



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

- **Introducción**
- Necesidad de la codificación de fuente
- Percepción de señales de televisión
 - o Sistema Visual Humano
 - o Sistema Auditivo Humano



Introducción: ¿Cómo hemos llegado a la TV de hoy día?

Ingeniería. Solución barata que funcione

- El receptor lo más barato

Adaptación TV a limitaciones y características propias del Sistema Visual Humano (SVH) = ojo humano + corteza visual

- Cómo ve el ojo condiciona el diseño de los sistemas TV.
 - o Qué ve el ojo y qué no ve (Cosas muy pequeñas, cosas muy rápidas, colores)
 - o Qué se puede suprimir (Redundancia psicofísica)

Evolución y compatibilidad con sistemas anteriores

- Lo técnicamente práctico => Realizable, barato (función de tecnología disponible)
- Lo que guarda compatibilidad (TV Analógica: B/N – Color)
- Convivencia analógico/digital



Introducción: codificación de fuente

La codificación de fuente tiene como objetivo reducir la tasa binaria de las señales a transmitir controlando la posible pérdida de calidad

- Codificación sin pérdidas
 - o Basada en estadísticos
 - o Límite == entropía
 - o Raramente necesaria en señales audiovisuales (e.g., imagen médica, señales forenses)
- Codificación con pérdida
 - o Basada en diversas técnicas
 - o Límite == curva tasa-distorsión
 - o La generalmente aplicada a señales audiovisuales
 - o Objetivo primordial == maximizar la calidad subjetiva



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

- *Introducción*
- **Necesidad de la codificación de fuente**
- Percepción de señales de televisión
 - o Sistema Visual Humano
 - o Sistema Auditivo Humano



Necesidad de la codificación de fuente (I) (*)

Imagen

- 960x1280 píxeles
- 3 colores
- 8 bpp
- Fichero RAW: 3.686.400 bytes
- Fichero JPEG: 598.016 bytes
- Tasa de compresión: 6,164



Video (TV):

- 720x576 píxeles/cuadro
- 3 colores
- 8 bpp
- 25 cuadros/segundo
- Total: 248.832.000 bps = 248,832 Mbps
- Canal TVD: 3-4 Mbps
- Tasa de compresión = 82,944 -62,208

Necesidad de la codificación de fuente (II) (*)

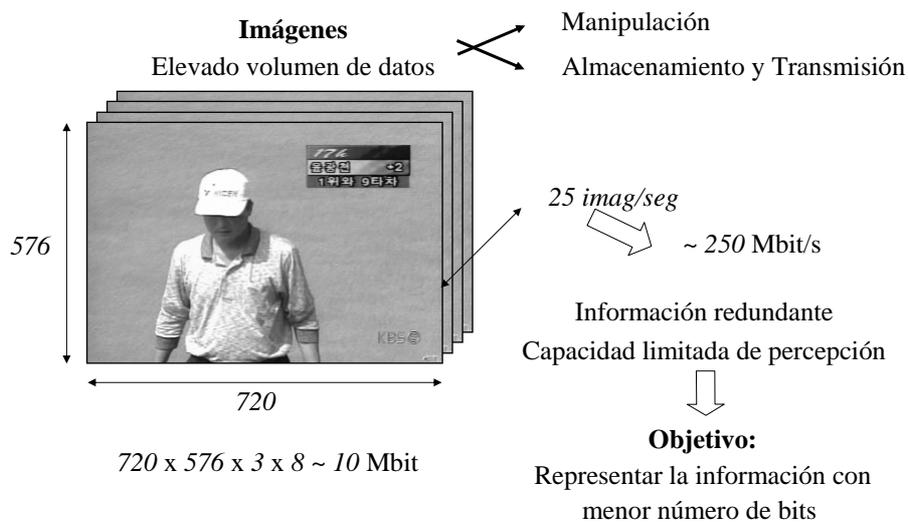
Audio (CD)

- 44100 muestras/segundo
- 16 bits por muestra
- 2 canales (estéreo)
- Total: 1.411.200 bps = 1,4112 Mbps
- Fichero (4:43 minutos): 399.369.600 bits = 49.921.200 bytes = 48.751,171 Kbytes = 47,609 Mbytes
- Fichero MP3 (4:43 minutos): 2.297.856 bytes = 2.244 Kbytes = 2,191 Mbytes
- Tasa de compresión = 21,725

Audio TVA (NICAM-728):

- 32000 muestras/segundo
- 14 bits por muestra
- 2 canales (estéreo)
- Total: 896.000 bps = 896 kbps
- Canal audio NICAM.728 = 640 kbps
- Tasa de compresión = 1,4

Necesidad de Reducción de la Información de vídeo



Escuela Politécnica Superior Universidad Autónoma de Madrid 

Reducción basada en Color (I): de RGB a YCrCb

Con autorización del autor (Luis Salgado Álvarez de Sotomayor, UPM)

Televisión Digital (JoseM.Martinez@uam.es, 2005-2010) Representación y Codificación AV en TVD: Introducción (9)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

Reducción basada en Color (II): Luminancia y Crominancia

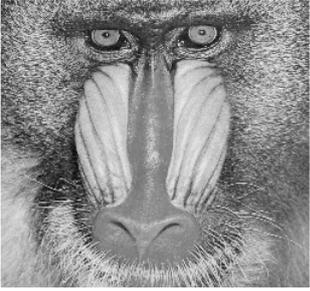
¿Sería posible reducir la información de color?

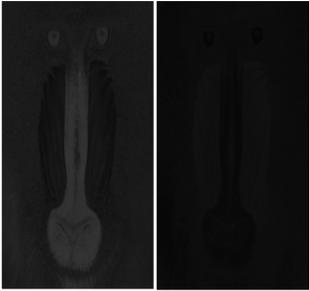
Con autorización del autor (Luis Salgado Álvarez de Sotomayor, UPM)

Televisión Digital (JoseM.Martinez@uam.es, 2005-2010) Representación y Codificación AV en TVD: Introducción (10)

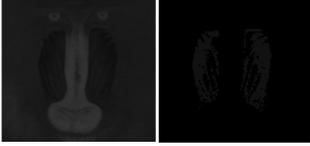
Escuela Politécnica Superior Universidad Autónoma de Madrid 

Reducción basada en Color (III): agudeza visual





4:2:2
~ 166 Mbit/s



4:2:0
~ 124,5 Mbit/s



Y según la aplicación, se utilizan imágenes de diferente tamaño

Con autorización del autor (Luis Salgado Álvarez de Sotomayor, UPM)
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Comunicación AV en TVD: Introducción (11)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

Reducción basada en percepción subjetiva de la distorsión



Degradar la imagen de manera **selectiva** para poder codificarla empleando menos bits **reduciendo el efecto perceptual de la distorsión**

Con autorización del autor (Luis Salgado Álvarez de Sotomayor, UPM)
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción (12)



Técnicas de codificación de fuente

¿Cómo puedo “compactar” eficientemente la salida (representación) de una fuente de información?

- Sin pérdidas: basándose en los estadísticos de la fuente
 - Símbolos más/menos probables menos/más bits.
- Con pérdidas
 - Destino final de la información: ser humano
 - Puedo eliminar información poco relevante, poco perceptible, ...
 - Sistema Visual Humano: Agudeza Visual (separación visible), Mezcla Aditiva Espacial (3 colores), Mezcla Aditiva Temporal (cuadros por segundo),
 - Sistema Auditivo Humano: Umbral de Audición (función de la frecuencia), Enmascaramiento (en tiempo y en frecuencia)
 - Calidad subjetiva
 - Si fuente analógica, siempre habrá pérdidas por el paso A/D (cuantificación)



Representación y Codificación de Señales Audiovisuales en Televisión Digital

Introducción

- *Introducción*
- *Necesidad de la codificación de fuente*
- **Percepción de señales de televisión**
 - Sistema Visual Humano
 - Sistema Auditivo Humano

Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

- *Introducción*
- *Necesidad de la codificación de fuente*
- **Percepción de señales de televisión**
 - **Sistema Visual Humano**
 - Luz y color
 - El fenómeno de la visión
 - El ojo humano
 - Respuesta del ojo humano como sensor de luz y color
 - Resolución del ojo
 - **Sistema Auditivo Humano**

SVH - Luz y color (I) (*)

Teorías científicas: dualidad onda-partícula

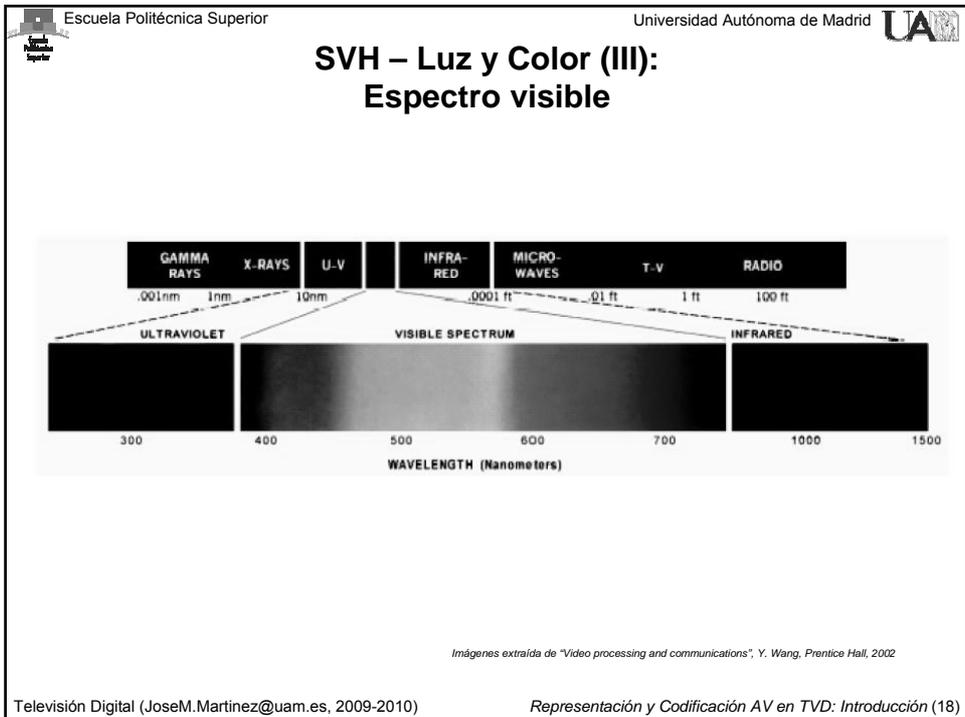
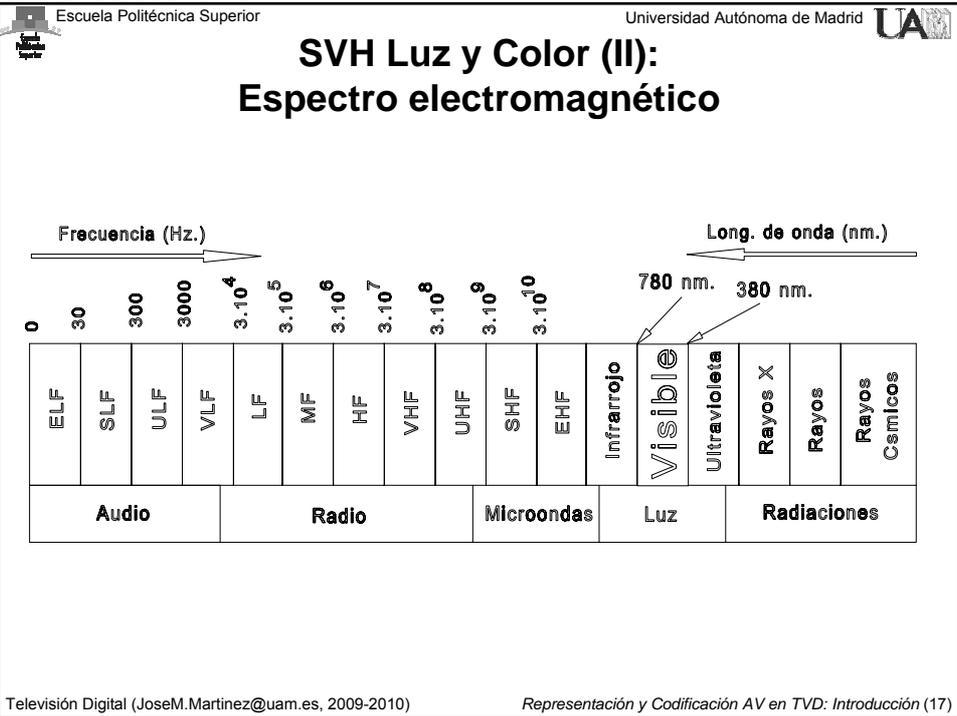
Teoría corpuscular. Partículas: fotones

- Explica fenómenos interacción luz-materia

Teoría ondulatoria. Onda electromagnética

- Variación de campo eléctrico y magnético que se propaga, difracción, etc.
- Longitud de onda

$$\lambda = \frac{c}{f} \quad \begin{array}{l} nm = 10^{-9} m \\ c = 3 \cdot 10^8 m / seg \end{array}$$



SVH – Luz y Color (IV): Fuentes de luz primarias y secundarias

Fuentes primarias o emisivas (aditivas) [illuminating]

- Emisores de luz (e.g., sol, bombilla, monitor TV)
- El color percibido depende de la frecuencia de la luz emitida
- Siguen ley aditiva
 - $R+G+B = \text{White}$

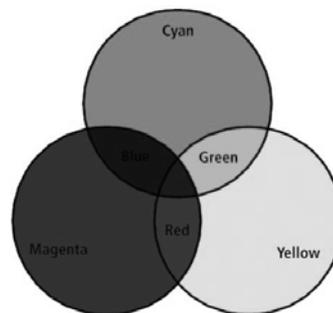
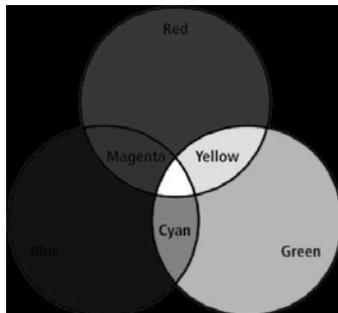
Fuentes secundarias o reflectivas (substractivas) [reflecting]

- Reflejan una luz incidente (e.g., pintura, tejido teñido, papel impreso)
- El color percibido depende de la frecuencia de la luz reflejada
 - Frecuencias emitidas – frecuencias absorbidas
- Siguen ley substractiva
 - $R+G+B = \text{Black}$

SVH – Luz y Color (V): Fuentes de luz primarias y secundarias

Fuentes primarias o emisivas (aditivas)

- Suma de “luces” generadas



Fuentes secundarias o reflectivas (substractivas)

- Resta de “luces” absorbidas = Suma de “luces” reflejadas

Imágenes extraída de “Video processing and communications”, Y. Wang, Prentice Hall, 2002

Escuela Politécnica Superior Universidad Autónoma de Madrid 

SVH – Luz y Color (VI): fuentes aditivas





red



Green



Blue

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción (21)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

SVH – Luz y Color (VII): Percepción de brillo y color

Percepción del brillo

- **Intensidad** de la radiación: función de densidad espectral de energía

$$C(\mathbf{X}, t, \lambda), \mathbf{X} = (x, y, z)$$

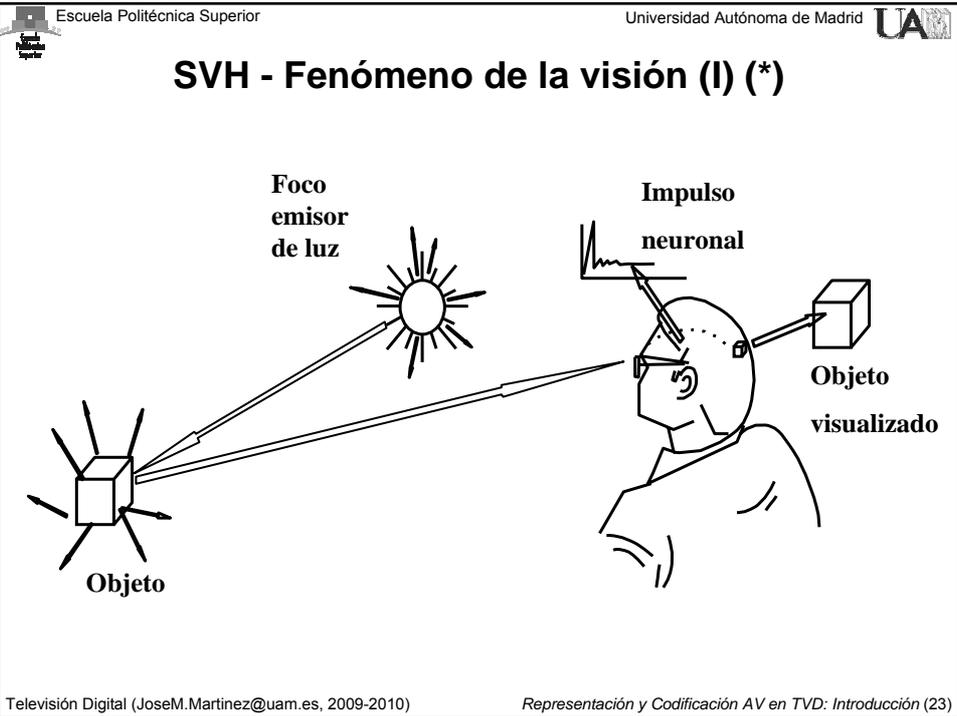
- Si se recibe de un objeto, la intensidad recibida es la incidente (E) afectada por la reflexión

$$C(\mathbf{V}, geom, t, \lambda) = r(\mathbf{V}, geom, t, \lambda) \cdot E(geom, t, \lambda)$$

Percepción del color

- **Distribución espectral** de energía

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



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

SVH - Fenómeno de la visión (II)

Fuentes de luz

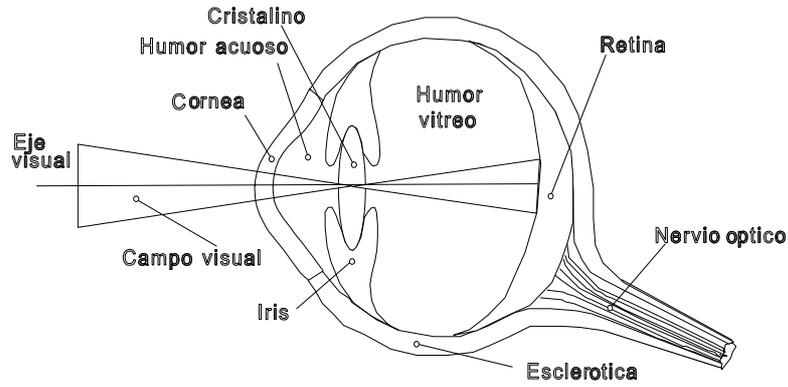
- Luz monocroma, frecuencia "pura", i.e. láser
- Luz policroma, mezcla de frecuencias, i.e. blanco
- Fuentes primarias y secundarias

Percepción: Respuestas del ojo y corteza visual (cerebro)

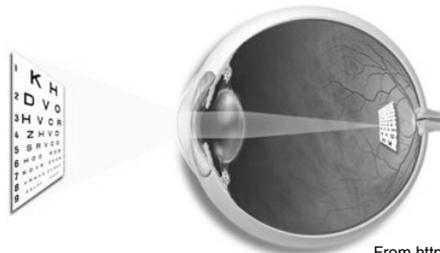
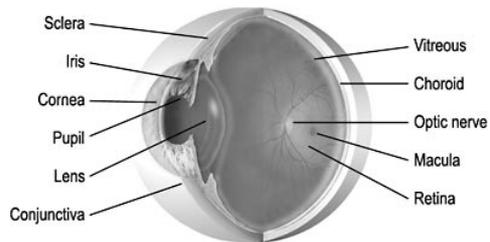
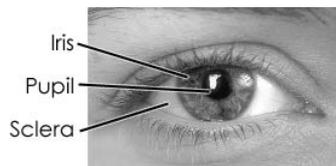
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción (24)



SVH - El ojo humano (I) (*)

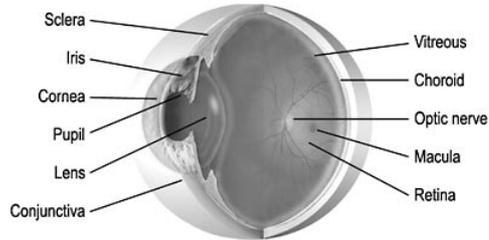


SVH - El ojo humano (II) (*)



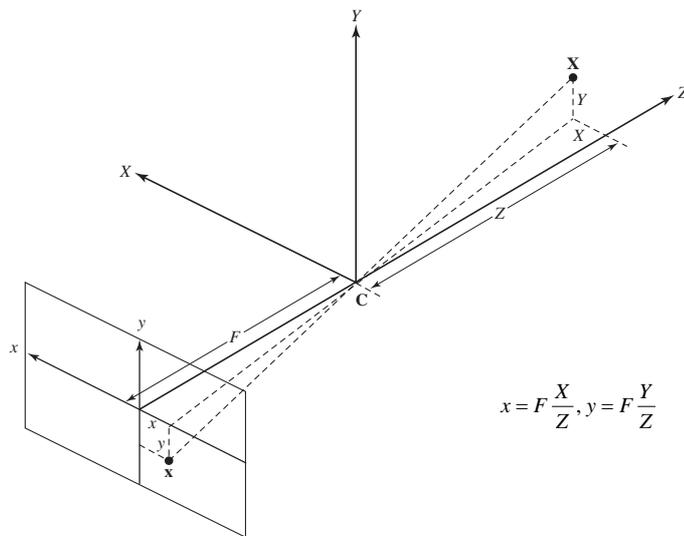
From <http://www.stlukeseye.com/Anatomy.asp>

