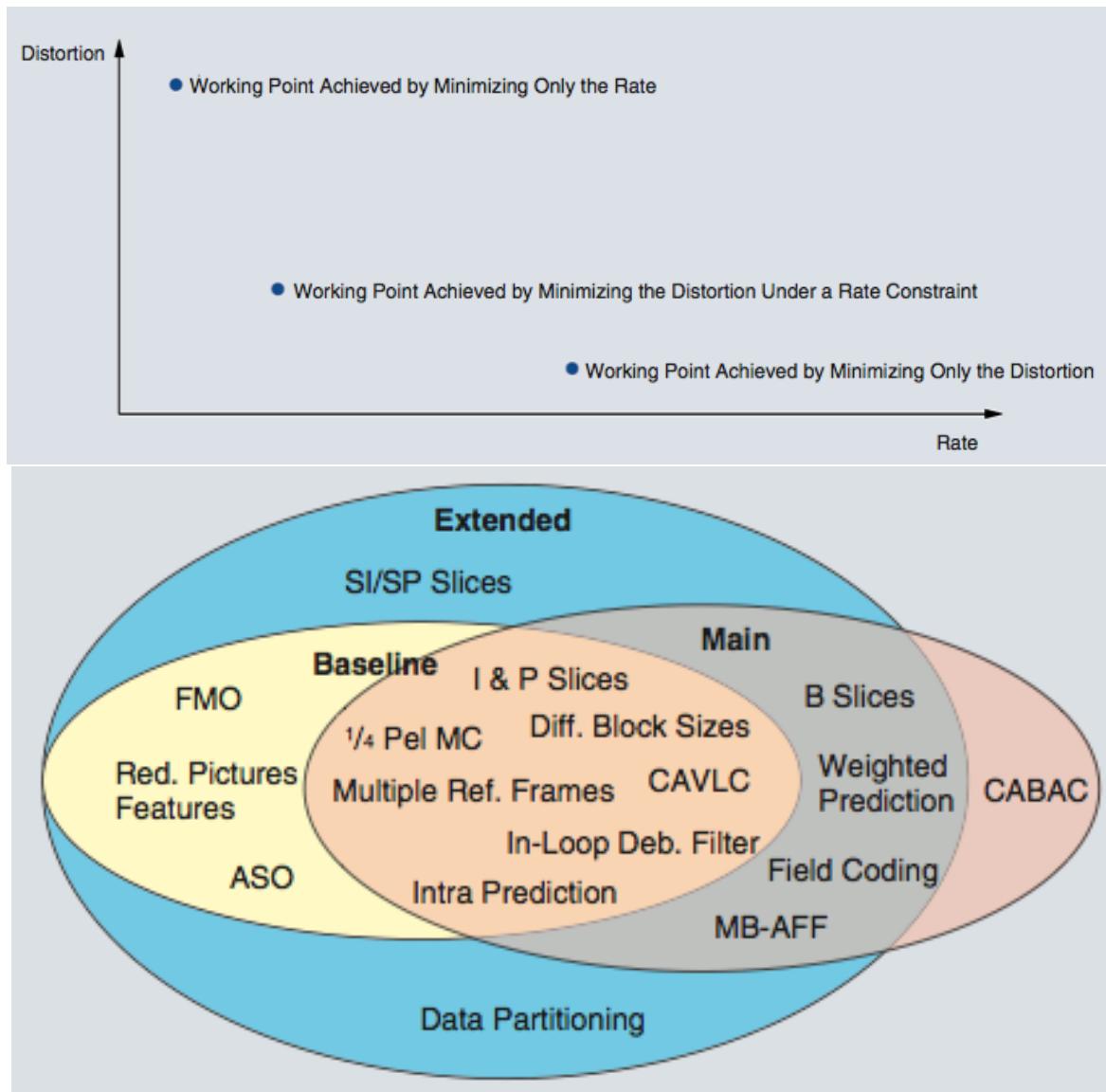


Vídeo DECODER

3. Sistemas estándar: estándares

Característica	MPEG-1	MPEG-2	MPEG-4	H.264/MPEG-4 Part 10/AVC
Tamaño del macro-bloque	16x16	16x16 (frame mode) 16x8 (field mode)	16x16	16x16
Tamaño del bloque	8x8	8x8	16x16 8x8, 16x8	8x8, 16x8, 8x16, 16x16, 4x8, 8x4, 4x4
Transformada	DCT	DCT	DCT/DWT	4x4 Integer transform
Tamaño de la muestra para aplicar la transformada	8x8	8x8	8x8	4x4
Codificación	VLC	VLC	VLC	VLC, CAVLC, CABAC
Estimación y compensación de movimiento	Si	Si	Si	Si, con hasta 16 MV
Perfiles	No	5 perfiles, varios niveles en cada perfil	8 perfiles, varios niveles en cada perfil	3 perfiles, varios niveles en cada perfil
Tipo de cuadros	I,P,B,D	I,P,B	I,P,B	I,P,B,SI,SP
Ancho de banda	Hasta 1.5 Mbps	2 a 15 Mbps	64 kbps a 2 Mbps	64 kbps a 150 Mbps
Complejidad del codificador	Baja	Media	Media	Alta
Compatibilidad con estándares previos	Si	Si	Si	No