SVH – El ojo humano (III): Cámara versus ojo humano



Cámara	Ojo	Función
Lentes (objetivo)	Cristalino, cornea: voluntario	Enfoque
Diafragma	Iris (pupila): involuntario	Regulación cantidad de luz
Película (fotoquímico) - CCD	Retina (fotoquímico+nervioso)	Sensibilidad y plano de imagen
Foto? - Cable/Memoria	Nervio óptico (al cerebro)	Transferencia de información

SVH – el ojo humano (IV): modelo de proyección perspectiva



SVH - Respuesta del ojo como sensor de luz y color (I)

“Sensación de luz”:

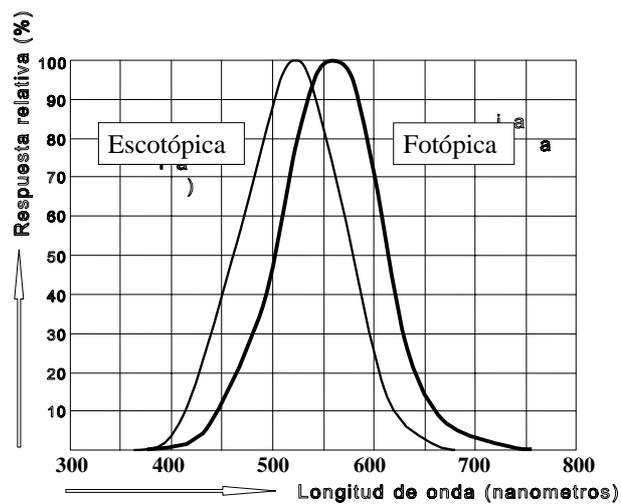
- Valoraciones empíricas/perceptuales

Función de la frecuencia de la luz

Función de la cantidad de luz

- Visión fotópica (luz suficiente, en TV)
- Visión mesotópica (luz media)
- Visión escotópica (poca luz)

SVH – Respuesta del ojo como sensor de luz y color (II): Curvas de respuesta (*)





SVH – Respuesta del ojo como sensor de luz y color (III): Tipos de receptores

Dos tipos de fotorreceptores en la retina

- Bastones (o cilindros): escotópica, poca luz => Intensidad luminosa
- Conos: fotópica, color, luz suficiente (en TV) => color

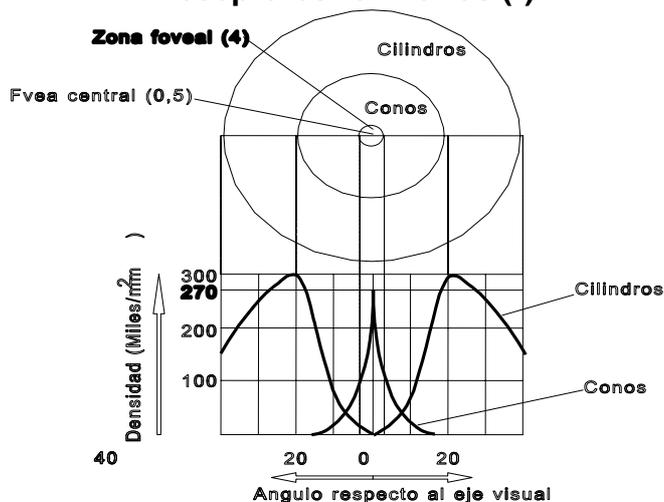
Distribución en la retina

- Centro, zona foveal. Más conos. Mucha luz
- Fuera de zona foveal. Más bastones (cilindros). Poca luz

Conos. 3 tipos, máximos de respuesta a los tres colores primarios, Rojo, Verde, Azul



SVH – Respuesta del ojo como sensor de luz y color (IV): Receptores retinianos (*)



SVH – Respuesta del ojo como sensor de luz y color (V): Teoría tri-receptor del color (*)

T. Young, “On the theory of light and colors”, Philosophical Transactions of the Royal Society of London”, 92:20-71, 1802.

Hay tres tipos de conos, con bandas de paso “overlapping”

- Pico en rojo (aprox. 570 nm)
- Pico en verde (aprox. 535 nm)
- Pico en azul (aprox. 445 nm)
o Distinta sensibilidad

$$a_i(c) = \int C(\lambda) \cdot a_i(\lambda) d\lambda, i \in \{1, 2, 3\}$$

El color percibido depende únicamente de tres números

- En lugar del espectro incidente

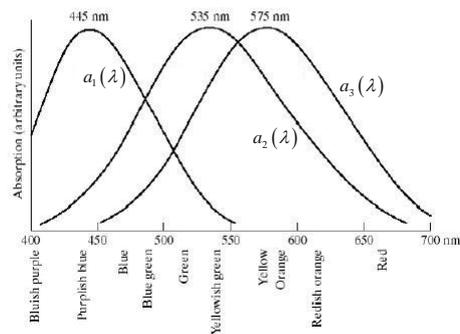


Imagen extraída de “Video processing and communications”, Y. Wang, Prentice Hall, 2002

SVH – Respuesta del ojo como sensor de luz y color (VI): Sensación de color (*)

Luminancia (brillo, energía total) y Crominancia (tinte [tono], saturación [pureza])

- Una luminancia $Y = \int C(\lambda) \cdot a_Y(\lambda) d\lambda, a_Y(\lambda) = a_1(\lambda) + a_2(\lambda) + a_3(\lambda)$

- Dos crominancias
o según el espacio de color

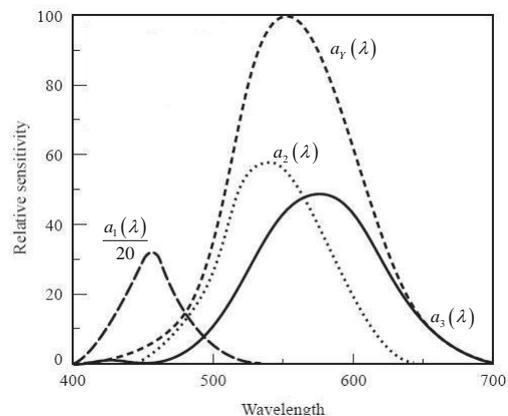


Imagen extraída de “Video processing and communications”, Y. Wang, Prentice Hall, 2002

SVH – Respuesta del ojo como sensor de luz y color (VII): Representación del color (*)

[Maxwell, 1885]: teoría tricromática de la mezcla de colores
La mayor parte de los colores pueden reproducirse mezclando
3 primarios (tristimulus values: T_k , positivos o negativos)

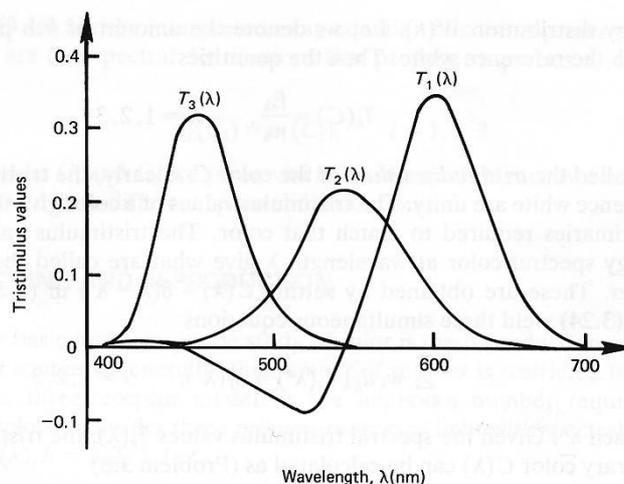
$$C = \sum_{k=1,2,3} T_k C_k$$

Una vez seleccionados los tres colores primarios, se obtienen
los valores T_k del color a representar

$$T_i = \int C(\lambda) T_i(\lambda) d(\lambda) \quad i = 1, 2, 3$$

Donde $T_k(\lambda)$ son las “color matching functions” que se
obtienen mediante experimentos visuales en condiciones
controladas

SVH – Respuesta del ojo como sensor de luz y color (VIII): Representación del color (*)



SVH – Respuesta del ojo como sensor de luz y color (IX): Representación del color (*)

Los valores $T_k(c), k = \{1, 2, 3\}$ definen un sólido de color del que sólo se pueden representar aquellos colores que verifiquen

$$T_k(c) \geq 0, \forall k \in \{1, 2, 3\}$$

El objetivo es buscar primarios que representen tantos como sea posible (*tristimulus values* positivos).

Esta teoría es la base de la captación y representación del color. Una cámara (o monitor) ha de tener sensores (o emisores) con una respuesta en frecuencia igual a las *color matching functions*.

Primarios más populares:

- Displays: RGB
- Printers: CMY (CMYK)

SVH – Respuesta del ojo como sensor de luz y color (X): Especificación del color (*)

Mediante ternas de valores tristimulus values (T_k)

- normalizados respecto al blanco de referencia ($T_k=1, i=1,2,3$)
- CIE RGB primary system ($R_0=700$ nm, $G_0=546.1$ nm, $B_0=435.8$ nm)

Mediante separación de luminancia y crominancia (tinte + saturación):

$$t_k(c) = \frac{T_k(c)}{T_1(c) + T_2(c) + T_3(c)}, k = \{1, 2, 3\} \quad \sum_{k=1}^3 t_k(c) = 1$$

- CIE XYZ, casi todos los colores visibles se generan con T_k positivos, pero XYZ no son realizables.
- Del CIE XYZ derivan YIQ e YUV (valores positivos, TVA), HSI (no lineal), LSI y Lab (diferencias de color).

SVH – Respuesta del ojo como sensor de luz y color (XI): Diagrama de cromaticidad del CIE (*)

x [red]

y [green]

$z = 1 - (x+y)$ [blue]

La línea que une dos puntos cualquiera da la gama de colores que se puede generar mezclándolos

La línea desde el punto de igual energía [blanco] hacia la frontera son los tonos de un mismo color

El triángulo RGB son los colores realizables con estos primarios

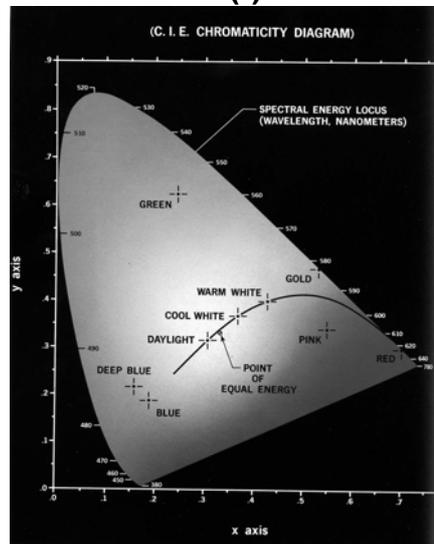


Imagen extraída de "Digital Image Processing", R.C. Gonzalez, P. Wlnts Prentice Hall, 1987

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

Representación y Codificación AV en TVD: Introducción (39)

SVH – Respuesta del ojo como sensor de luz y color (XII): Rendimiento del ojo como sensor de luz

Muy buen localizador posicional

Muy mal analizador espectral

- En la longitud de onda, colores
- En el tiempo, variación de estímulo visual

Diseño de los Sistemas de TV

- Calidad donde se necesite => donde el ojo lo vea

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

Representación y Codificación AV en TVD: Introducción (40)

SVH - Resolución del ojo (I)

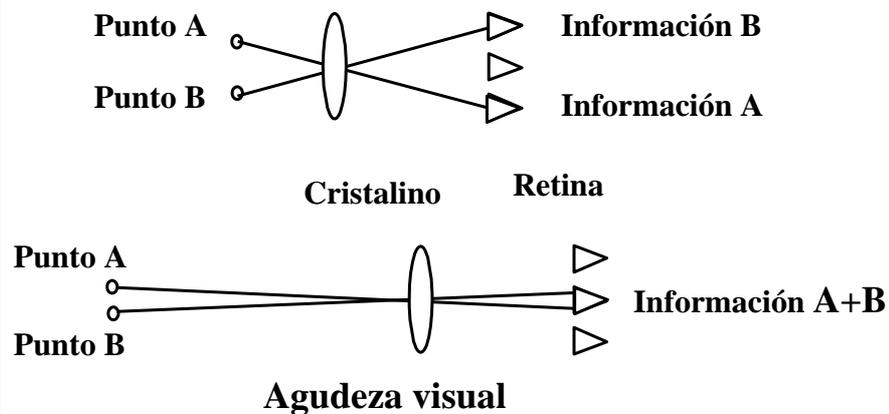
Se forma una imagen sobre la retina físicamente, sobre los receptores.

- Limitados en número y tiempo de reacción

El ojo no es perfecto, no ve cosas:

- Arbitrariamente pequeñas: imperfecta resolución espacial
- Arbitrariamente rápidas: imperfecta resolución temporal

SVH – Resolución del ojo (II): Mezcla aditiva espacial –MAE–





SVH – Resolución del ojo (III): Agudeza Visual

Retina, puntos separados ángulo $>1\text{min}$:

- Sensores distintos: fibras nerviosas distintas
- **Dos** informaciones al cerebro

Puntos separados ángulo $< 1\text{ min}$:

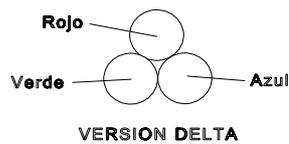
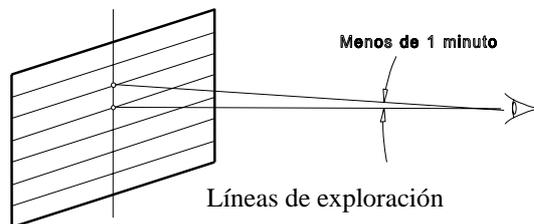
- Mismo sensor: una única fibra nerviosa
- **Una única** información al cerebro

Agudeza visual: capacidad de distinguir dos puntos próximos.



SVH- Resolución del ojo (IV): Uso de la MAE

Ejemplos: líneas TV, triadas fósforo



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

SVH – Resolución del ojo (V): Uso de la MAE

Ejemplos: triadas fósforo

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción (45)

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

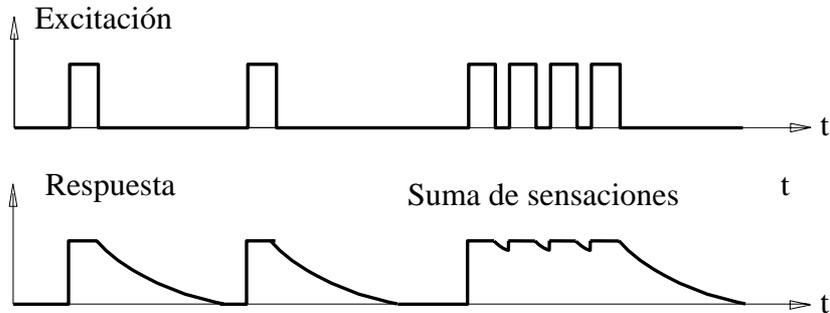
SVH – Resolución del ojo (VI): Mezcla aditiva temporal –MAT–

Respuesta del ojo a un impulso

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Introducción (46)

SVH – Resolución del ojo (VII): Mezcla Aditiva Temporal –MAT-

Respuesta de ojo a varios impulsos seguidos



Memoria Visual

SVH – Resolución del ojo (VIII): Memoria Visual

Serie de estímulos luminosos

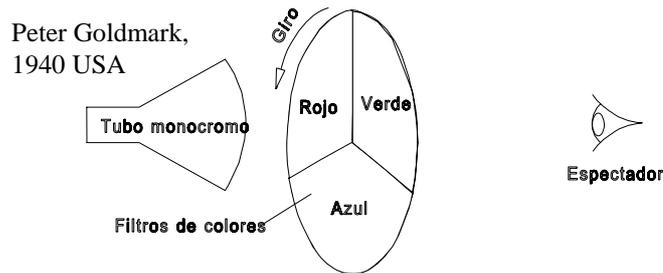
- Separados más de 50ms se perciben como independientes/distintos
- Menos de 50ms: cada nuevo impulso provoca una respuesta que se suma a la anterior
- Alrededor de 50ms: parpadeo o *flicker*, incómodo, ni mezclados ni separados

Integración temporal de impulsos

- Memoria Visual de lo que se ha visto

SVH – Resolución del ojo (IX): Uso de la MAT

Ejemplos: Cine, TV número de imágenes limitado, color como suma de tres imágenes



Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

- *Introducción*
- *Necesidad de la codificación de fuente*
- **Percepción de señales de televisión**
 - **Sistema Visual Humano**
 - Luz y color*
 - El fenómeno de la visión*
 - El ojo humano*
 - Respuesta del ojo humano como sensor de luz y color*
 - Resolución del ojo*
 - **Sistema Auditivo Humano**
 - El oído**
 - Percepción de sonidos**
 - Enmascaramiento**

SAH – el oído (I)

Comportamiento a grandes rasgos

- Analizador espectral (con limitaciones)
- Capaz de localizar la fuente sonora
- La codificación perceptual de audio aprovecha estas propiedades

Fisiología del aparato auditivo humano

- Oído externo. Audición binaural y adaptación de la señal (filtrado)
- Oído medio. Transmisión eficaz del sonido al caracol (amplificación)
- Oído interno. Analiza la señal (distingue frecuencias)

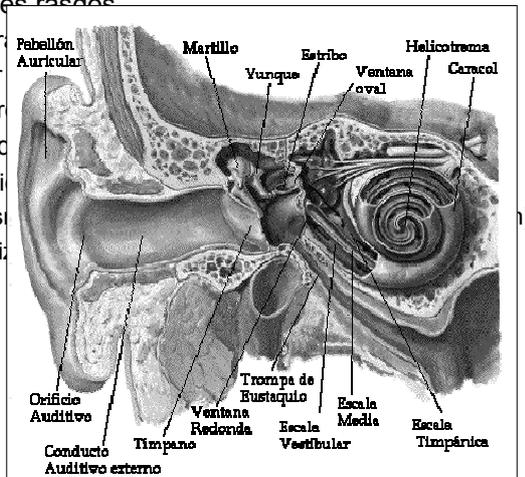
SAH – el oído (II)

Comportamiento a grandes rasgos

- Analizador espectral (con limitaciones)
- Capaz de localizar la fuente sonora
- La codificación perceptual de audio aprovecha estas propiedades

Fisiología del aparato auditivo humano

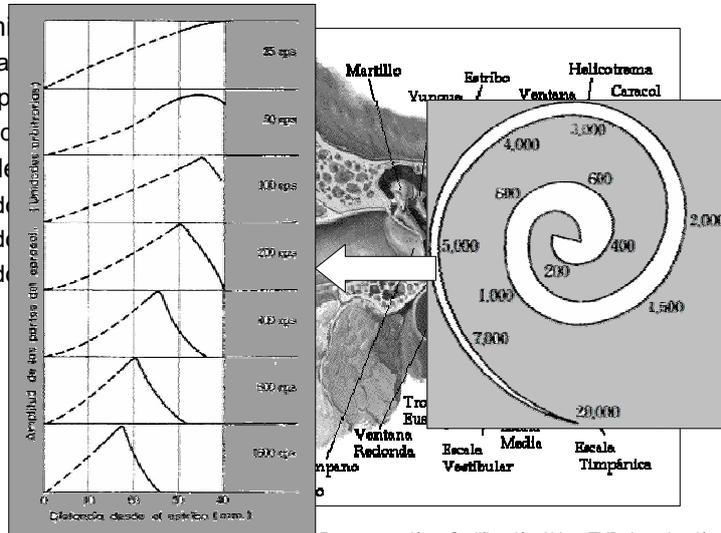
- Oído externo. Audición binaural y adaptación de la señal (filtrado)
- Oído medio. Transmisión eficaz del sonido al caracol (amplificación)
- Oído interno. Analiza la señal (distingue frecuencias)



SAH – el oído (IIII)

Comportam

- Ana
 - Cap
 - La c
- Fisiología de
- Oíd
 - Oíd
 - Oíd



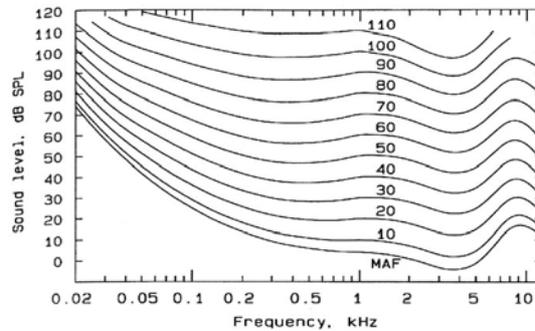
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)

Representación y Codificación AV en TVD: Introducción (53)

SAH - Percepción de sonidos (I)

Factores que intervienen

- Nivel sonoro (altura de un sonido). Mide la intensidad percibida (en fonos). Depende de la frecuencia
 - Umbral absoluto de audición (MAF). Depende de la frecuencia
 - Saturación (umbral de dolor)



- Banda crítica ...
- Enmascaramiento ...