3. Sistemas estándar: estándares



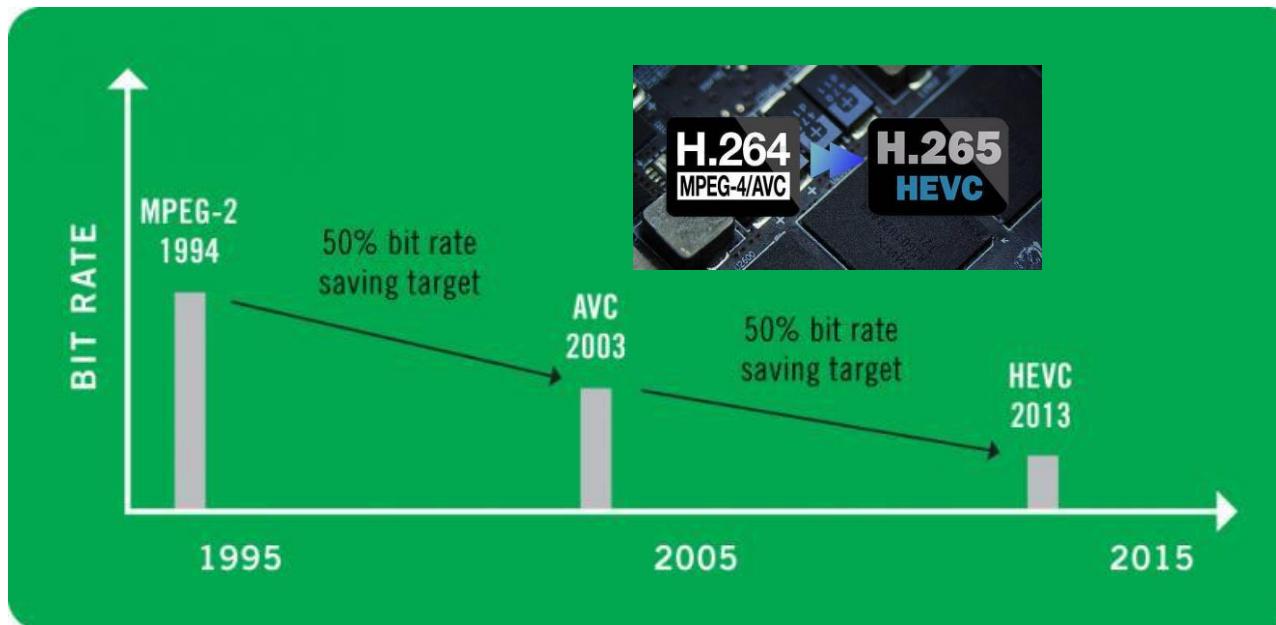
3. Sistemas estándar: estándares

Criteria \ Technique	HEVC	H.264/AVC
Year of the Implementation	2013	2003
Publisher	ISO/IEC and ITU	ISO/IEC and ITU
Name	MPEG-H, HEVC, Part 2	MPEG 4 Part 10 AVC
Application Area	Ultra High Definition video with up to 8k x 4k resolution	Blu-ray DVD, internet video streaming and High Definition video
Key Improvement	Reducing the video bit rate by 40-50% compared with H.264/AVC with the same video quality.	Reducing the video bit rate by 40-50% compared with MPEG-2 with the same video quality.