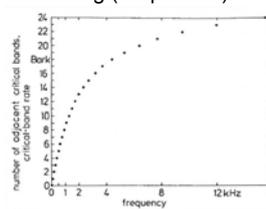
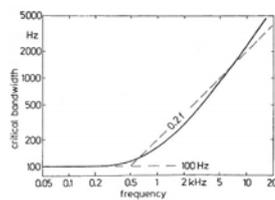
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)

Representación y Codificación AV en TVD: Introducción (54)

SAH - Percepción de sonidos (II)

Banda crítica

- Determinada por la capacidad de discriminar dos tonos muy próximos en frecuencia (el sistema auditivo no puede separar dos tonos que estén dentro de una banda crítica)
- Modelo: banco de filtros (banda estrecha) no uniforme. Ancho de la banda de paso del filtro mayor a mayor frecuencia central
- Se pasa a una escala lineal: tonalidad (medida en barks, cada banda crítica tiene un ancho de 1 bark)
 - Frecuencias < 500 Hz : 1 Bark \approx freq/100
 - Frecuencias > 500 Hz : 1 Bark \approx 9 + 4 log (freq / 1000)



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

Representación y Codificación AV en TVD: Introducción (55)

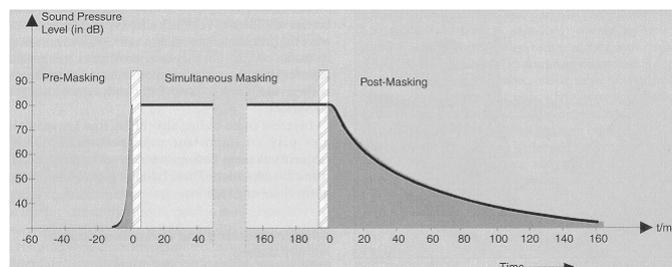
SAH - Enmascaramiento (I)

Enmascaramiento

- Fenómeno por el cual el oído no es capaz de separar dos sonidos
- Enmascaramiento en tiempo y frecuencia

Enmascaramiento en tiempo

- Un tono más alto enmascara otro tono posterior (o anterior) de menor intensidad



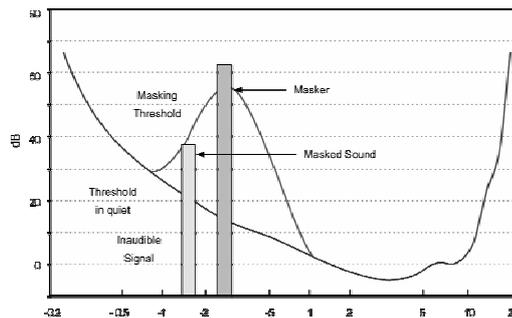
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010)

Representación y Codificación AV en TVD: Introducción (56)

SAH - Enmascaramiento (II)

Enmascaramiento en frecuencia

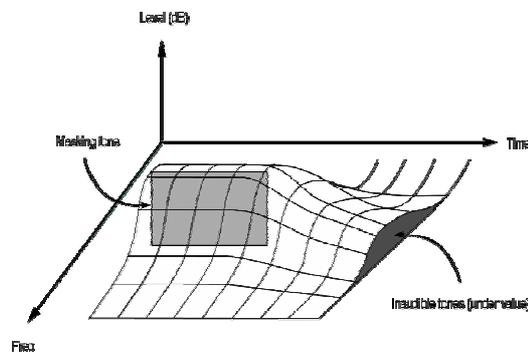
- Fuera de la banda crítica también
- Dos tipos
 - Tono enmascarando otro tono (umbral depende de P_{ton})
 - Tono enmascarando ruido (umbral depende de la frecuencia)



SAH - Enmascaramiento (III)

Enmascaramiento en tiempo y frecuencia

- Los enmascaramientos en tiempo y en frecuencia no se dan de forma aislada, sino que se combinan en una envolvente de máscara que se extiende a lo largo de la frecuencia y del tiempo, por debajo de la cual los sonidos son inaudibles.





Representación y Codificación de Señales Audiovisuales en Televisión Digital Introducción

- *Introducción*
- *Necesidad de la codificación de fuente*
- *Percepción de señales de televisión*
 - *Sistema Visual Humano*
 - Luz y color*
 - El fenómeno de la visión*
 - El ojo humano*
 - Respuesta del ojo humano como sensor de luz y color*
 - Resolución del ojo*
 - *Sistema Auditivo Humano*
 - El oído*
 - Percepción de sonidos*
 - Enmascaramiento*



Créditos

Para la elaboración de algunas de estas transparencias se ha hecho uso de material

- Jesús Bescós Cano, EPS-UAM
- Luis Herranz, EPS-UAM
- Enrique Rendón Angulo, E.T.Ing.Telecomunicación, UPM
- Luis Salgado, E.T.S.Ing.Telecomunicación, UPM
- Yao Wang, Polytechnic University, Brooklyn



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

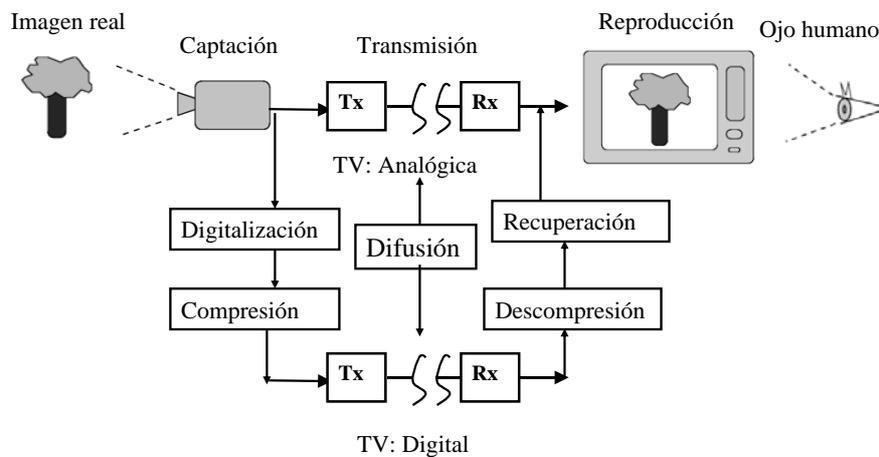
- Introducción
- Digitalización de señales de vídeo
 - Introducción
 - UIT-R BT.601
- Multiplexación y Codificación de la señal de vídeo digital
 - Introducción
 - UIT-R BT.656 y 1302
 - Interfaces paralelo y serie

Representación y Codificación de Señales Audiovisuales en Televisión Digital

Digitalización y Codificación de señales de vídeo

- **Introducción**
- Digitalización de señales de vídeo
 - Introducción
 - UIT-R BT.601
- Multiplexación y Codificación de la señal de vídeo digital
 - Introducción
 - UIT-R BT.656 y 1302
 - Interfaces paralelo y serie

Introducción: Sistema de Televisión





Introducción: Ventajas de la TV Digital

- No existe intermodulación luminancia/crominancia
- Multiregeneración de señal sin pérdida de calidad
- Mayor calidad (más ancho de banda)
- Compresión de la información
- Mejores posibilidades de tratamiento (ordenadores)
- Otros servicios añadidos



Introducción: Terminología de tipos vídeo

Vídeo en componentes

- Tres componentes de color almacenados/transmitidos por separados
- Analógico: RGB o YUV (YIQ)
- Digital: YCrCb

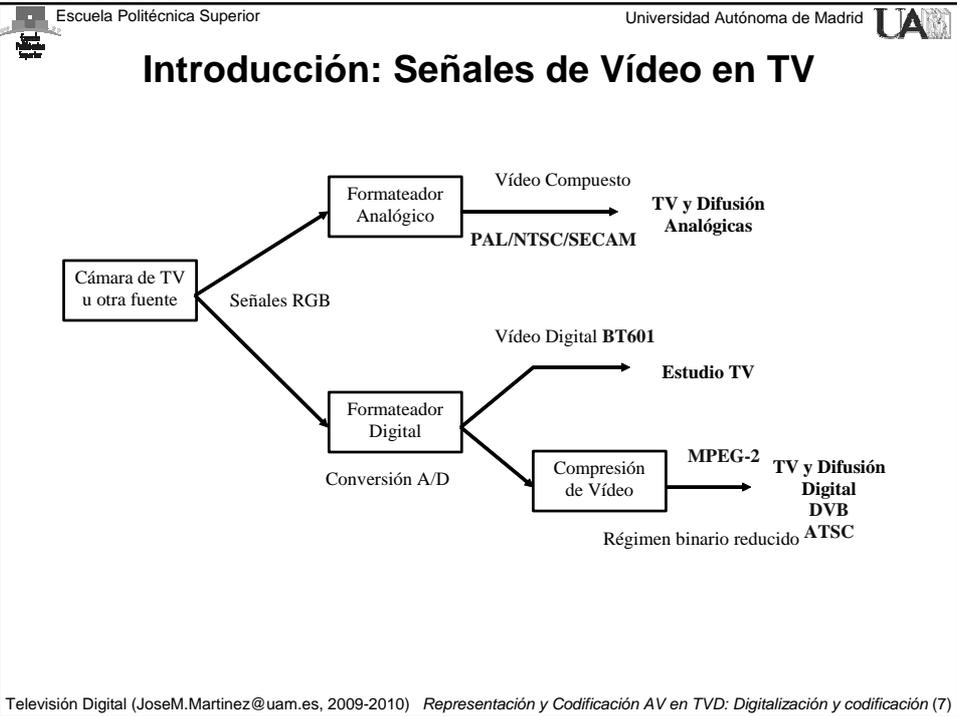
Vídeo compuesto

- Convertir RGB a YUV (YIQ)
- Multiplexar YUV en una única señal
- Aparatos de consumo de vídeo analógico

S-video

- Y y croma (QAM de UV –IQ-) se almacenan por separado
- Aparatos de vídeo de gama alta

Los monitores de gama alta admiten todas estas entradas



Escuela Politécnica Superior Universidad Autónoma de Madrid

Introducción: Formatos de vídeo digital

Video Format	Y Size	Color Sampling	Frame Rate (Hz)	Raw Data Rate (Mbps)
HDTV Over air, cable, satellite, MPEG2 video, 20-45 Mbps				
SMPTE296M	1280x720	4:2:0	24P/30P/60P	265/332/664
SMPTE295M	1920x1080	4:2:0	24P/30P/60I	597/746/746
Video production, MPEG2, 15-50 Mbps				
BT.601	720x480/576	4:4:4	60I/50I	249
BT.601	720x480/576	4:2:2	60I/50I	166
High quality video distribution (DVD, SDTV), MPEG2, 4-10 Mbps				
BT.601	720x480/576	4:2:0	60I/50I	124
Intermediate quality video distribution (VCD, WWW), MPEG1, 1.5 Mbps				
SIF	352x240/288	4:2:0	30P/25P	30
Video conferencing over ISDN/Internet, H.261/H.263, 128-384 Kbps				
CIF	352x288	4:2:0	30P	37
Video telephony over wired/wireless modem, H.263, 20-64 Kbps				
QCIF	176x144	4:2:0	30P	9.1

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (8)



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- *Introducción*
- **Digitalización de señales de vídeo**
 - **Introducción**
 - UIT-R BT.601
- **Multiplexación y Codificación de la señal de vídeo digital**
 - **Introducción**
 - UIT-R BT.656 y 1302
 - Interfaces paralelo y serie



Introducción a la digitalización de señales (I)

Proceso general

- Digitalización y Recuperación

Digitalización

- **Muestreo**
 - Necesidad de filtrado
- **Cuantificación**
 - Error de cuantificación y saturación
- **Codificación**
 - Codificación de línea (asignación de símbolos a señales digitales)
 - Codificación con pérdidas (MPEG) => compresión



Introducción a la digitalización de señales (II): Proceso de Digitalización (Analógico \Rightarrow Digital)

Señal Analógica: Información en el voltaje (mV)

- Ruido y Distorsiones cambian la información

Filtrado: Limitación de la banda de la señal analógica

- Para evitar el *aliasing*

Muestreo: Selección de una serie de datos

- Valores tomados de forma periódica (sigue siendo tensión)

Cuantificación: Discretización de la amplitud

- A cada grupo de voltajes se asigna un número (nivel de cuantificación: Q): error/ruido de cuantificación

Codificación: Generación de la señal digital

- Representar el número Q en binario ('0' y '1')
- Asignar voltajes a cada valor de bit (NRZ)
 - o Ruido y Distorsiones cambian voltajes pero 'no' bits

(Multiplexación y Compresión)



Introducción a la digitalización de señales (III): Proceso de recuperación (Digital \Rightarrow Analógico)

(Descompresión y Desmultiplexación)

Decodificación : de 0's y 1's a nivel Q

Reconstrucción: de nivel Q a valor de tensión

Interpolación : de valores aislados a señal

Filtrado posterior para limitar en banda (igual al de la entrada)



Introducción a la digitalización de señales (IV): Conversión A/D y D/A

Funciones realizadas en:

★ Conversor Analógico/Digital

- o Muestreo
- o Cuantificación uniforme
- o Codificación en binario natural

★ Conversor Digital/Analógico

- o Descodificación
- o Reconstrucción
- o Interpolación



Introducción a la Digitalización de Vídeo (I): ¿qué señal se digitaliza?

Vídeo compuesto

- No aporta ninguna ventaja
- Presenta problemas con la subportadora de color

Componentes RGB.

- Excesiva velocidad binaria
- Escasa ganancia de calidad, por lo que se usa menos

Luminancia y diferencias de color

- Compatibilidad entre sistemas
- Aprovechamiento de las limitaciones del SVH



Introducción a la Digitalización de Vídeo (II): Digitalización de vídeo raster

Muestreo de la señal raster = muestrear las líneas (dimensión horizontal)

La tasa de muestreo debe ser seleccionada adecuadamente

- Para que estén alineadas verticalmente, la tasa tiene que ser múltiplo de la frecuencia de línea
- Intervalo de muestreo horizontal = Intervalo de muestreo vertical
- Tasa de muestreo común para los diferentes sistemas

$$f_s = 858 f_l (\text{NTSC}) = 864 f_l (\text{PAL}) = 13.5 \text{ MHz}$$

$$f_l (\text{NTSC}) = 15.734 \text{ KHz} \quad \square \quad 525 \text{ l/c} * 30 \text{ c/s} = 15.75 \text{ Khz} \quad (525 * 29,94 = 15,7185)$$

$$f_l (\text{PAL}) = 15.625 \text{ KHz} = 625 \text{ l/c} * 25 \text{ c/s} = 15.625 \text{ Khz}$$

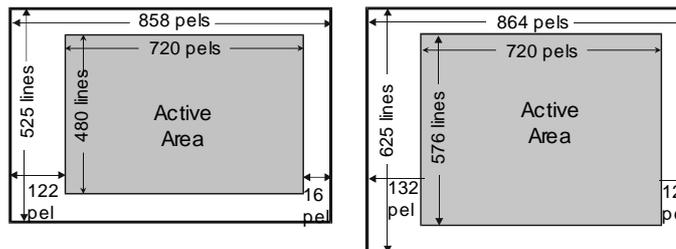


Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- *Introducción*
- **Digitalización de señales de vídeo**
 - *Introducción*
 - **UIT-R BT.601**
- Multiplexación y Codificación de la señal de vídeo digital
 - *Introducción*
 - UIT-R BT.656 y 1302
 - Interfaces paralelo y serie

UIT BT.601 (I): Introducción

Estándar para la digitalización y codificación (binaria) de señales de vídeo digital para estudio con un máximo de valores comunes en parámetros de sistemas de 525 y 625 líneas



525/60: 60 field/s

NTSC

625/50: 50 field/s

PAL

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (17)

UIT BT.601 (II): Introducción

Posibilidades de codificación de componentes

- Componentes Y, Cr, Cb
- Componentes R, G, B

Las estructuras de muestreo son estáticas espacialmente

Las muestras de color tienen que estar co-situadas

Para sistemas 525 y 625 líneas el número de muestras activas por cada línea es el mismo (720)

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (18)

**UIT-R BT.601 (III): Luminancia**

Se utiliza:

$$Y = E'_Y = 0.299 \cdot E'_R + 0.587 \cdot E'_G + 0.114 \cdot E'_B$$

Primarios gamma-correctados y normalizados (0 a 1 V)
Y con valores entre 0 y 1 V

Norma AES/EBU **N-10**

- Niveles Y de 0 a 700 mV

**UIT-R BT.601 (IV): Diferencias de color**

$$(E'_R - E'_Y) = 0.701 \cdot E'_R - 0.587 \cdot E'_G - 0.114 \cdot E'_B$$

$$(E'_B - E'_Y) = -0.299 \cdot E'_R - 0.587 \cdot E'_G + 0.886 \cdot E'_B$$

Al estar los componentes primarios normalizados, la luminancia está normalizada, pero no así las diferencias de color que varían entre [-0.701,+0.701] y [-0.886,+0.886] respectivamente.

Condition	E'_R	E'_G	E'_B	E'_Y	$E'_R - E'_Y$	$E'_B - E'_Y$
White	1.0	1.0	1.0	1.0	0	0
Black	0	0	0	0	0	0
Red	1.0	0	0	0.299	0.701	-0.299
Green	0	1.0	0	0.587	-0.587	-0.587
Blue	0	0	1.0	0.114	-0.114	0.886
Yellow	1.0	1.0	0	0.886	0.114	-0.886
Cyan	0	1.0	1.0	0.701	-0.701	0.299
Magenta	1.0	0	1.0	0.413	0.587	0.587



UIT-R BT.601 (V): Diferencias de color renormalizadas

Se calculan los coeficientes para que las diferencias de color varíen entre $-0,5$ a $0,5$ V \Rightarrow renormalización

$$\begin{aligned} Cr &= Kr \cdot (E'_R - Y) & \Rightarrow & Cr = 0.713 \cdot (E'_R - Y) \\ Cb &= Kb \cdot (E'_B - Y) & \Rightarrow & Cb = 0.564 \cdot (E'_B - Y) \end{aligned}$$

AES/EBU **N-10** establece Cr y Cb con excursión (rango dinámico) de 700mV, entre $-0,35$ y $0,35$ V



UIT-BT.601 (VI): Cuantificación (8 bits)

Se aplica cuantificación uniforme

Los extremos se dejan para señalización (0 y 255)

La señal de luminancia se reduce a 220 niveles (salvo sobrecarga)

$$\bar{Y} = 219 \cdot (E'_Y) + 16$$

y se redondea al entero más cercano

Las diferencias de color ocupan 225 niveles, siendo el 128 el valor nulo

$$\bar{C}_R = 224(Cr) + 128 = 160 \cdot (E'_R - Y) + 128$$

$$\bar{C}_B = 224(Cb) + 128 = 126 \cdot (E'_B - Y) + 128$$



UIT-R BT.601 (VII): Esquemas de muestreo

Patrón de muestreo espacial

- Ortogonal y Estático
- Repetido en cada línea, campo (UIT: trama) y cuadro

Jerarquía de muestreo o familia ampliable

Muestras de componentes cosituadas

Señales Cr y Cb submuestreadas un número entero respecto a Y

Descrito con secuencia de tres dígitos, relación entre frecuencias de muestreo. Ej: 4:2:2



UIT-R BT.601 (VIII): Posibles esquemas de muestreo

4:4:4 en R,G,B ó Y,Cr,Cb

- misma frecuencia de muestreo f_m en las tres componentes

4:2:2 en Y,Cr,Cb

- Muestreo completo (f_m) en Y
- Submuestreo horizontal 2:1 en Cr y Cb ($f_m/2$)

4:1:1 en Y,Cr,Cb (solamente como ejemplo; no se usa)

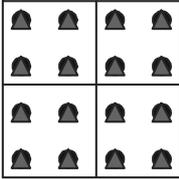
- Ídem en Y, submuestreo horizontal 4:1 en Cr y Cb

4:2:0 en Y,Cr,Cb (no significa que Cb se pierda)

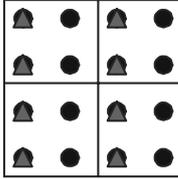
- Submuestreo horizontal y vertical 2:1 en Cr y Cb

Escuela Politécnica Superior Universidad Autónoma de Madrid 

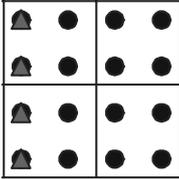
UIT-R BT.601 (IX): “Co-situación” de las muestras



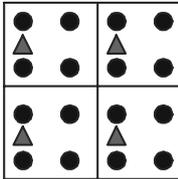
4:4:4
For every 2x2 Y Pixels
4 Cb & 4 Cr Pixel
(No subsampling)



4:2:2
For every 2x2 Y Pixels
2 Cb & 2 Cr Pixel
(Subsampling by 2:1 horizontally only)



4:1:1
For every 4x1 Y Pixels
1 Cb & 1 Cr Pixel
(Subsampling by 4:1 horizontally only)



4:2:0
For every 2x2 Y Pixels
1 Cb & 1 Cr Pixel
(Subsampling by 2:1 both horizontally and vertically)

● Y Pixel ▲ Cb and Cr Pixel

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (25)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.601 (X): Representación en 8 y 10 bits

Las fuentes de señal y los monitores trabajan con 8 bits por muestra

Para añadir precisión durante el tratamiento

- Se añaden 2 bits adicionales
- Los equipos de 8 bits los ponen a cero

En la representación se considera

● 8 bits parte entera	10010001 ⇒ 145d ⇒ 91h
● 2 bits parte fraccionaria	10010001 <u>01</u> ⇒ 145. <u>25</u> d ⇒ 91. <u>4</u> h

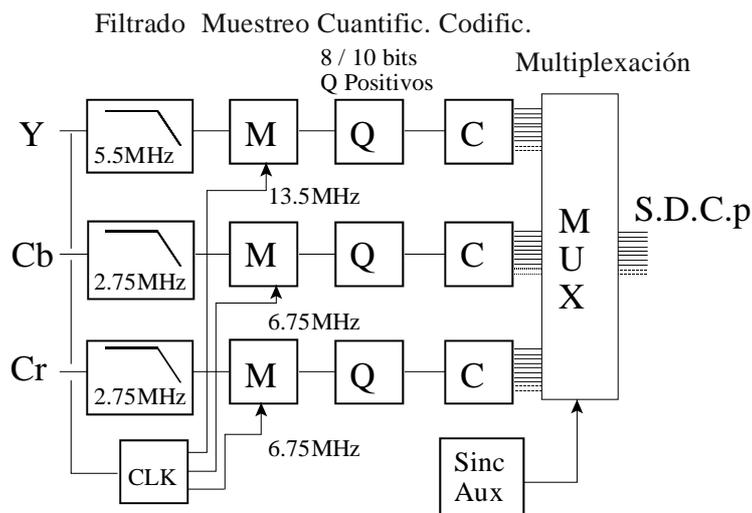
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (26)

UIT-R BT.601 (XII): Posibilidades en UIT-R BT.601

- Sistemas de 625 / 525 líneas por cuadro
- Esquema 4:4:4 en R,G,B e Y,Cr,Cb
 - Equipamiento de estudio y aplicaciones de video de alta calidad
- Esquema 4:2:2 en Y,Cr,Cb
 - Interfaz principal entre equipamiento de estudio y formato de intercambio internacional de programas (contribución)
- Frecuencia de muestreo y relación de aspecto
 - **Convencional:** $f_m=13,5$ MHz para 4:3
 - **Anamórfico:** $f_m=13,5$ MHz para 16:9
 - **Muestreo extendido:** $f_m=18,0$ MHz sólo para 16:9
- Cuantificación con 8 ó 10 bits

Nos centraremos en analizar el 4:2:2 convencional al ser el más utilizado

UIT-R BT.601 (XII): 4:2:2, Y,Cr,Cb a 13,5 MHz Esquema del tratamiento





UIT-R BT.601 (XIII): 4:2:2, Y,Cr,Cb a 13,5 MHz Muestreo y filtrado

Frecuencia principal 13,5MHz a 4:2:2

- Luminancia $f_m=13,5 \text{ Mhz}$ $\Rightarrow f_c$ Nyquist 6,75MHz
- Crominancias $f_m/2=6,75 \text{ MHz}$ $\Rightarrow f_c$ Nyquist 3,375 MHz

Periodo de muestreo de Y: $T=74\text{ns}$

Periodo de diferencias de color $2T=148\text{ns}$

- Las muestras de Cr y Cb **coinciden** en el espacio con las muestras impares de Y (1^a,3^a...) (n^{os} 0,2...)