Progression	Successor H.264/AVC	Successor to MPEG-2
Resolutions	Supports Up to 8K	Support Up to 4K
Number of frames/second	up to 300 fps	up to 59.94 fps
Profiles	3 profiles; 13 levels	21 profiles ; 17 levels
Picture Partitioning	Code Tree Unit (CTU) with size 64x64	Macroblock with size 16x16
Block Partitioning	Partitioned to quadtree Prediction Unit (PU) with size	Partitioned to sub-blocks With size 4x4

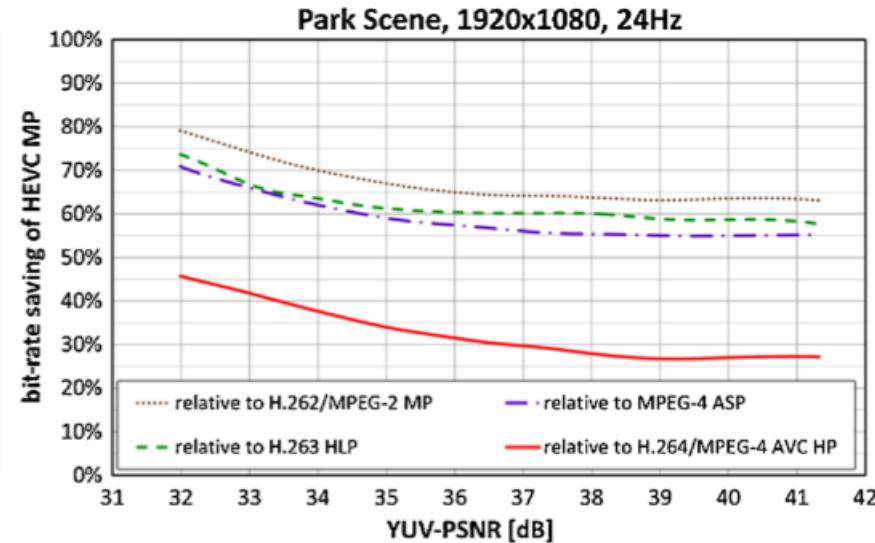
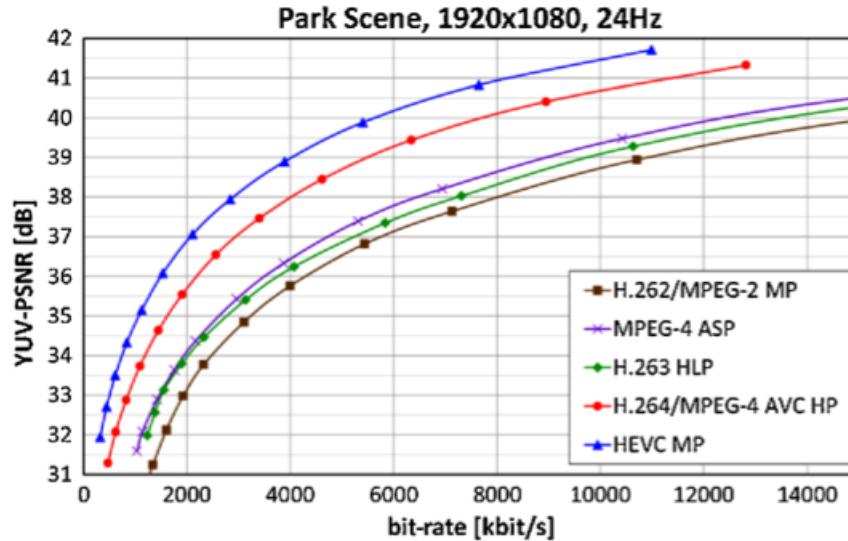
3. Sistemas estándar de compresión

¿Cuál es la **mejora** de cada
nuevo estándar?

3. Sistemas estándar: estándares



3. Sistemas estándar: estándares



AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR ENTERTAINMENT APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/ H.262 MP
HEVC MP	35.4%	63.7%	65.1%	70.8%
H.264/MPEG-4 AVC HP	-	44.5%	46.6%	55.4%
MPEG-4 ASP	-	-	3.9%	19.7%
H.263 HLP	-	-	-	16.2%

Vídeo H.264/AVC HP HDTV 1080i: 1920 x1080 pixels y 25 fps

Formato: 4:2:0

Resolución Y: 1920x1080 muestras (8bits/muestra)

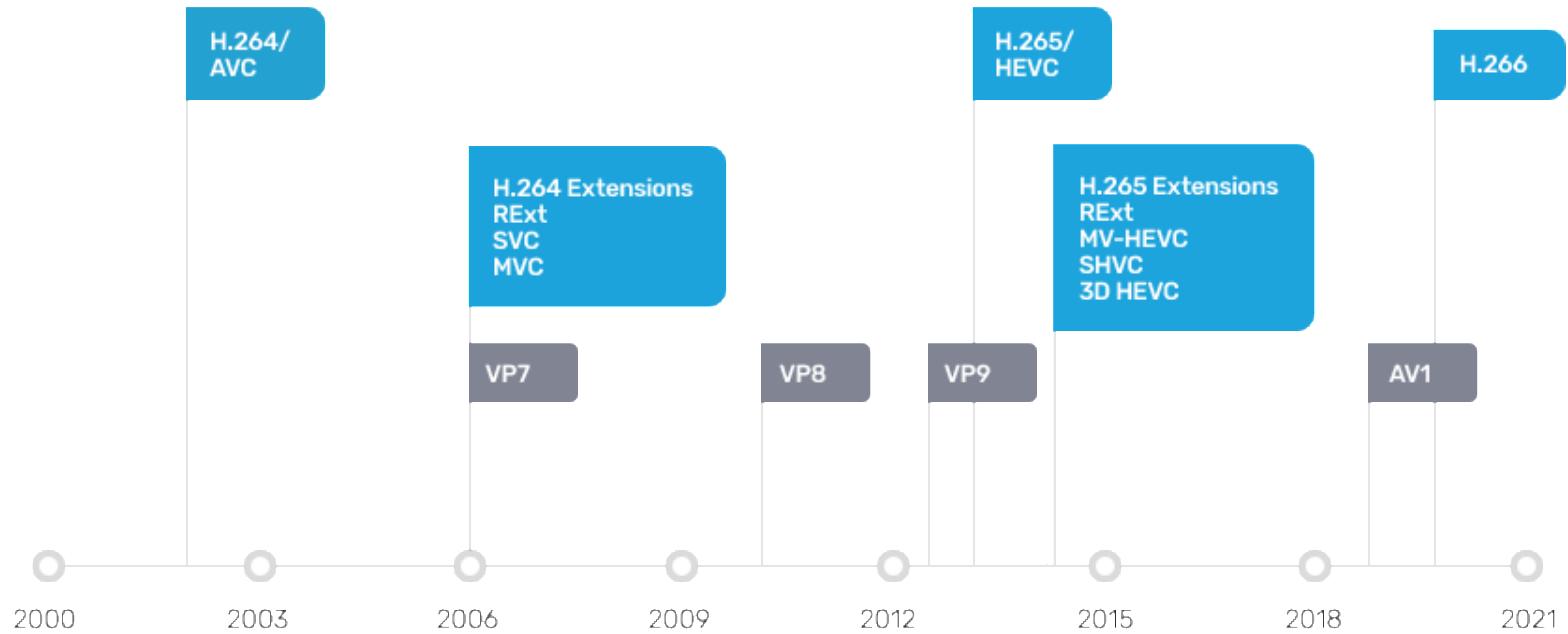
Resolución Cb, Cr: 960x540 muestras (8bits/muestra)