Número entero de muestras por línea $T_h=64\mu\text{s}$

- Se obtienen 864 muestras de Y y 432 de cada crominancia (63,93 μs)
 - o (858 muestras Y en NTSC)
- Se repite en cada línea



UIT-R BT.601 (XIV): 4:2:2, Y,Cr,Cb a 13,5 MHz Temporización

Se consideran muestras de vídeo válidas durante

$720 T = 53,33\mu\text{s} > 52 \mu\text{s}$ de analógico

- **720** muestras de Y, **360** muestras de Cr y **360** de Cb
- 18T de borrado analógico (18* 74=1,332 μs)
 - o 10+702+8

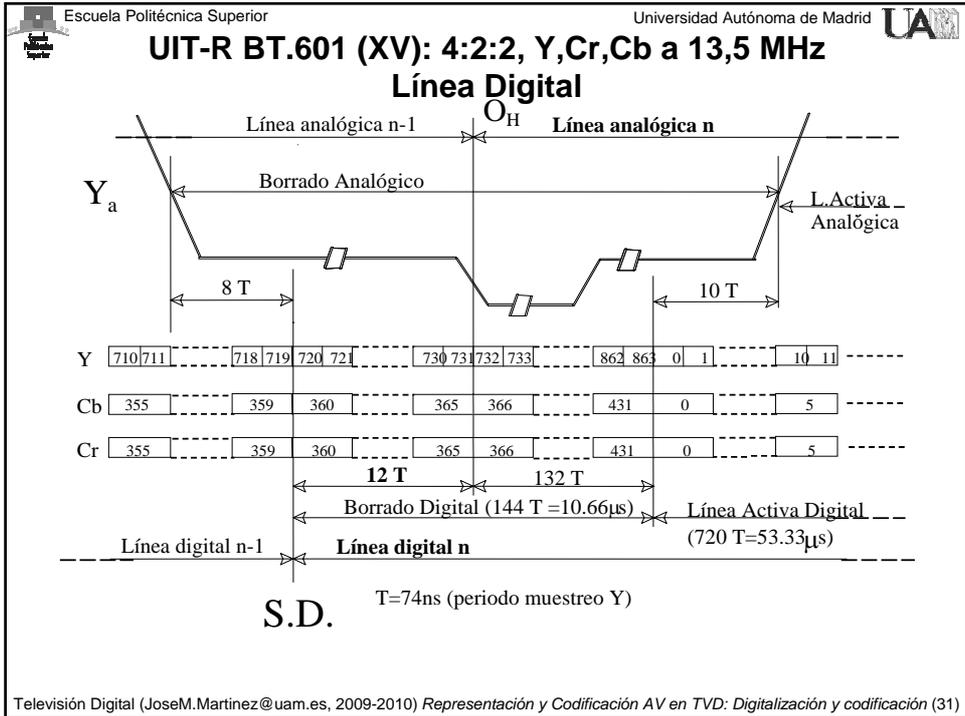
Este periodo se llama Línea Activa Digital (LAD)

Periodo de borrado digital 144 T =10,66 μs

- La línea digital comienza **12 T** antes de ISH (OH)

Línea digital

- 144T BD + (10T BA + 702 T vídeo + 8T BA)



Escuela Politécnica Superior Universidad Autónoma de Madrid

UIT-R BT.601 (XVI): 4:2:2, Y,Cr,Cb a 13,5 MHz

Cuadro Digital

		Digital line numbers			
		525 lines	625 lines		
Digital field blanking (V = 1)		4	1		
		9	22		
Digital active field (V = 0)		10	23	Digital field 1 (F = 0)	
		263	310		
		264	311		
Digital field blanking (V = 1)		265	312		
		266	313		
		272	335		
Digital active field (V = 0)		273	336	Digital field 2 (F = 1)	
		525	623		
Digital field blanking (V = 1)		1	624		
		2	625		
		3			

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (32)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.601 (XVII): 4:2:2, Y,Cr,Cb a 13,5 MHz

Cuantificación y Codificación

Uniforme con 256 ó 1024 niveles para codificar con 8 ó 10 bits

No se aprovechan todos los niveles Q

- Se deja margen para sobreniveles (saturación)
 - o Tanto por debajo del mínimo como por encima del máximo
- Se reservan para señalización el primer nivel (0) y el último (255 ó 1023, todos 1's en binario)

La norma considera señales de 1 V de excursión de amplitud

En la práctica 0.7V de excursión: **AES/UNE N-10**

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (33)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.601 (XVIII): 4:2:2, Y,Cr,Cb a 13,5 MHz

Niveles de Cuantificación en Y (*)

Niveles Q	8 bits	10 bits	Tensión
Reservado	255 = FFh	1020 a 1023	
Sobreniveles	236 a 254	941 a 1019	Saturación
Blanco	235	940	1 / 0.7 V
nº útiles	220	877	
Negro	16 = 10h	64	0 V
Sobreniveles	1 a 15	4 a 63	< 0 V
Reservado	0 = 00h	0 a 3	

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

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.601 (XIX): 4:2:2, Y,Cr,Cb a 13,5 MHz Cálculo de Niveles Q de Y (*)

Fórmula para norma (Y de 0 a 1V), 8 bits:

- $Y_q = \text{redondeo}(219 Y_{am} + 16)$

Fórmula para norma **N-10** (700mV), 8 bits:

- $Y_q = \text{redondeo}(313 Y_{am} + 16)$

Otras para 10 bits
 Y_{am} señal analógica muestreada
 Y_q señal cuantificada
redondeo() entero más próximo

Relleno con $16 = 10h$

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (35)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.601 (XX): 4:2:2, Y,Cr,Cb a 13,5 MHz Niveles de Cuantificación en Cr y Cb (*)

Niveles Q	8 bits	10 bits	Tensión
Reservado	255 = FFh	1020 a 1023	
Sobreniveles	241 a 254	961 a 1019	saturación>
Máximo	240	960	0.5 / 0.35 V
nº útiles	225	897	
Sin color	128 = 80h	512	0 V
Mínimo	16 = 10h	64	-0.5/- 0.35 V
Sobreniveles	1 a 15	4 a 63	saturación<
Reservado	0 = 00h	0 a 3	

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (36)



UIT-R BT.601 (XXI): 4:2:2, Y,Cr,Cb a 13,5 MHz Cálculo de Niveles Q de Cr y Cb (*)

Fórmula para norma (de -0,5 a 0,5V), 8 bits:

- $Cr_q = \text{redondeo}(224 C_{ram} + 128)$
- $Cb_q = \text{redondeo}(224 C_{bam} + 128)$

Fórmula para norma **N-10** ($\pm 350\text{mV}$), 8 bits:

- $Cr_q = \text{redondeo}(320 C_{ram} + 128)$
- $Cb_q = \text{redondeo}(320 C_{bam} + 128)$

C_{xam} señal analógica muestreada

C_{xq} señal cuantificada

redondeo() entero más próximo

Otras para 10 bits

Relleno con $128 = 80h$



UIT-R BT.601 (XXII): 4:2:2 Anamórfico 16:9 a 13,5 MHz, 4:2:2, Y,Cr,Cb (*)

Igual salvo en relación de aspecto

Frecuencia de muestreo la misma

- Y con 13,5 MHz
- Cr y Cb con 6,75 MHz

Mismos filtros \Rightarrow mismo ancho de banda

Mismo número de muestras en horizontal

Diferente relación de aspecto:

- Pantalla más ancha
- Pérdida de resolución horizontal
- Píxeles "estirados" horizontalmente: **anamórfico**



UIT-R BT.601 (XXIII): 4:2:2 Muestreo extendido 18 MHz, 16:9, 4:2:2, Y,Cr,Cb (*)

Frecuencia de muestreo principal más alta para mantener calidad con mayor relación de aspecto:

- Luminancia $f_m=18$ MHz \Rightarrow fc Nyquist 9 MHz
- Crominancias $f_m/2=9$ MHz \Rightarrow fc Nyquist 4,5 MHz

Periodo de muestreo de Y: $T=55,5$ ns

Periodo de diferencias de color: $2T=111,1$ ns

- Muestras de Cr y Cb **cosituadas** con impares de Y

Número entero de muestras por línea

- Salen 1152 muestras de Y y 576 de cada crominancia
- Se repite también en cada línea



UIT-R BT.601 (XXIV): 4:2:2 Muestreo extendido Temporización (*)

Se consideran muestras de vídeo válidas durante

$960 T = 53,33 \mu s > 52 \mu s$ de analógico

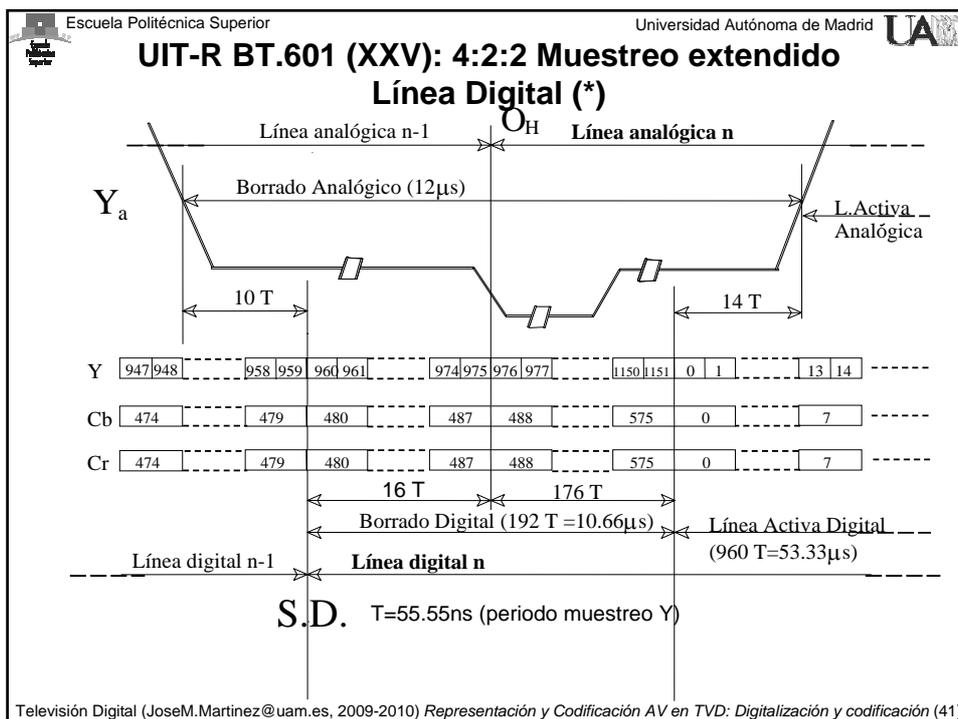
- **960** muestras de Y, **480** muestras de Cr y **480** de Cb

Línea Activa Digital (LAD) con más muestras

Periodo de borrado digital $192 T=10,66 \mu s$

- La línea digital comienza **16 T** antes de ISH (OH)

Por otro lado: misma cuantificación y codificación



Escuela Politécnica Superior Universidad Autónoma de Madrid

Representación y Codificación de Señales Audiovisuales en Televisión Digital

Digitalización y Codificación de señales de vídeo

- *Introducción*
- *Digitalización de señales de vídeo*
 - *Introducción*
 - *UIT-R BT.601*
- **Multiplexación y Codificación de la señal de vídeo digital**
 - **Introducción**
 - UIT-R BT.656 y 1302
 - Formato de señal común
 - Interfaz paralelo
 - Interfaz serie

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (42)



Multiplexación y codificación de vídeo digital (I): Introducción

La norma UIT-R BT.601 genera señales de vídeo digital con valores comunes para sistemas de 525 y 625 líneas

Sin embargo no define los interfaces y el tren de datos (multiplexados) de ese vídeo digital codificado

Las normas UIT-R BT.656 y UIT-R BT.1302 describen el medio de interconexión de equipos (interfaces) de vídeo digital que funcionen conforme a la norma UIT-R BT.601

- Formato de señal común a las interfaces serie y paralelo
- Interfaz para bits en paralelo
- Interfaz para bits en serie



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- *Introducción*
- *Digitalización de señales de vídeo*
 - *Introducción*
 - *UIT-R BT.601*
- **Multiplexación y Codificación de la señal de vídeo digital**
 - *Introducción*
 - **UIT-R BT.656 y 1302**
 - Formato de señal común
 - Interfaz paralelo
 - Interfaz serie



UIT-R BT.656 y BT.1302 (I): Introducción

Las normas UIT-R BT.656 y UIT-R BT.1302 describen el medio de interconexión de equipos (interfaces) de vídeo digital que funcionen conforme a la norma UIT-R BT.601

- Formato de señal común a las interfaces serie y paralelo
- Interfaz para bits en paralelo
- Interfaz para bits en serie

- Las diferencias entre ambas radican en que la UIT-R BT.656 aplica a señales UIT-R BT.601 con muestreo convencional o anamórfico (13,5 MHz), mientras que UIT-R BT.1302 aplica a señales con muestreo extendido (18 MHz)



UIT-R BT.656 y 1302 (II): Formato de señal común Múltiplex de Vídeo Digital

Datos de Vídeo (resultado de la digitalización)

- Multiplexado de datos de LAD
- Muestras en intervalos de borrado H y V \Rightarrow se desechan

Datos de Sincronización.

- Sincronismos Horizontal y Vertical

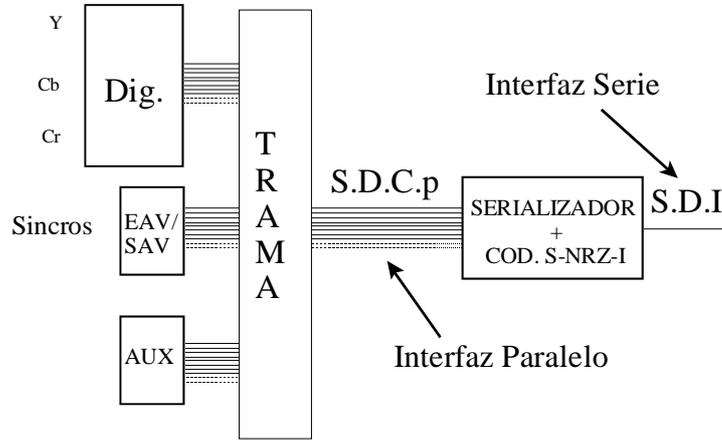
Datos Auxiliares

- Otras informaciones \Rightarrow no vídeo

Datos de Relleno.

- Cuando no hay información que transmitir

UIT-R BT.656 y 1302 (III): Formato de señal común Esquema General de codificación

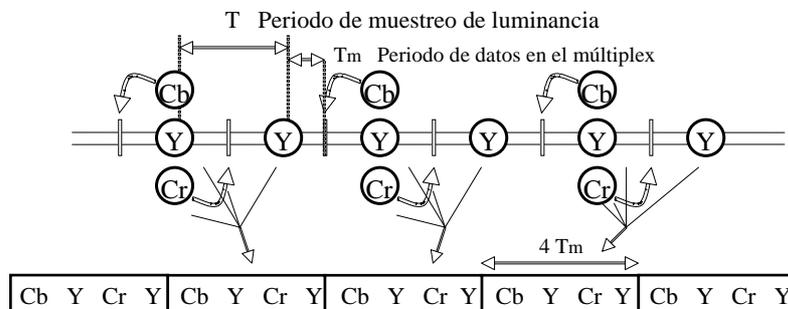


Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (47)

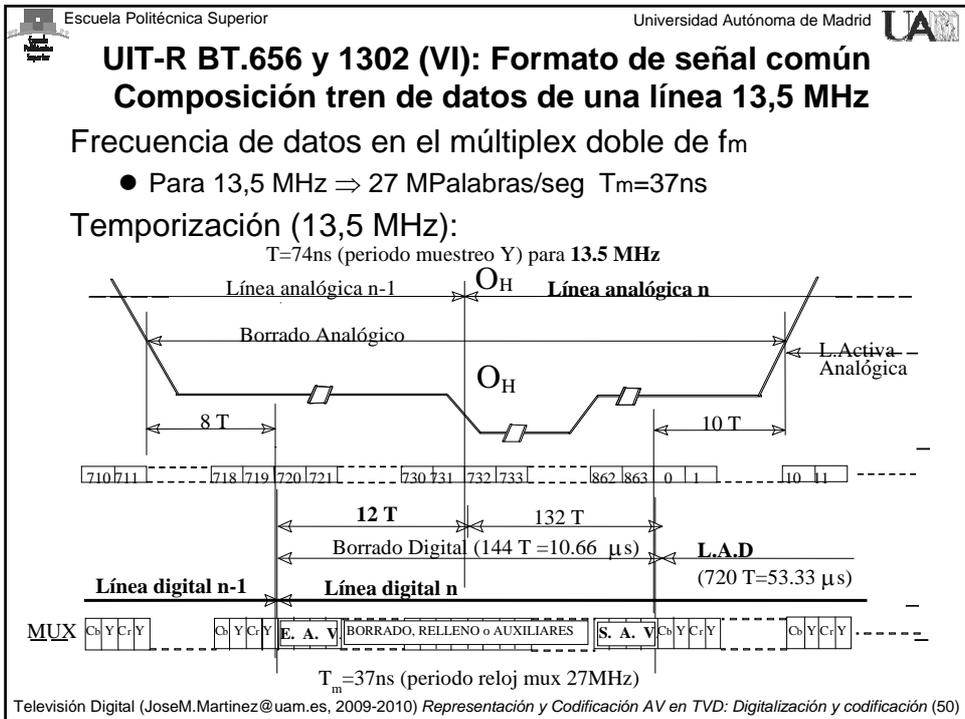
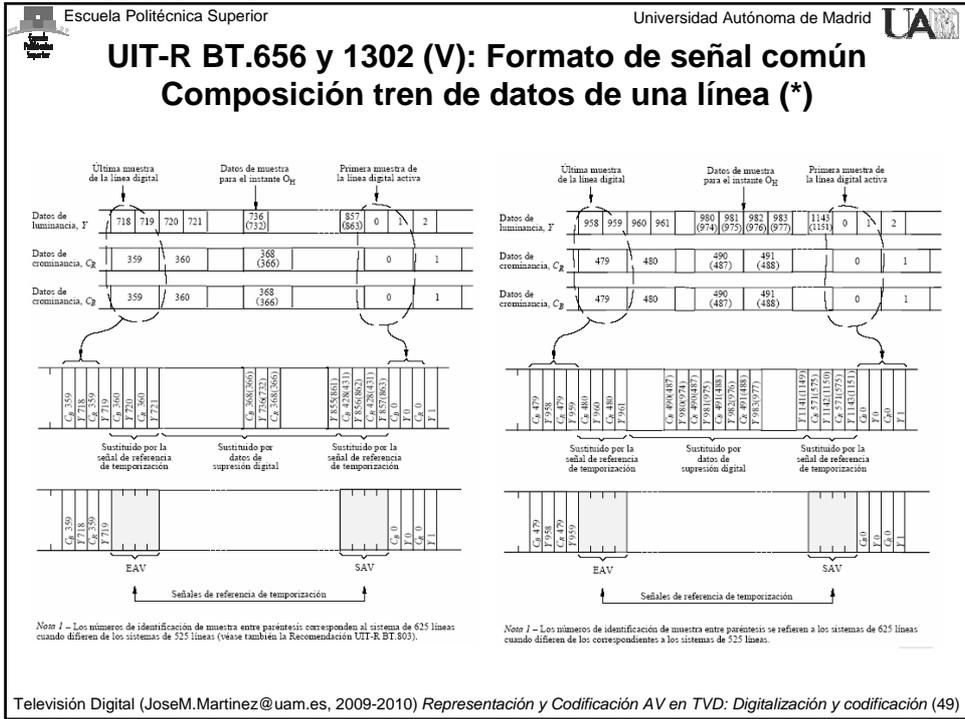
UIT-R BT.656 y 1302 (IV): Formato de señal común Multiplexación de Datos de LAD

Muestreo de Y a f_m y de Cr y Cb a $f_m/2$ (4:2:2)

- Cada dos palabras/datos de Y \Rightarrow Un dato de Cr y un dato de Cb
o Palabra = {8,10} bits
- Cada $2T$ ($T=2 T_m$) se transmiten 4 palabras en el orden que muestra la figura



Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (48)

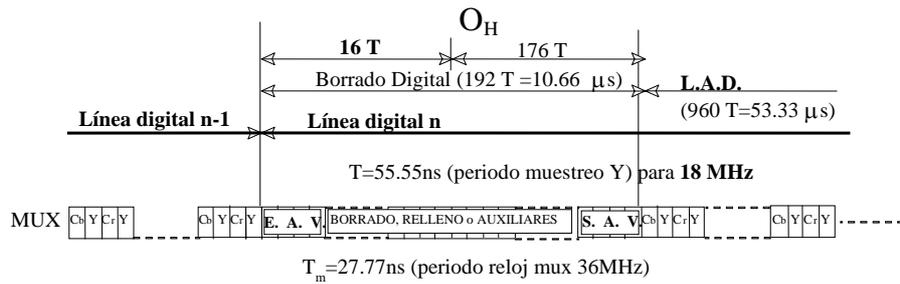


UIT-R BT.656 y 1302 (VII): Formato de señal común Composición tren de datos de una línea 18 MHz(*)

Frecuencia de datos en el múltiplex doble de f_m

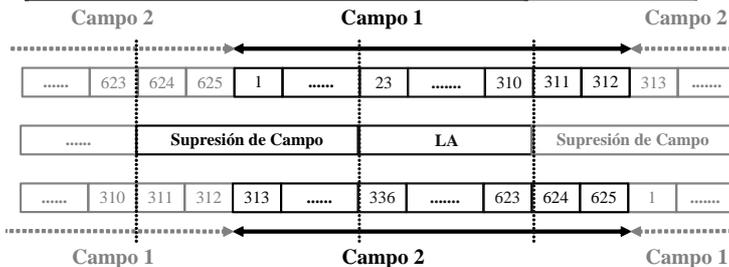
- Para 18,0 MHz \Rightarrow 36 MPalabras/seg $T_m=27,77ns$

Temporización (18 MHz):



UIT-R BT.656 y 1302 (VIII): Formato de señal común Líneas de un cuadro de vídeo digital

	Líneas	Nº en Campo	Nº en Cuadro
Campo 1			
Total	1 a 312	312	625
Líneas Activas	23 a 310	288	576
Borrado Vertical	624 a 22	24	49
Campo 2			
Total	313 a 625	313	
Líneas Activas	336 a 623	288	
Borrado Vertical	311 a 335	25	





UIT-R BT.656 y 1302 (IX): Formato de señal común

Palabras de datos y sincronismo en una línea

$f_m = 13,5 \text{ MHz}$:

- Datos totales en una Línea
 - $864 (Y) + 2 \times 432 (Cr \text{ y } Cb) = 1728$
- Datos en LAD
 - $720 (Y) + 2 \times 360 (Cr \text{ y } Cb) = 1440$
- Datos durante el Borrado horizontal
 - $1728 - 1440 = 288 = 280 + 8 (EAV + SAV)$
 - Se sustituyen, no se usan los muestreados

$f_m = 18,0 \text{ MHz}$

- Datos totales en una Línea
 - $1152 (Y) + 2 \times 576 (Cr \text{ y } Cb) = 2304$
- Datos en LAD
 - $960 (Y) + 2 \times 480 (Cr \text{ y } Cb) = 1920$
- Datos durante el Borrado horizontal
 - $2304 - 1920 = 384 = 376 + 8 (EAV + SAV)$
 - Se sustituyen, no se usan los muestreados

Dos grupos de 4 palabras para sincronización

- **EAV** (End of Active Vídeo) final de LAD
- **SAV** (Start of Active Vídeo) principio de LAD



UIT-R BT.656 y 1302 (X): Formato de señal común

Códigos de sincronismo de línea

Los Códigos de sincronismo (4 palabras) van en todas las líneas (incluso las de borrado vertical)

- **EAV** (End of Active Vídeo)
 - **Final de LAD** y de la **Línea Digital**
 - **Comienzo de Línea** y borrado horizontal
- **SAV** (Start of Active Vídeo) principio de LAD
 - Final del periodo de borrado y **comienzo de LAD**
 - Las palabras en LAD se numeran de 0 en adelante
- Se utilizan como **sincronismos** horizontal y vertical
- Formato FF.00.00.XY
 - En XY se definen 3 bits de información (F,V,H) y 4 de protección frente a errores (Pi)
 - Pi genera un código (7,3,2,1)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.656 y 1302 (XI): Formato de señal común Códigos de sincronismo de línea

Número del bit de datos	Primera palabra (FF)	Segunda palabra (00)	Tercera palabra (00)	Cuarta palabra (XY)
9 (MSB)	1	0	0	1
8	1	0	0	F
7	1	0	0	V
6	1	0	0	H
5	1	0	0	P ₃
4	1	0	0	P ₂
3	1	0	0	P ₁
2	1	0	0	P ₀
1 (véase la Nota 2)	1	0	0	0
0	1	0	0	0

NOTA 1 – Los valores indicados son los recomendados para las interfaces de 10 bits.
 NOTA 2 – Para la compatibilidad con las interfaces existentes de 8 bits, no se definen los valores de los bits D₁ y D₀.

F = 0 durante la trama 1
 F = 1 durante la trama 2
 V = 0 fuera de la supresión de trama vertical
 V = 1 durante la supresión de trama vertical
 H = 0 en SAV
 H = 1 en EAV

F	V	H	P ₃	P ₂	P ₁	P ₀
0	0	0	0	0	0	0
0	0	1	1	1	0	1
0	1	0	1	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	1	1
1	0	1	1	0	1	0
1	1	0	1	1	0	0
1	1	1	0	0	0	1

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (55)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

UIT-R BT.656 y 1302 (XII): Formato de señal común Datos Auxiliares 'HANC'

Situados en la zona de borrado digital horizontal

- Entre EAV y SAV de línea digital

Posible contenido:

- Relleno con Nivel de Borrado [Cb Y Cr Y] (palabras de supresión)
 - o Señal sin vídeo (Cb Y Cr Y): **80h 10h 80h 10h: 128 16 128 16**
- Datos Auxiliares (según norma UIT-R BT.1364)
 - o "Audio Embedded" (hasta 4 canales de audio digital)
 - o Otros (**EDH**: Error Detection and Handling, If.Serie)
- En total son **280** (376) Palabras

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (56)



UIT-R BT.656 y 1302 (XIII): Formato de señal común Datos Auxiliares 'VANC'

Situados en la zona de borrado digital vertical (de campo)

- Cambio de Bit de Campo en el EAV

Durante las líneas de supresión de campo

- Uso de LAD para datos auxiliares llamados 'VANC'
- 2 líneas por cuadro reservadas (códigos de tiempo)

Capacidad Datos auxiliares en Borrado Vertical

- 1440 palabras/LAD x 49 líneas/cuadro x 25 cuadros/s =
- 1,764 MPalabras/s ! (13,5 MHz)

Datos auxiliares todos en formato normalizado

- Paquetes con estructura definida (UIT-R BT.1364)



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- *Introducción*
- *Digitalización de señales de vídeo*
 - o *Introducción*
 - o *UIT-R BT.601*
- **Multiplexación y Codificación de la señal de vídeo digital**
 - o *Introducción*
 - o **UIT-R BT.656 y 1302**
 - Formato de señal común*
 - Interfaz paralelo**
 - Interfaz serie

Escuela Politécnica Superior		Universidad Autónoma de Madrid	
 UIT-R BT.656 y 1302 (XIV): Interfaz Digital Paralelo Interfaz eléctrico (I) (*)			
Reloj	27 MHz	36 MHz	
Cable Conexión	8/10 pares balanceados para datos 1 par balanceado para reloj síncrono		
Alcance	50 m sin ecualizar 300m ecualizado	40 m sin ecualizar 160/225m ecualiz.	
Conector	Sub D 25 pines		
Código de Forma de Onda	NRZ-L bipolar (Non Return to Zero) Nivel positivo para '1' lógico Nivel negativo para '0' lógico		
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (59)			

Escuela Politécnica Superior		Universidad Autónoma de Madrid	
 UIT-R BT.656 y 1302 (XV): Interfaz Digital Paralelo Interfaz eléctrico (II) (*)			
Señal de Reloj	Onda Cuadrada. Activación por flanco a positivo: en punto medio de los datos		
	27 MHz Ancho $18,5 \pm 3$ ns Periodo: 37 ns	36 MHz Ancho $13,9 \pm 2$ ns Periodo: 27,8 ns	
Tipo Conexión	Balanceada unidireccional Interfaz eléctrico simétrico		
Zout	Impedancia de salida: 110 Ω máxima		
Amplitud señal de salida	0,8 Vpp a 2,0 Vpp (sobre 110 Ω)		
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (60)			



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- *Introducción*
- *Digitalización de señales de vídeo*
 - *Introducción*
 - *UIT-R BT.601*
- **Multiplexación y Codificación de la señal de vídeo digital**
 - *Introducción*
 - **UIT-R BT.656 y 1302**
 - Formato de señal común*
 - Interfaz paralelo*
 - Interfaz serie**



UIT-R BT.656 y 1302 (XVI): Interfaz Digital Serie Señal SDI

Estándar más utilizado

- Conexión de equipos de vídeo digital
- Un único cable coaxial o fibra óptica

En el mismo canal

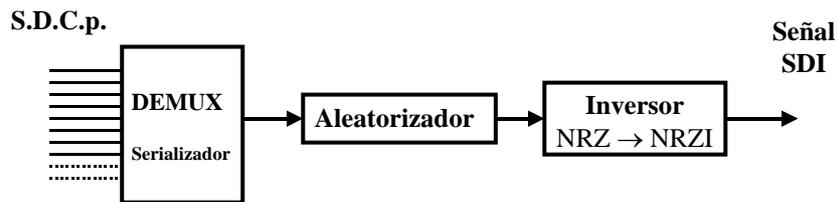
- Información: 10 bits (posiblemente 2 LSB relleno: '0')
 - Todos los datos del múltiplex digital
- Señal de Reloj: código **autosincronizable** (NRZI)



UIT-R BT.656 y 1302 (XVII): Interfaz Digital Serie Proceso de Serialización / Codificación

Proceso de Codificación: Pasos

- **1. Aleatorización:** Equilibrado estadístico 0's y 1's y aumento del número de transiciones
- **2. Inversión:** Información en transiciones de polaridad



UIT-R BT.656 y 1302 (XVIII): Interfaz Digital Serie Aleatorización (Scrambling) (*)

Señal de partida: NRZ-L (Primero LSB)

- Desequilibrio entre número de 0's y 1's
- Amplio margen de frecuencias

Objetivos

- Disminuir componente continua de la señal
- Estrechar el ancho de banda que ocupa

Proceso con generador pseudoaleatorio

- Iguala nº de bits de cada valor (0 ó 1)
- Estrecha margen de frecuencias (mejora equalización)

- Polinomio: $A(x) = x^9 + x^4 + 1$



UIT-R BT.656 y 1302 (XIX): Interfaz Digital Serie Inversión (*)

Objetivo: Incluir una señal de reloj en los datos e independizar de la polaridad

Conversión de señal NRZ a codificación NRZI

- Non Return to Zero Inverted (Bipolar)

Operación con polinomio

$$I(x) = x + 1$$

Resultado

- Codificación de 0: la señal mantiene su nivel
- Codificación de 1: la señal cambia de nivel \Rightarrow utilizado para obtener reloj en receptor



UIT-R BT.656 y 1302 (XX): Interfaz Digital Serie Transmisión Señal SDI – Problemas (*)

Respuesta en frecuencia (del cable o fibra)

- **Distorsión** en amplitud y fase

Desplazamiento de continua

Desadaptación de impedancias

Problemas interpretación información digital

- Interferencia entre símbolos
- Transiciones imperfectas
 - o Tiempos de subida
 - o Fluctuaciones de fase del reloj ('Jitter')
 - o Sobreniveles ('Ringing')

Escuela Politécnica Superior Universidad Autónoma de Madrid

UIT-R BT.656 y 1302 (XXI): Interfaz Digital Serie Proceso de Recuperación

Ecualización para corregir problemas del canal

- Aún así puede haber errores

```

    graph LR
      SDI[Señal SDI] --> Ecualizador[Ecualizador]
      Ecualizador --> Inversor[Inversor NRZI -> NRZ]
      Recuperacion[Recuperación señal de reloj] --> Inversor
      Inversor --> Desaleatorizador[Desaleatorizador]
      Desaleatorizador --> MUX[MUX]
      MUX --- SDCP[S.D.C.p.]
  
```

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (67)

Escuela Politécnica Superior Universidad Autónoma de Madrid

UIT-R BT.656 y 1302 (XXII): Interfaz Digital Serie Interfaz eléctrico (I) (*)

Reloj Serie	270 MHz $f_m=13,5$ MHz	360 MHz $f_m=18,0$ MHz
Cable de Conexión	Coaxial (impedancia 75Ω) Pérdidas retorno > 15 dB (5-360 MHz) (fibra óptica con norma 1367)	
Tipo de Conector	BNC normalizado (hasta 850 MHz en circuitos de 75Ω)	
Señal en el canal	Bipolar 800 mVpp $\pm 10\%$ Código NRZ-I sin polaridad	

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: Digitalización y codificación (68)



UIT-R BT.656 y 1302 (XXIII): Interfaz Digital Serie Interfaz eléctrico (II) (*)

Tiempos de subida y bajada (20-80%)	Entre 0.75 y 1,5 ns (diferencia entre ellos < 0,5 ns)
Fluctuación de fase	< 10% periodo de reloj
Ringing	< 10% de amplitud
Sensibilidad en Recepción	Para detectar correctamente señal con niveles estándar en conexión directa con pérdidas de 40 dB a 270 MHz y curva $1/\sqrt{f}$ desde 1MHz



Representación y Codificación de Señales Audiovisuales en Televisión Digital Digitalización y Codificación de señales de vídeo

- **Introducción**
- **Digitalización de señales de vídeo**
 - *Introducción*
 - *UIT-R BT.601*
- **Multiplexación y Codificación de la señal de vídeo digital**
 - *Introducción*
 - *UIT-R BT.656 y 1302*
 - Formato de señal común*
 - Interfaz paralelo*
 - Interfaz serie*



Referencias

- ★ Rec. UIT-R BT.601: Encoding parameters of digital television for studios (1994)
- ★ Rec. UIT-R BT.656 (1998)
- ★ Rec. UIT-R BT.1302 (1997)

Créditos

- Estas transparencias están adaptadas de las elaboradas originalmente por Enrique Rendón Angulo, Profesor Titular de la E.T.Ing.Telecomunicación de la UPM, para la asignatura Televisión.
- También se han incorporado transparencias de Yao Wang, Polytechnic University, Brooklyn



Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

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

JoseM.Martinez@uam.es
tel:+34.91.497.22.58

2009-2010



Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

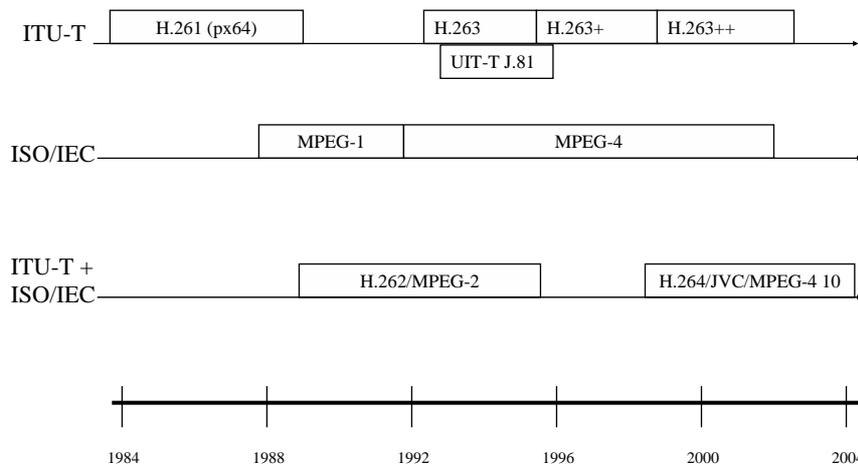
- Introducción
- Descripción del sistema.
 - Codificación de una imagen de la secuencia.
 - Detalle del codificador y del decodificador.
 - Obtención de la imagen predicción.
 - Diagrama de bloques.
- Estrategias de predicción.
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora

Introducción

El esquema de codificación híbrido transformacional es el que se ha impuesto como base de la mayor parte de los estándares

- La predicción se realiza en el dominio de la imagen (espacial).
 - o Generalmente se usa predicción temporal con compensación de movimiento
- La transformación se aplica sobre el error de predicción.
 - o Se suele usar la DCT

Introducción: Estándares de Codificación con esquema híbrido

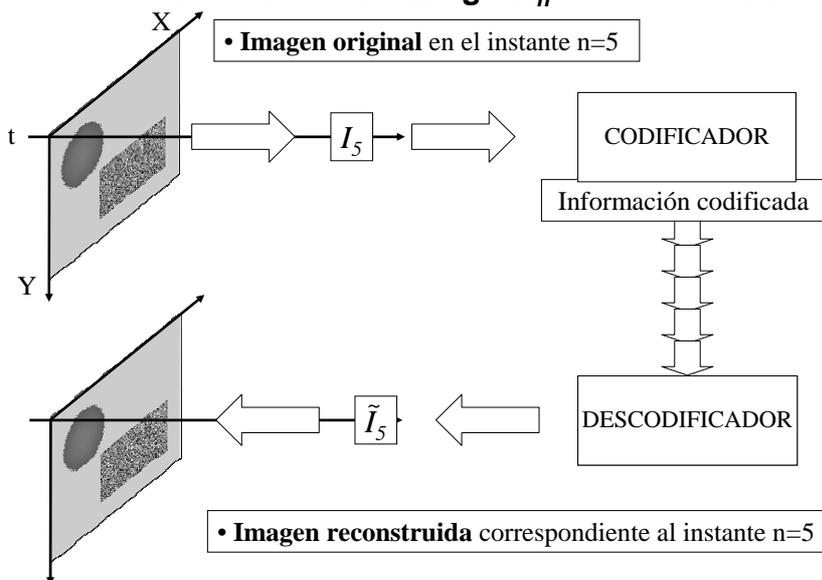


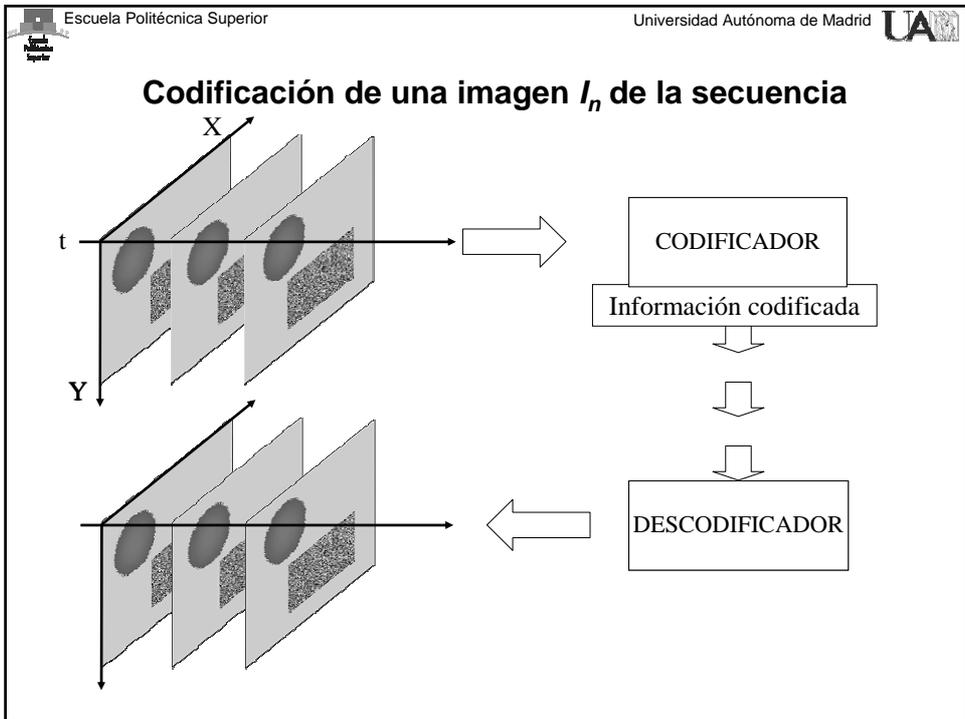
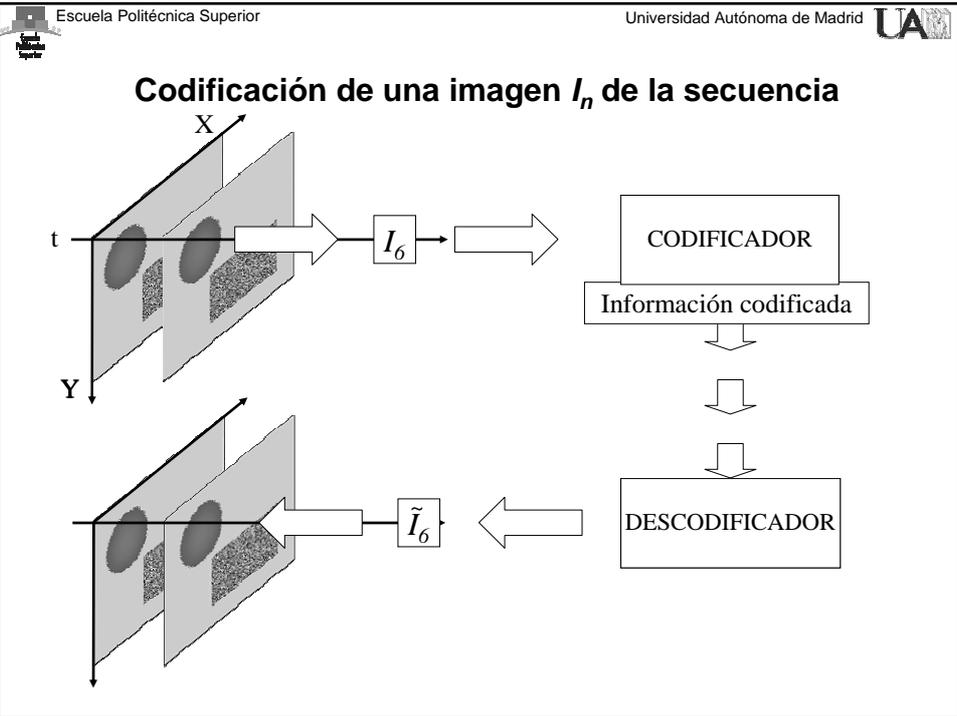
Representación y Codificación de Señales Audiovisuales en Televisión Digital