Bit Rate ≈ 12,5 Mbps

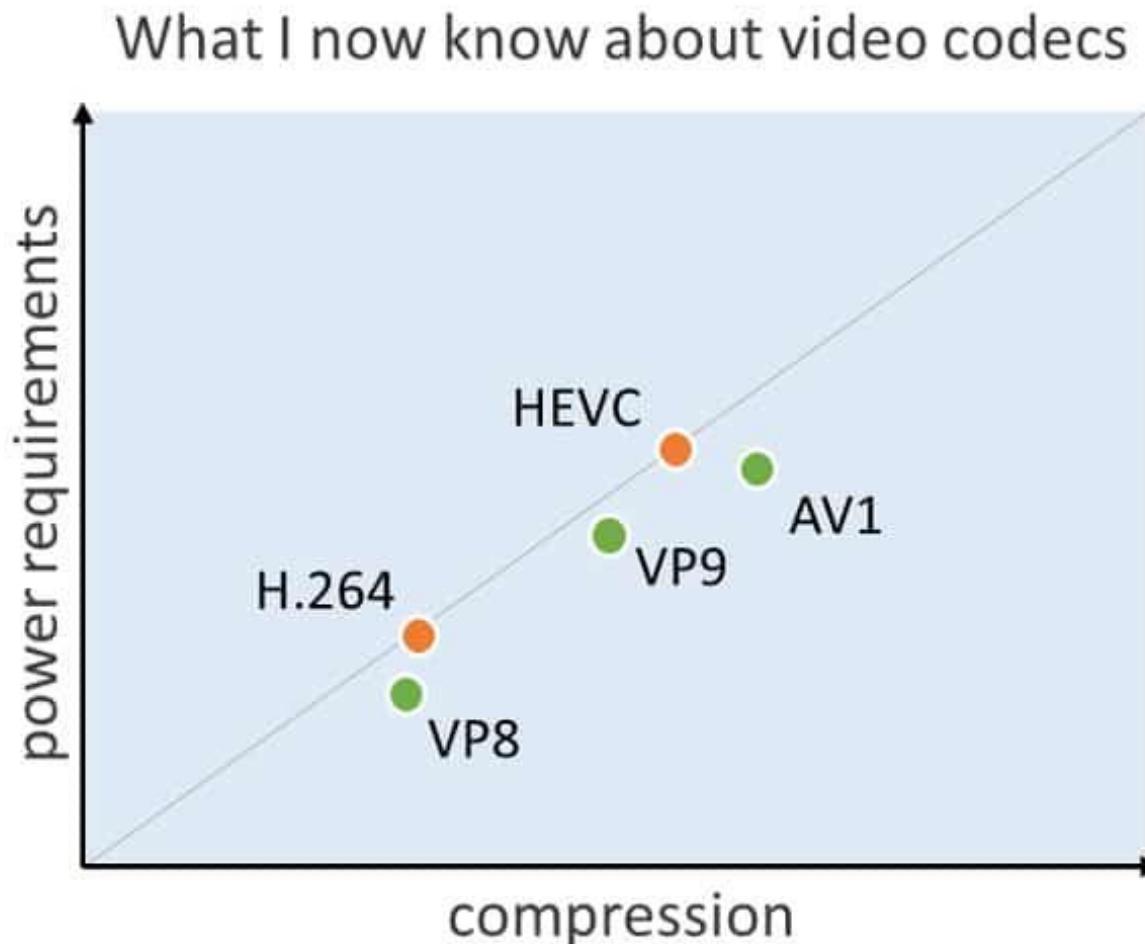
3. Sistemas estándar: estándares

● ITU-T & ISO/IEC

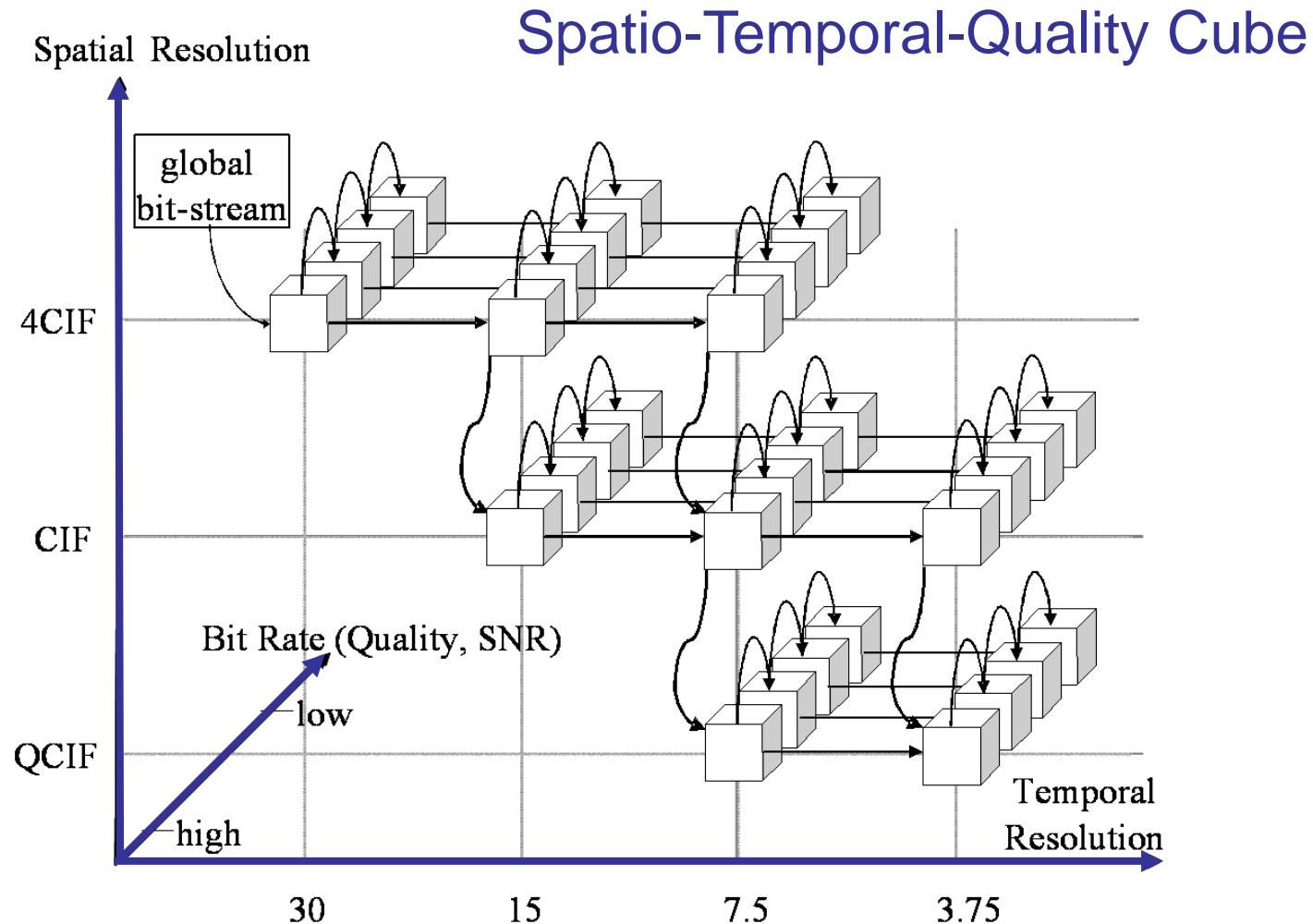
● OTHER



3. Sistemas estándar: estándares



3. Sistemas estándar: estándares



3. Sistemas estándar: contenedores

Un formato contenedor es un formato de archivo que puede contener varios tipos de datos, comprimidos mediante una serie de códecs. Permiten almacenar audio, video, subtítulos, capítulos y meta-datos (tags), junto con la información de sincronización requerida para la reproducción.

Los formatos contenedores mas utilizados son:

- **AVI** (contenedor estándar de Windows)
- **MOV** (contenedor estándar de QuickTime)
- **MPEGTS** (contenedor adoptado para la TV Digital)
- **MP4** (contenedor estándar para MPEG-4, H.264)
- **Ogg** (contenedor estándar de Xiph.org códecs)
- **Matroska** (estándar abierto, adoptado por DivX, Inc. para DivX Plus HD)
- **WebM** (estándar abierto, subconjunto de Matroska, codec VP8, VP9)
- **ASF** (contenedor estándar para Microsoft WMA y WMV)
- **RealMedia** (contenedor estándar para RealVideo y RealAudio)

Índice

Objetivos

Contenido

1. Introducción

2. Conceptos básicos de compresión de vídeo

3. Sistemas estándar de compresión de vídeo

4. Actividades y prácticas

5. Referencias bibliográficas

4. ACTIVIDADES Y PRÁCTICAS

4. Actividades y prácticas



A screenshot of a web browser displaying the website for the Telecommunication Networks Group (TKN) at the Technical University of Berlin (TU Berlin). The page title is "Institut für Telekommunikationssysteme: Telecommunication Networks Group". The header includes the TU Berlin logo and links for Contact, Impressum, Sitemap, Deutsch, Index A-Z, Mobil, and search. The main content area features a large image of a modern building complex, a "Welcome to TKN - the Telecommunication Networks Group" message, and a "Quick Access" sidebar with a search bar. The left sidebar lists categories such as News, Members, Teaching, Research, Publications, Soft/Hardware Components, Visitor Information, and Internals.

ASIGNATURA:
COMUNICACIONES MULTIMEDIA

PRÁCTICA:
COMPRESIÓN Y PERFIL DE VÍDEO COMPRIMIDO

OBJETIVO:

Analizar el perfil del tráfico de vídeo comprimido con H.264 y H.265 a diferentes calidades, y obtener información sobre el tamaño de los frames y bitrate del vídeo codificado, así como el PSNR resultante de la codificación.

DOCUMENTACIÓN:

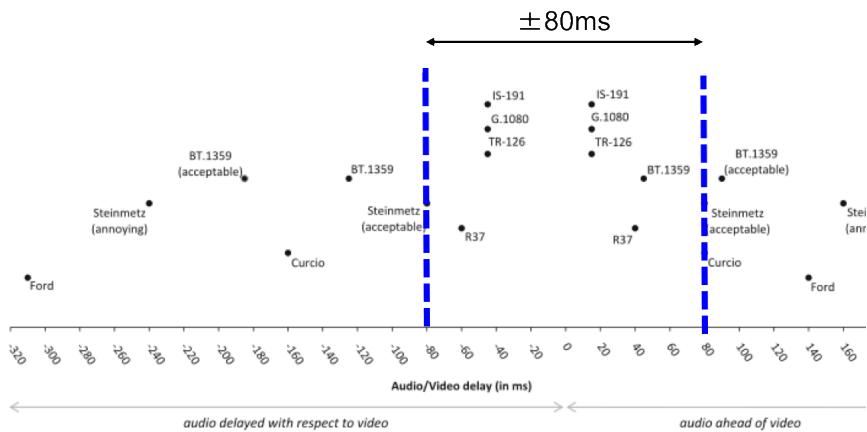
Contenidos de vídeo:
Ejemplos: www.blender.org (www.sintel.org)

Software para la codificación:
Ffmpeg (<http://ffmpeg.org>)

METODOLOGÍA:

- Trabajo Práctico individual (2h)
- Presentación de una memoria

4. Actividades y prá



ASIGNATURA:
COMUNICACIONES
MULTIMEDIA

ACTIVIDAD:
CUESTIONES SOBRE SINCRONIZACIÓN EN
SISTEMAS MULTIMEDIA

OBJETIVO:

Complementar aspectos vistos en teoría relacionados con la sincronización.

METODOLOGÍA:

- Trabajo Práctico (1h)
- Presentación de una memoria individual

Índice

Objetivos

Contenido

- 1. Introducción**
- 2. Conceptos básicos de compresión de vídeo**
- 3. Sistemas estándar de compresión de vídeo**
- 4. Actividades y prácticas**
- 5. Referencias bibliográficas**

5. REFERENCIAS BIBLIOGRÁFICAS

5. Referencias bibliográficas

- Vídeos Bit Experience 2018
(http://www.ifema.es/broadcast_01/Informacion_general/Streaming/index.htm)
- “Next-Generation Video Coding and Streaming”, Benny Bing, Wiley, 2015.
- “Fundamental of multimedia”, Li,Ze-Nian, Springer, 2014.
- “H.264 and MPEG-4 Video Compression”, I. E. G. Richardson, Wiley, 2010. 2nd Edition.
- “Multimedia over IP and Wireless Networks”, M. Van der Schaar- Philip A. Chou, 2007.
- “Standard Codecs: Image Compression to Advanced Video Coding”, M. Ghanbari, IEEE Telecommunications Series 49, 2003.