Compresión de vídeo híbrido transformacional

- *Introducción*
- **Descripción del sistema.**
 - **Codificación de una imagen de la secuencia.**
 - Detalle del codificador y del decodificador.
 - Obtención de la imagen predicción.
 - Diagrama de bloques.
- Estrategias de predicción.
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora

Codificación de una imagen I_n de la secuencia



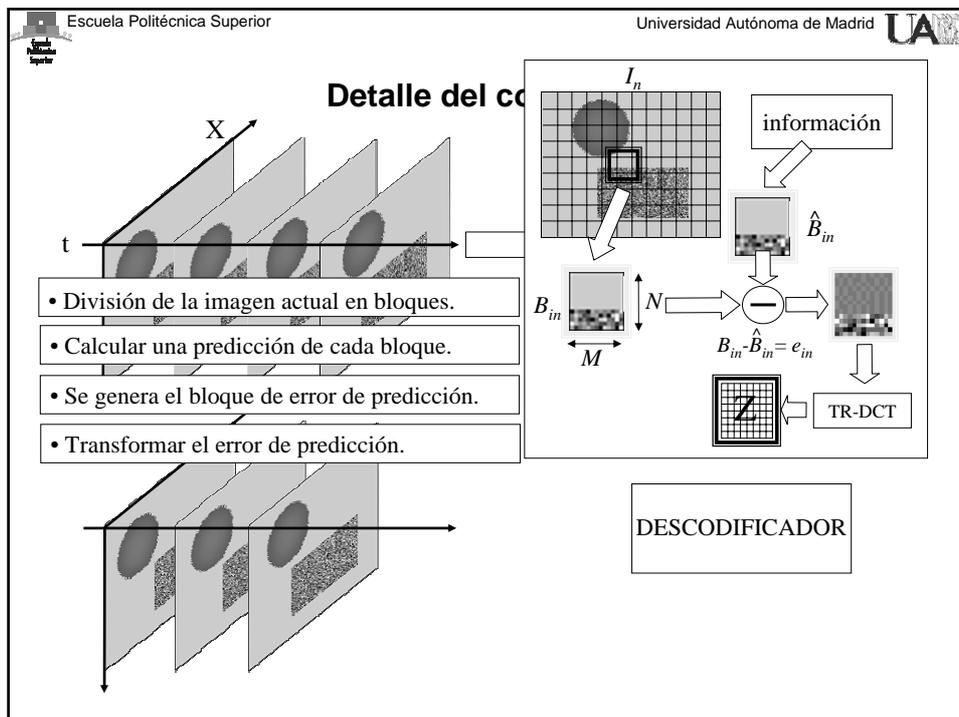


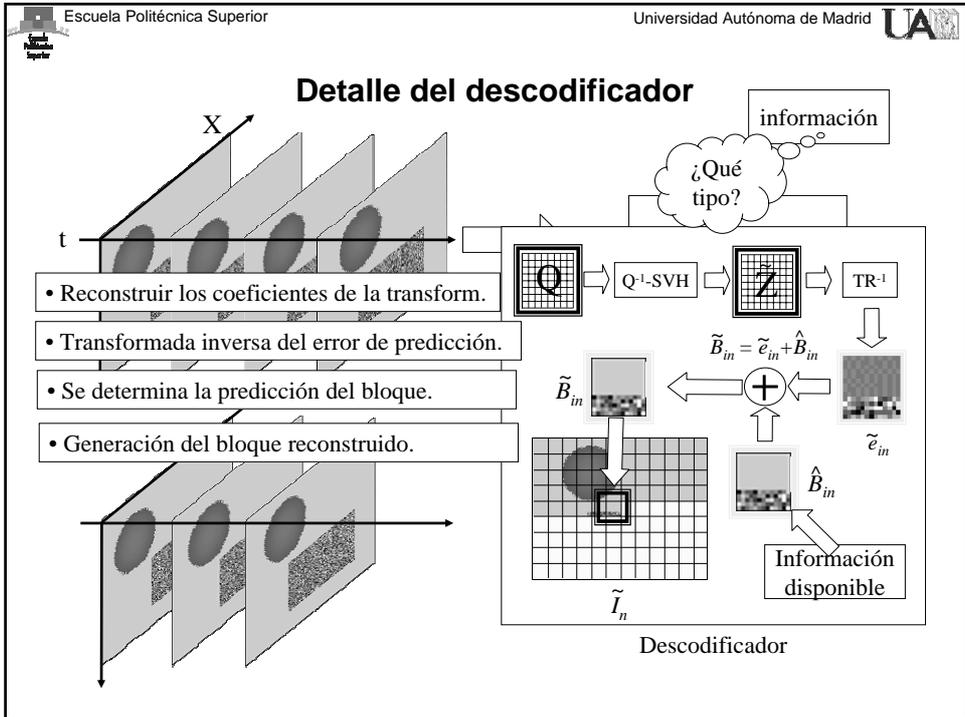
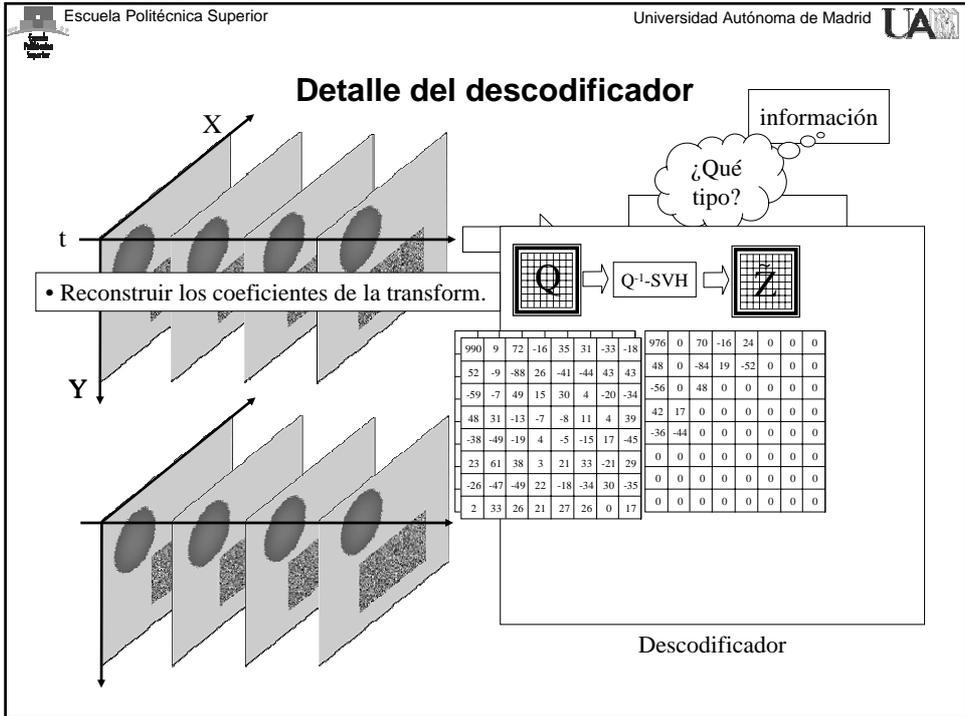
Representación y Codificación de Señales Audiovisuales en Televisión Digital

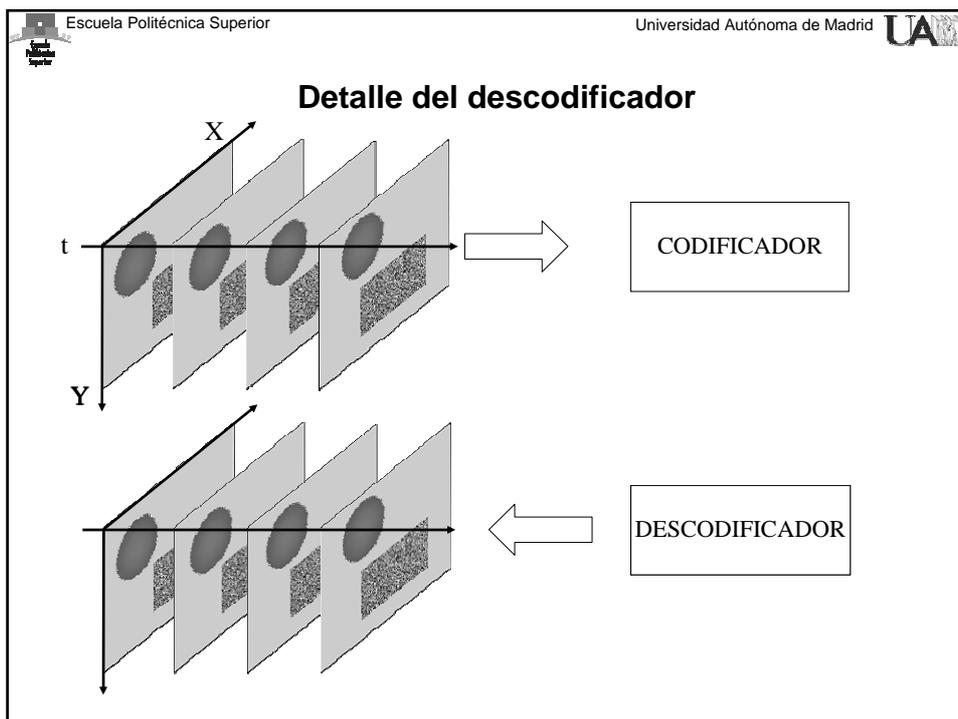
Compresión de vídeo híbrido transformacional

- *Introducción*
- **Descripción del sistema.**
 - *Codificación de una imagen de la secuencia.*
 - **Detalle del codificador y del decodificador.**
 - Obtención de la imagen predicción.
 - Diagrama de bloques.
- Estrategias de predicción.
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora

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







Escuela Politécnica Superior Universidad Autónoma de Madrid

Representación y Codificación de Señales Audiovisuales en Televisión Digital

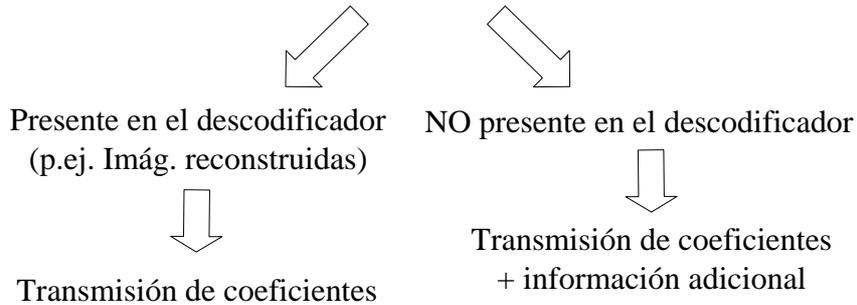
Compresión de vídeo híbrido transformacional

- *Introducción*
- **Descripción del sistema.**
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del descodificador.*
 - **Obtención de la imagen predicción.**
 - Diagrama de bloques.
- Estrategias de predicción.
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: cod. híbrido transformacional (16)

Obtención de la imagen predicción

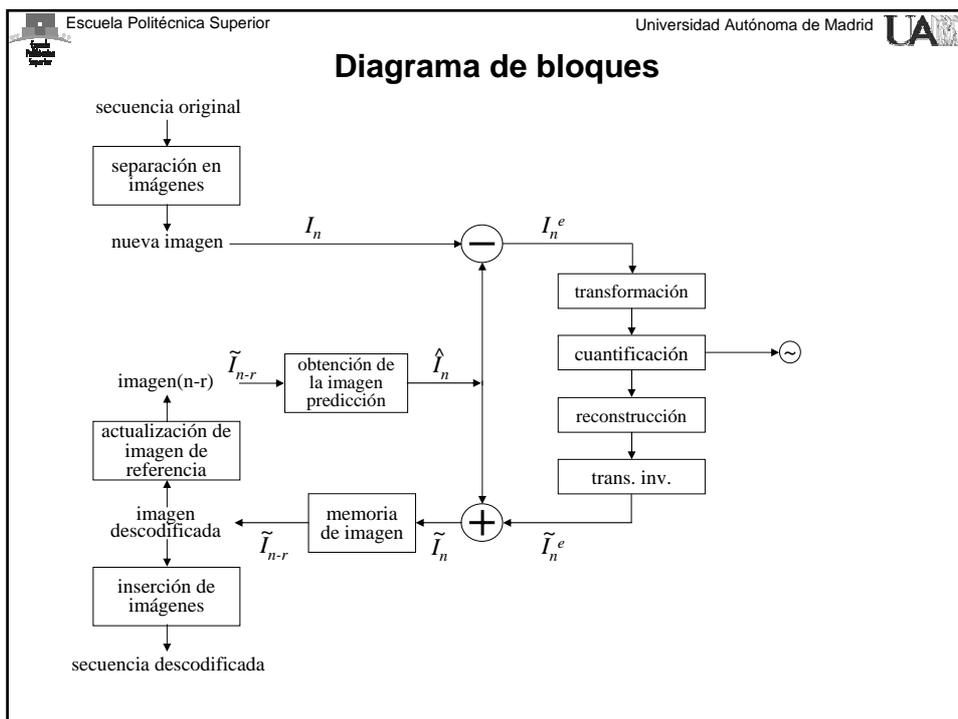
La unión de los bloques de predicción $\{\hat{B}_{in}\} \Rightarrow \hat{I}_n$ Imagen Predicción
 Predicción generada a partir de información



El codificador incorpora el decodificador
 con objeto de obtener la información a partir de la cual generar la predicción

Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- **Descripción del sistema.**
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - **Diagrama de bloques.**
- Estrategias de predicción.
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora



Escuela Politécnica Superior Universidad Autónoma de Madrid

Representación y Codificación de Señales Audiovisuales en Televisión Digital

Compresión de vídeo híbrido transformacional

- **Introducción**
- **Descripción del sistema.**
 - Codificación de una imagen de la secuencia.
 - Detalle del codificador y del descodificador.
 - Obtención de la imagen predicción.
 - Diagrama de bloques.
- **Estrategias de predicción.**
 - **Introducción**
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- **Sistemas de contención.**
- **Errores de transmisión y necesidad de refresco.**
- **Líneas de mejora**

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: cod. híbrido transformacional (20)

Introducción

Considerando imágenes estáticas, éstas contienen una gran cantidad de información redundante. Se han presentado:

- Estrategias para tratar de eliminar esta redundancia.
- Entre ellas, los esquemas predictivos analizados generan la predicción de los valores de los píxeles de la imagen a partir de píxeles próximos espacialmente.



píxeles próximos → píxeles similares

Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- **Estrategias de predicción.**
 - *Introducción*
 - **Predicción temporal.**
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- Sistemas de contención.
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora

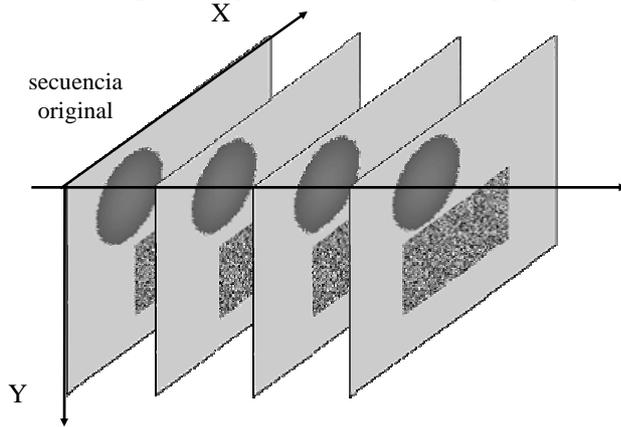


Predicción temporal

Extensión directa al tratamiento de secuencias



píxeles próximos en el tiempo → píxeles similares

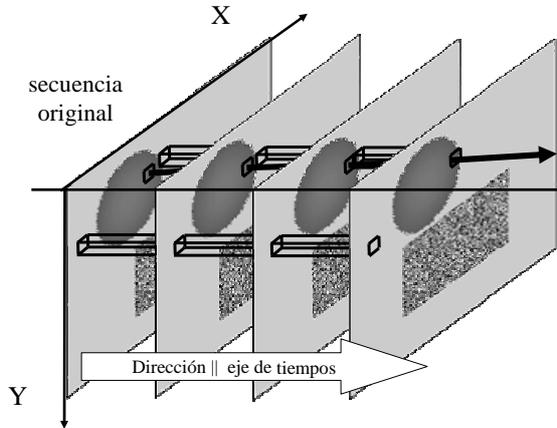


Predicción temporal

Extensión directa al tratamiento de secuencias



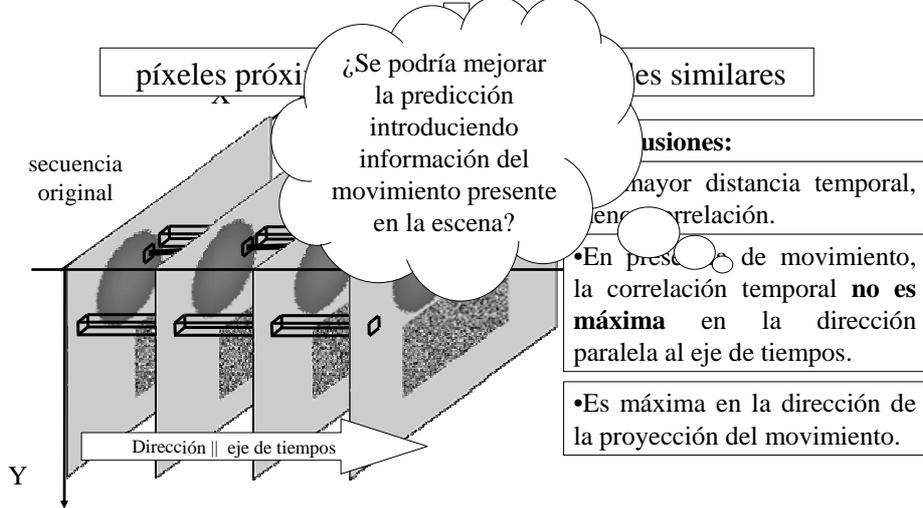
píxeles próximos en el tiempo → píxeles similares



Conclusiones:
• A mayor distancia temporal, menor correlación.
• En presencia de movimiento, la correlación temporal no es máxima en la dirección paralela al eje de tiempos.
• Es máxima en la dirección de la proyección del movimiento.

Predicción temporal

Extensión directa al tratamiento de secuencias



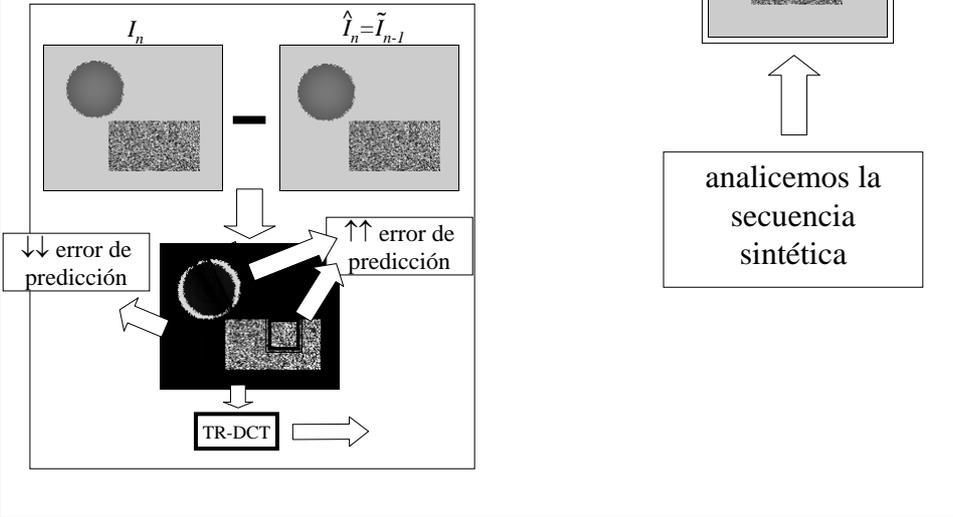
Representación y Codificación de Señales Audiovisuales en Televisión Digital

Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- **Estrategias de predicción.**
 - *Introducción*
 - *Predicción temporal.*
 - **Introducción del movimiento en el cálculo de la predicción.**
 - *Estimación y compensación de movimiento.*
 - *Métodos de Estimación de movimiento*
 - *Métodos de Compensación de movimiento*
- *Sistemas de contención.*
- *Errores de transmisión y necesidad de refresco.*
- *Líneas de mejora*

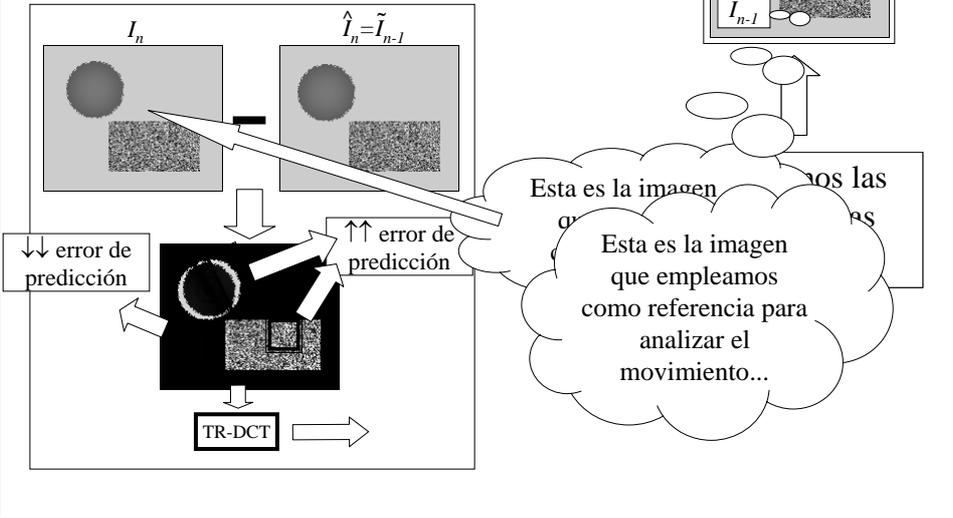
Introducción del movimiento en la predicción

Codificación sin emplear movto.

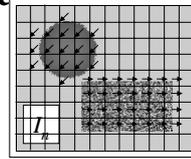
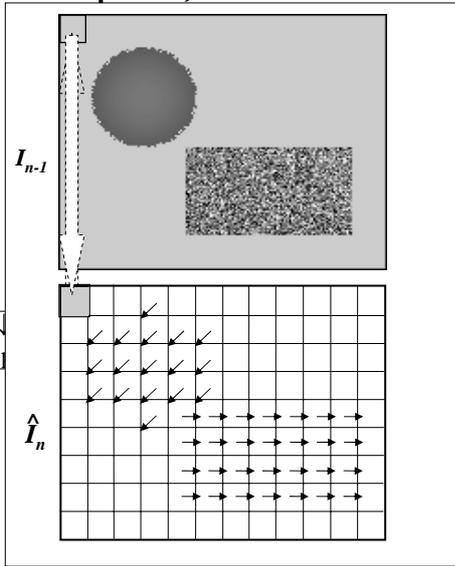


Introducción del movimiento en la predicción

Codificación sin emplear movto.



Introducción del movimiento en la



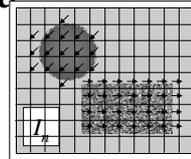
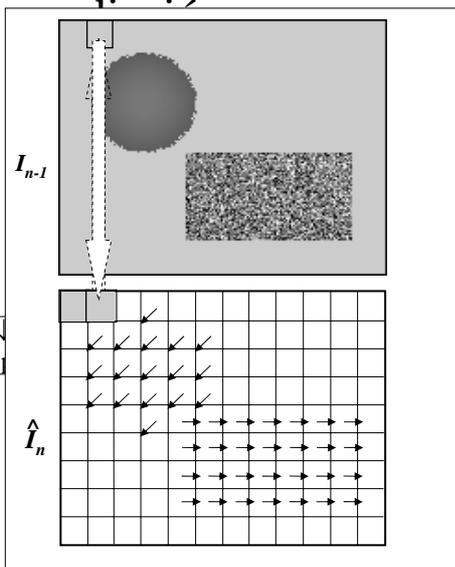
Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Introducción del movimiento en la



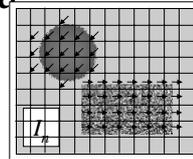
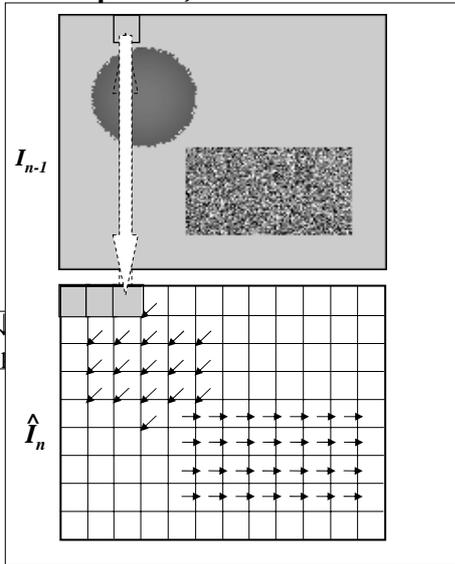
Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Introducción del movimiento en la



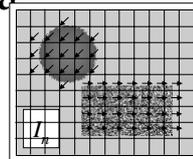
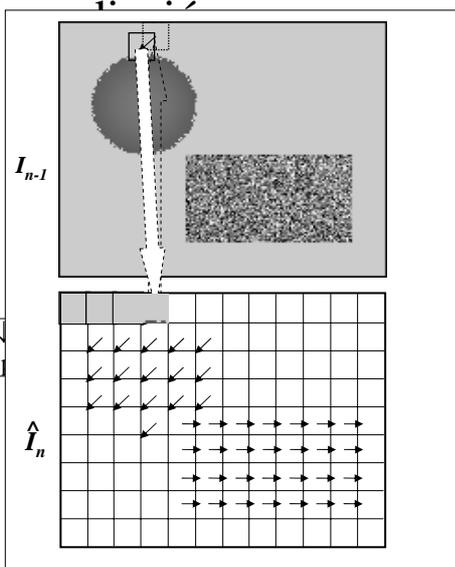
Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Introducción del movimiento en la



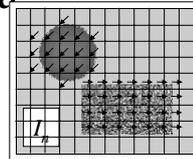
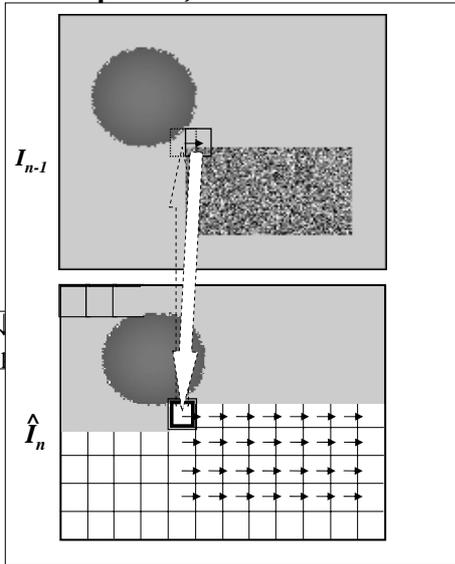
Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Introducción del movimiento en la



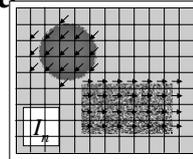
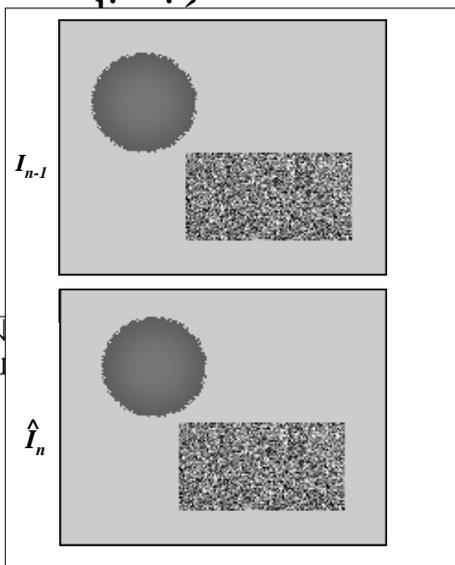
Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Introducción del movimiento en la



Supongamos conocido:

- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

vectores de movimiento.

¿Es posible generar la imagen predicción \hat{I}_n empleando para ello la información de movimiento?

Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- **Estrategias de predicción.**
 - *Introducción*
 - *Predicción temporal.*
 - *Introducción del movimiento en el cálculo de la predicción.*
 - **Estimación y compensación de movimiento.**
 - *Métodos de Estimación de movimiento*
 - *Métodos de Compensación de movimiento*
- *Sistemas de contención.*
- *Errores de transmisión y necesidad de refresco.*
- *Líneas de mejora*

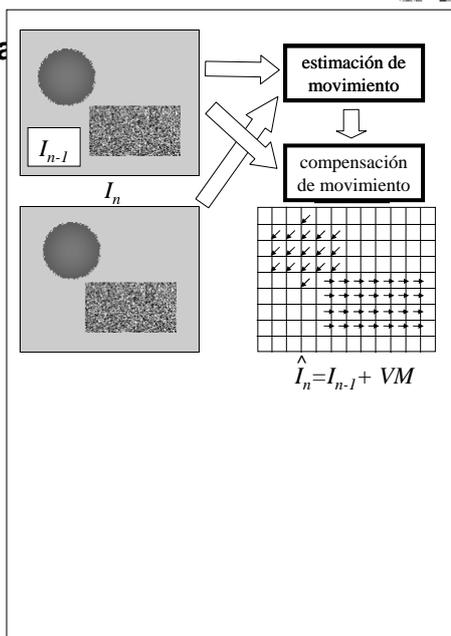
Estimación y compensación de movimiento

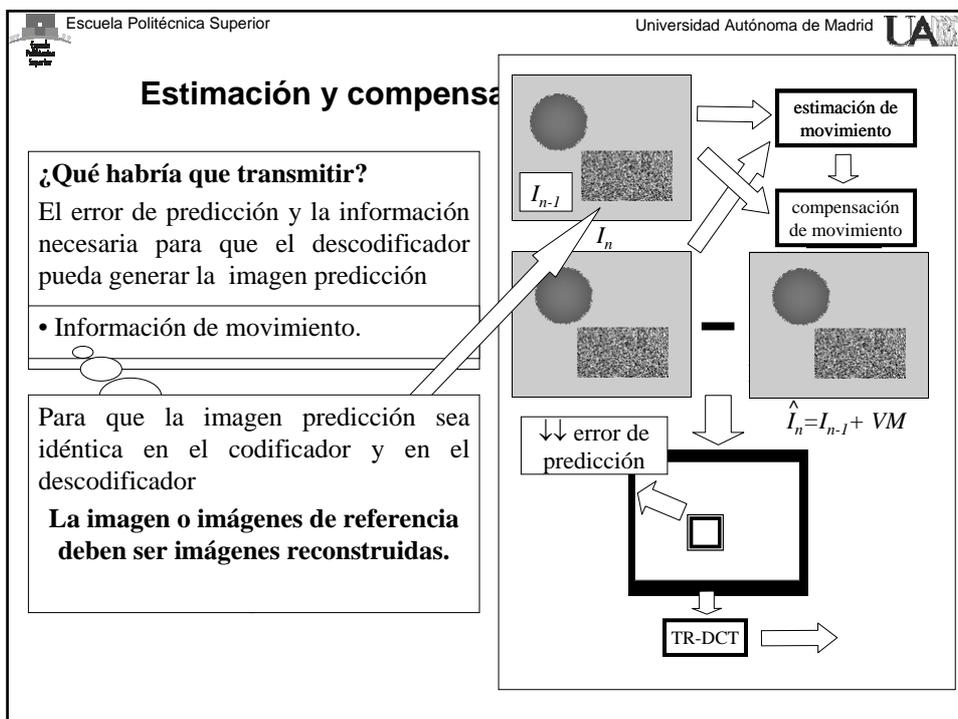
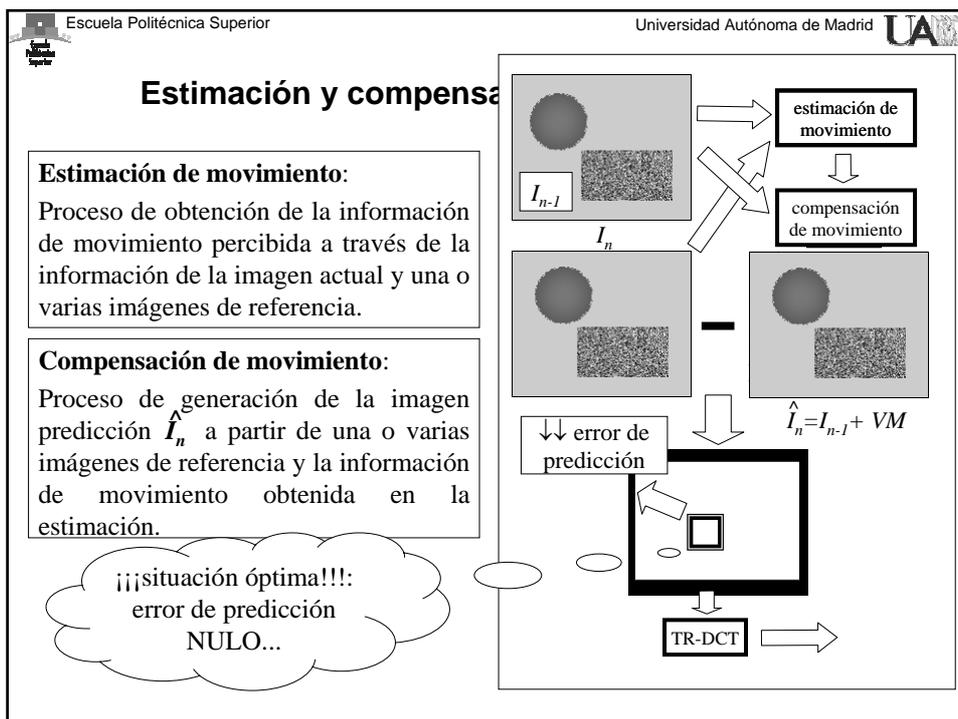
Estimación de movimiento:

Proceso de obtención de la información de movimiento percibida a través de la información de la imagen actual y una o varias imágenes de referencia.

Compensación de movimiento:

Proceso de generación de la imagen predicción \hat{I}_n a partir de una o varias imágenes de referencia y la información de movimiento obtenida en la estimación.





Escuela Politécnica Superior Universidad Autónoma de Madrid

Estimación y compensación

¿Qué habría que transmitir?
El error de predicción y la información necesaria para que el decodificador pueda generar la imagen predicción

- Información de movimiento.

Para que la imagen predicción sea idéntica en el codificador y en el decodificador
La imagen o imágenes de referencia deben ser imágenes reconstruidas.

The diagram illustrates the motion estimation and compensation process. It starts with the current frame I_n and the previous reconstructed frame \tilde{I}_{n-1} . The process involves motion estimation and motion compensation to produce a predicted frame $\hat{I}_n = I_{n-1} + VM$. The prediction error, indicated by two downward arrows, is then processed by TR-DCT.

Escuela Politécnica Superior Universidad Autónoma de Madrid

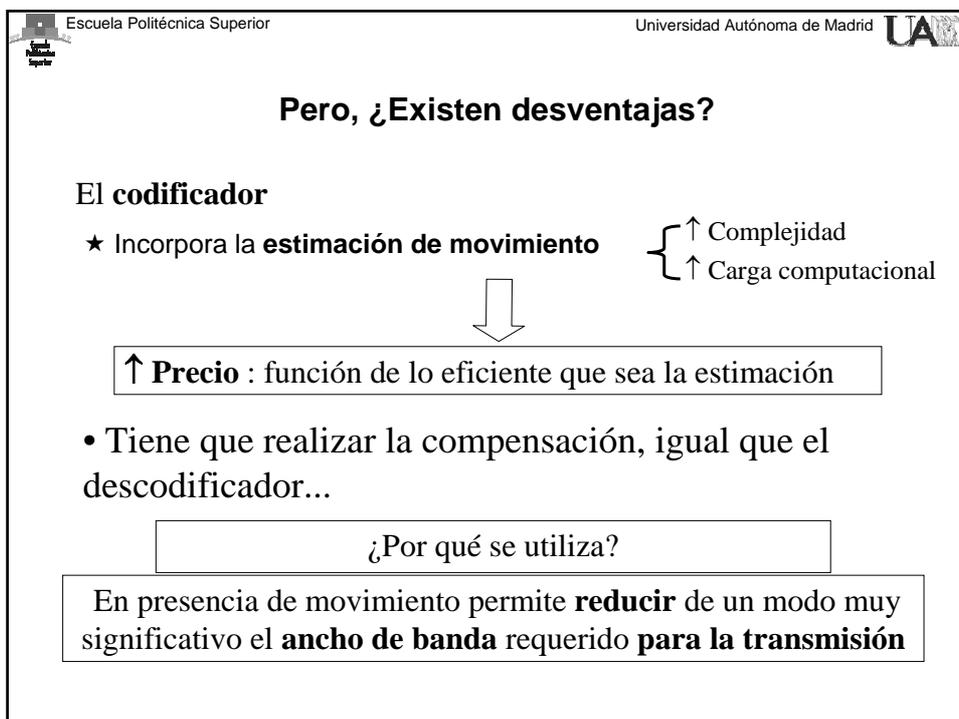
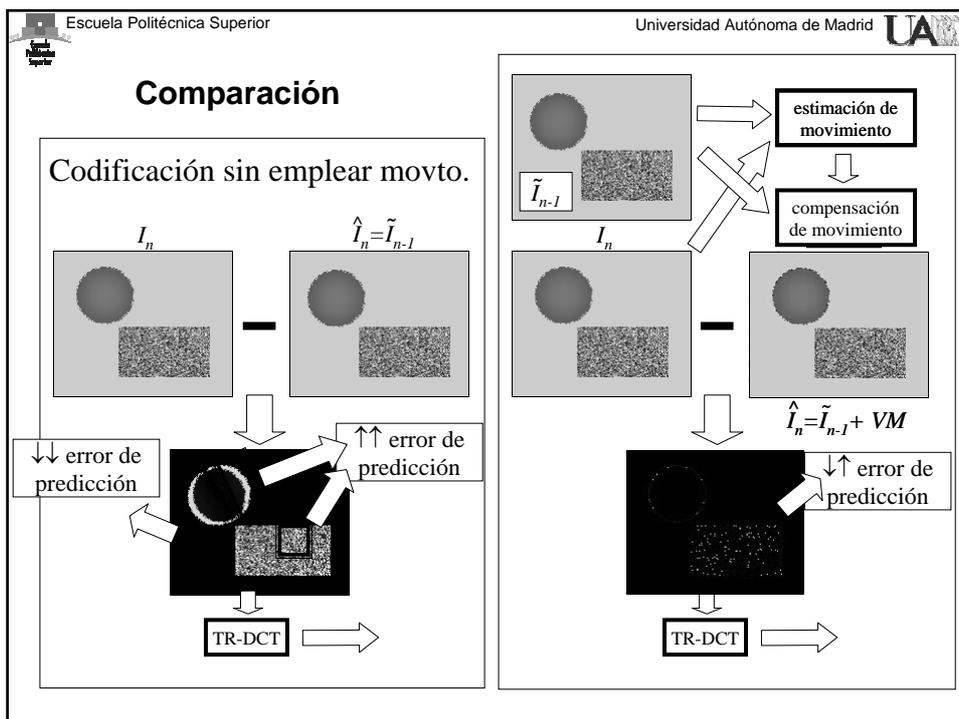
Estimación y compensación

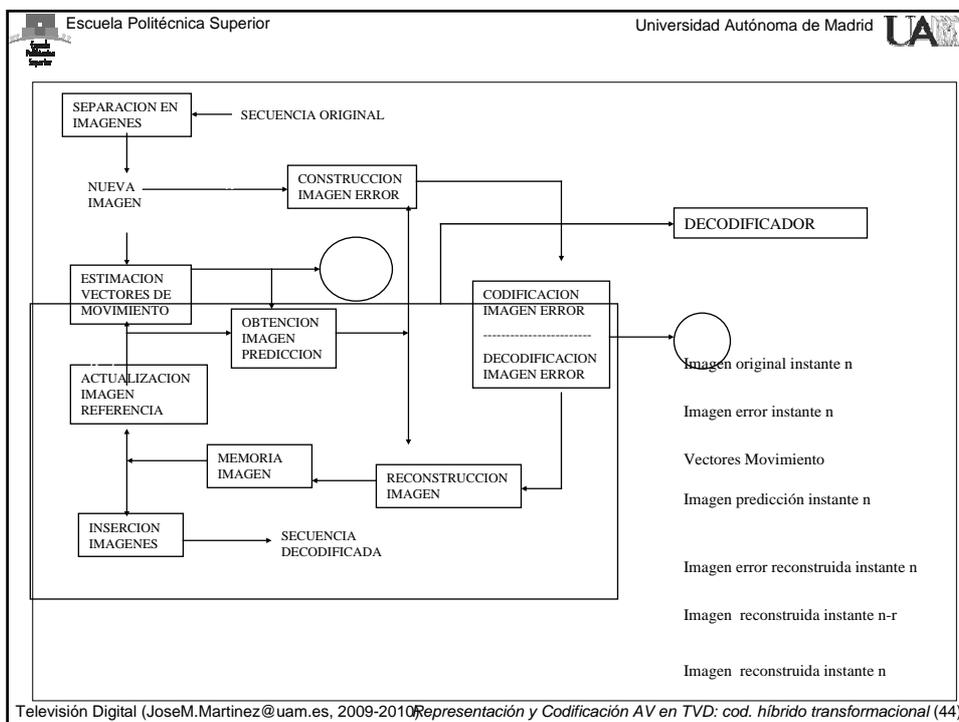
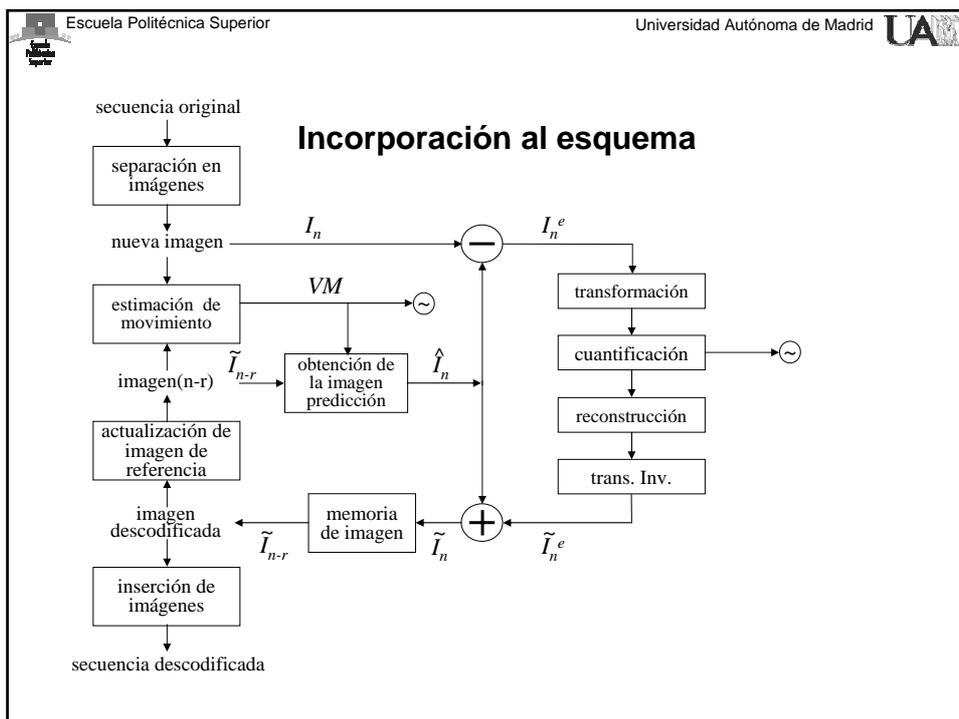
¿Qué habría que transmitir?
El error de predicción y la información necesaria para que el decodificador pueda generar la imagen predicción

- Información de movimiento.

Para que la imagen predicción sea idéntica en el codificador y en el decodificador
La imagen o imágenes de referencia deben ser imágenes reconstruidas.

The diagram illustrates the motion estimation and compensation process. It starts with the current frame I_n and the previous reconstructed frame \tilde{I}_{n-1} . The process involves motion estimation and motion compensation to produce a predicted frame $\hat{I}_n = \tilde{I}_{n-1} + VM$. The prediction error, indicated by a downward arrow, is then processed by TR-DCT.







Predicción temporal con estimación y compensación del movimiento

Procesamiento de los cambios entre imágenes de una secuencia a través de correspondencias.

Mejorar la predicción temporal mediante la reducción del efecto del movimiento entre imágenes de una secuencia.

Estimación/Predicción:

- Muy costosa
- En el codificador solamente
- Vectores de movimiento (a codificar y transmitir)

Compensación

- A partir de la compensación se compensan los cambios antes de la predicción de la imagen
- En codificador y decodificador
- Tras la compensación, se haya el error de predicción (a codificar y transmitir)



Estimación de movimiento en tratamiento digital de vídeo (*)

Es uno de los problemas fundamentales en TDV (Tratamiento Digital de Vídeo)

- El movimiento de objetos y de la cámara es fuente de variaciones temporales
- Si se conocen los parámetros de este movimiento 3D, se puede predecir (con más o menos acierto) el movimiento
 - o Compresión, segmentación, filtrado, tratamiento de imágenes
- El movimiento 3D relativo entre cámara y objetos induce un movimiento 2D en el plano de la imagen (sistema de proyección)
 - o Proyección 2D: movimiento aparente o flujo óptico
 - o Es lo que podemos estimar (recuperar) de la señal de vídeo (intensidad)
 - o Movimiento real y flujo óptico no tienen porque coincidir
 - Cambio de iluminación genera un flujo óptico
 - Esfera uniforme que rota sobre su eje genera un flujo óptico (en condiciones de iluminación homogénea)

Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- **Estrategias de predicción.**
 - *Introducción*
 - *Predicción temporal.*
 - *Introducción del movimiento en el cálculo de la predicción.*
 - *Estimación y compensación de movimiento.*
 - **Métodos de Estimación de movimiento**
 - *Métodos de Compensación de movimiento*
- *Sistemas de contención.*
- *Errores de transmisión y necesidad de refresco.*
- *Líneas de mejora*

Métodos de estimación (*)

Ajustes de bloques

- El método más popular en codificación de vídeo por su simplicidad y buenos resultados.

Recursivos de píxel

- Para cada píxel del bloque se parte de una estimación a priori que se corrige iterativamente. Muy costoso.

Correlación de fases

- Basada en que la translación en una imagen cambia la fase de la Transformada de Fourier pero no su amplitud.

Basados en modelos

- **paramétricos**
 - Mediante modelos afines se logra modelar movimientos más complejos que la translación. Mayor complejidad y necesidad de una adecuada cuantificación de los parámetros del modelo afín.
- **determinísticos**
- **estocásticos**

Flujo Óptico

- Muy usado en segmentación.

Estimación para codificación

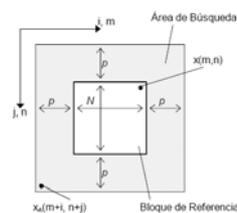
- el objetivo es reducir el error de predicción
- se trabaja por bloques
 - Más robusto frente al ruido
 - Menos vectores a transmitir
 - Consistente con la codificación por macrobloques DCT
- Se busca para la imagen a codificar el bloque que más se parece en la imagen de referencia (en el caso más simple, la anterior)
 - Se asume que la proyección de los objetos en el plano sufre un desplazamiento, lo que simplifica el método. En un caso general, los objetos sufren rotaciones, zooms, oclusiones, ...
 - Se trabaja con imágenes de referencia reconstruidas
- Lo usado en H.261, MPEG-1, MPEG-2, MPEG-4, H.264/AVC
 - Las diferencias están en los tamaños de bloque, áreas de búsqueda, ...
 - H.264 permite que los bloques sean de diferentes tamaños (modos) y relación de aspecto
- Contras: imposibilidad de manejar rotaciones, deformaciones, zooms, ...; efecto de bloques (no tienen generalmente correlación con los objetos ni el movimiento)

Ajuste de bloques (I)

Bloques de $M \times N$ (generalmente $N \times N$)

Área de búsqueda

- Imagen anterior
- (x_0, y_0) : esquina superior izquierda del área de búsqueda
- (x_m, x_n) : esquina superior izquierda bloque referencia (imagen actual)
- $x_0 + p = x_m$ / $y_0 + q = x_n$
- Área: $(M+2p) \times (N+2q)$



Búsqueda

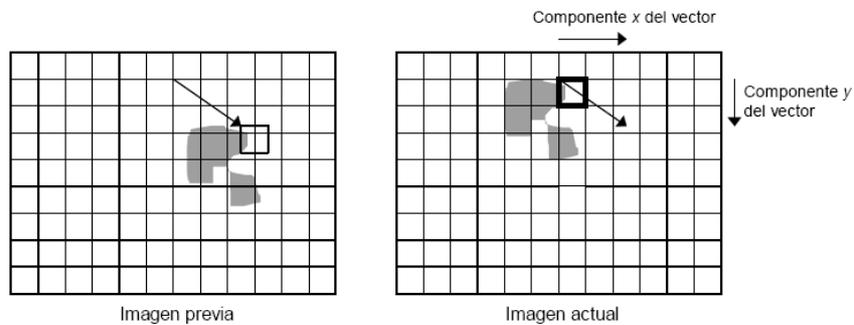
- $M \times N(x_0+i, y_0+j)$ vs $M \times N(x_m, x_n)$
- $i=0..2p$ / $j=0..2q$
 - $(2p+1)(2q+1)$ comprobaciones (exhaustivo)

Vector de movimiento

- $(u, v) = (x_0+i-x_m, x_0+j-x_n)$ para (i, j) mejor "matching"
 - $(x_0+i-x_0-p, y_0+j-x_0-q) = (i-p, j-q)$
 - $([-p, +p], [-q, +q])$

Ajuste de bloques (II)

Búsqueda del bloque (de igual tamaño) que más se parece



Demostración

Posibilidades del ajuste de bloques (I)

Tamaño de bloque

- Si pequeño: menor error de predicción, más vectores, menor fiabilidad (menos píxeles a "estimar")
- Si grande: más rápidos

Área de búsqueda

- Velocidad frente a posibles mejoras (métodos exhaustivo)
- Tamaño y "entropía" de vectores de búsqueda (hay que codificarlos para su transmisión)

Precisión de cálculo

- Píxel, $\frac{1}{2}$ píxel, $\frac{1}{4}$ píxel
- Requiere interpolación (completa o tras estimación entera –algoritmo de dos pasos–)
- Mayor fiabilidad a costa de mayor cómputo



Posibilidades del ajuste de bloques (II)

Criterio de selección/similitud (función de coste) escogido

- Minimizar función de coste:
 - o Error cuadrático medio (Mean Square Error –MSE-)

$$D(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N [x(m, n) - x_A(m + i, n + j)]^2$$

- o Error absoluto medio (Mean Absolute Error –MAE-)

$$D(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N |x(m, n) - x_A(m + i, n + j)|$$

- o son las más utilizadas por su bajo coste computacional
- o A nivel implementación a veces se usan sin dividir (no medio) para ahorrar cálculos (cuidado con rango de variables)

Algoritmos de ajuste de bloque escogidos

- Búsqueda exhaustiva vs. Búsqueda rápida



Algoritmos de ajuste de bloques

Búsqueda exhaustiva (*full search*)

- El más preciso ;)
- Su coste es proporcional al tamaño del área de búsqueda
 - o $(2p+1)(2q+1)$ evaluaciones de la función de error
 - o sobre MN píxeles
 - o Cada píxel: una diferencia, un cuadrado (MSE) o un valor absoluto (MAE), una suma (más el tiempo de acceso a píxel)
 - o Asumiendo un vídeo de T cuadros/seg, FxC píxeles
 - o Operaciones/seg= $\{(FCT)/MN\} 3 (2p+1)(2q+1)$ MN = FCT $3 (2p+1)(2q+1)$
 - o T=25, C=720, F= 576, p=q=15 => 29,98 G OPS (p=q=7 => 6,99 G OPS)

Búsqueda rápida

- Reducen al menos un orden de magnitud, pero pueden dar un óptimo local en lugar de uno global
 - o Búsqueda logarítmica
 - o Búsqueda piramidal (método multiresolución)
 - o Combinaciones

Algoritmos de ajuste de bloques

Búsqueda logarítmica (I)

Búsqueda logarítmica

- Reducción de los orígenes sobre los que se busca
 - Direction of Minimum Distorsion (DMD)
"Displacement Measurement and Its Application in Interframe Image Coding" J.R. Jain, A.K. Jain. IEEE Trans. on Communications, 29(12), Dec 1981.
 - Three Step Search (TSS)
"Motion compensated interframe coding for video conferencing". T. Koga, K. Iinuma, A. Hirano, Y. Iijima, T. Ishiguro. Proc. Nat. Telecommun. Conf. New Orleans, L.A. Nov 1981.
 - One-at-a-Time Search (OTS)
Simplificación de Conjugate Direction Search (CDS)
"Predictive coding based on efficient motion estimation". R. Srinivasan, K.R. Rao. IEEE Trans. on Communications, 33(8), Aug 1985.
- Por supuesto, existen otros muchos métodos de búsqueda rápida: hexagonal, en diamante... pero en todos casos el principio es el mismo: reducir los pasos del algoritmo exhaustivo ordenando la búsqueda de otra manera.

Algoritmos de ajuste de bloques

Búsqueda logarítmica (II) (DMD)

El algoritmo DMD (Direction of Minimum Distortion)

Función de coste

- MSE

La búsqueda comienza en cinco posiciones

- la central, una a cada lado de ella en dirección vertical y otras dos en dirección horizontal
- con un paso de tamaño igual a $p/2$ o a la potencia de 2 más próxima a $p/2$.

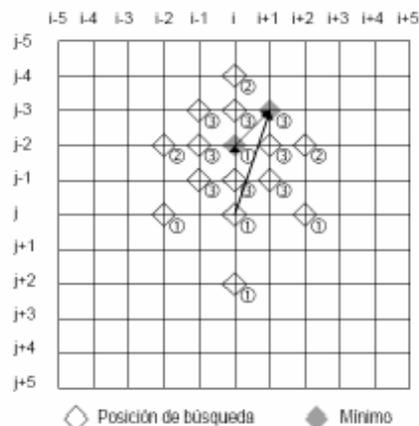
En la segunda iteración se calculan tres posiciones más con el mismo paso y tomando como posición central la del mínimo obtenido en la búsqueda anterior.

Esto se repite hasta que los mínimos de dos etapas coinciden.

En ese momento se divide el tamaño del paso por 2 y se repite el proceso hasta que el paso es igual a 1

En este momento se busca en los ocho puntos que rodean al central.

El mínimo de esta última búsqueda se toma como posición óptima.

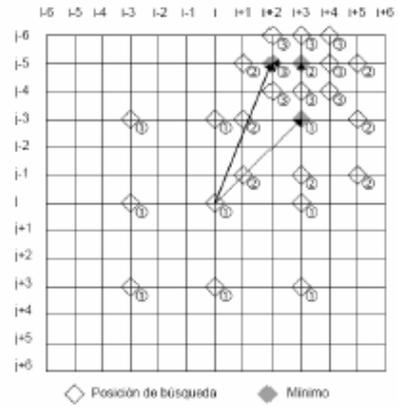


◇ Posición de búsqueda ◆ Mínimo

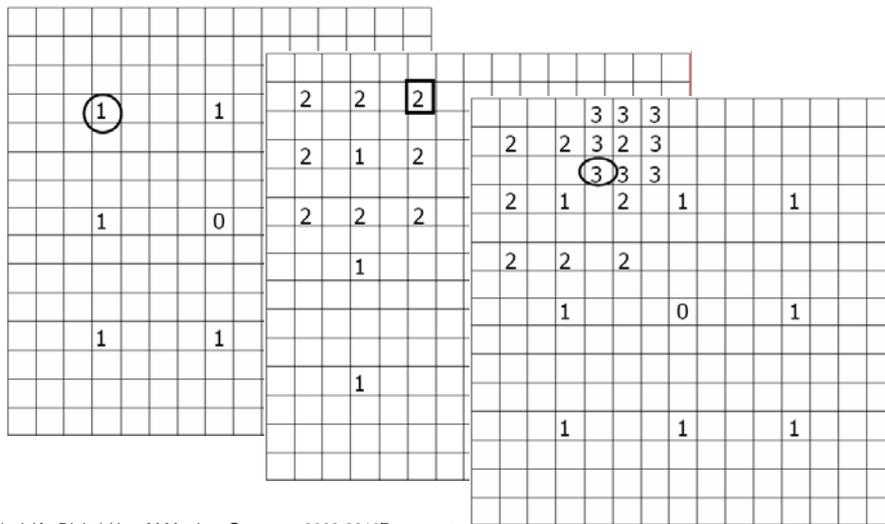
Algoritmos de ajuste de bloques Búsqueda logarítmica (III) (TSS)

El algoritmo TSS (Three Step Search) es similar al anterior pero con dos diferencias:

- la función de coste empleada es MAE
- el patrón de búsqueda varía tal como se ilustra en la figura.
- Se calcula $d_{1p/q} = 2^{kp/q} - 1$
- $kp/q = \text{ceil}(\log_2 p/q)$
- Se evalúa en los bloques con origen en
 - $(0,0)$, (d_{1p}, d_{1q}) , $(-d_{1p}, d_{1q})$, $(-d_{1p}, -d_{1q})$, $(d_{1p}, -d_{1p/q})$, $(d_{1p}, 0)$, $(-d_{1p}, 0)$, $(0, d_{1q})$, $(0, -d_{1q})$
- El mínimo es el nuevo origen y se dividen las distancias entre 2
- Se repite hasta que los orígenes estén a 1 píxel (kp/q indica el número de iteraciones)
- TSS ($p=q=7 \Rightarrow k=3$)



Algoritmos de ajuste de bloques Búsqueda logarítmica (IV) (TSS)



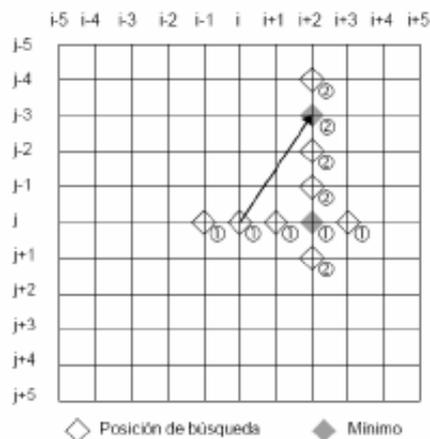
Algoritmos de ajuste de bloques Búsqueda logarítmica (V) (TSS)

- Siendo $p=q$
 - Número de comparaciones por bloque $(8k+1)$
vs $(2p+1)(2q+1)$
 - Píxeles del bloque MN
 - $(8k+1)MN$ evaluaciones de la función de coste
 - Cada píxel: una diferencia, un cuadrado (MSE) o un valor absoluto (MAE), una suma (más el tiempo de acceso a píxel)
 - Asumiendo un vídeo de T cuadros/seg, $F \times C$ píxeles
 - Operaciones/seg = $\{(FCT)/MN\} 3 (8k+1)MN = FCT 3 (8k+1)$
 - $T=25, C=720, F= 576, p=q=15 \Rightarrow 1,03 \text{ G OPS}$ ($p=q=7 \Rightarrow 0,77 \text{ G OPS}$)
vs $29,98 \text{ G OPS}$ ($p=q=7 \Rightarrow 6,99 \text{ G OPS}$)

Algoritmos de ajuste de bloques Búsqueda logarítmica (VI) (OTS)

El algoritmo One-at-a-Time Search (OTS) –simplificación de CDS-.

Se comienza la búsqueda en la dirección horizontal (eje i en la figura) con paso de un píxel comenzando por la posición central y las dos adyacentes y continuando en el sentido de la que produzca menor error hasta que se encuentre un mínimo. A partir de dicho mínimo, se continúa la búsqueda en dirección vertical (eje j) hasta encontrar un nuevo mínimo que se toma como predicción.



Algoritmos de ajuste de bloques

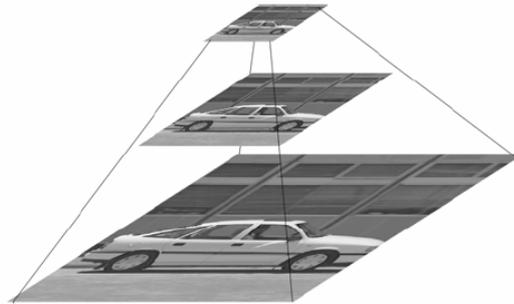
Búsqueda Piramidal (método multiresolución) (I)

Exhaustiva sobre imagen reducida con bloque reducido

Resolution level 1

Resolution level 3

Resolution level 3



Algoritmos de ajuste de bloques

Búsqueda Piramidal (método multiresolución) (II)

- El primer paso es construir para cada imagen una representación piramidal multiresolución. Esta representación consiste en submuestrear la imagen a la mitad para cada nivel. Previamente al submuestreo es importante filtrar la imagen paso bajo (teoría submuestreo).
- Se realiza una búsqueda exhaustiva de los vectores de movimiento en la imagen de menor resolución (donde además el tamaño de bloque puede ser menor, pues el filtrado elimina los detalles) y a continuación usar esos vectores de movimiento para refinar la búsqueda en la imagen de mayor resolución siguiente (y así hasta encontrar los vectores definitivos).
- La reducción se logra al usar uno (o pocos) de los mínimos de cada nivel para refinar exhaustivamente el nivel anterior.
- Además, en ciertos codificadores de vídeo donde se dispone de antemano de versiones de menor resolución del *frame* de vídeo (codificadores con escalabilidad), se puede utilizar este método sin necesidad de construir primero la pirámide multiresolución, con el ahorro en tiempo de computación que esto supone.

Algoritmos de ajuste de bloques

Búsqueda Piramidal (método multiresolución) (III) (*)

- Distintos parámetros se pueden variar en este método, como por ejemplo el número de niveles de la pirámide, el tamaño del *kernel* del filtro usado o el rango de búsqueda y el tamaño de bloque en cada uno de los niveles. Un ejemplo típico de valores de estos parámetros se encuentra en la siguiente tabla:

Parámetro vs. Nivel (en píxeles)	5	4	3	2	1
Tamaño del kernel LPF	10x10	10x10	5x5	5x5	3x3
Máximo desplazamiento	±31	±31	±7	±3	±1
Tamaño de bloque	128x128	64x64	64x64	28x28	12x12

- Mención especial requiere el filtro que se utilice para el diezmo de la imagen. Éste puede ser tan sencillo como la media de píxeles vecinos (filtro cuadrado) o tan complejo como un filtrado Gaussiano o Laplaciano.

Algoritmos de ajuste de bloques

Búsqueda Piramidal (método multiresolución) (IV)

A cada nivel (se suele submuestrear por 2 cada dimensión)

- $F_L = F/2^L$, $C_L = C/2^L$
- $(x_L, y_L) = (x/2^L, y/2^L)$
- En nivel L se calcula el vector de movimiento mediante uno de los algoritmos ya vistos (en principio con uno exhaustivo)
 - o niveles = L (L, L-1, ..., 1, 0). Nivel 0 == nivel imagen original
 - o $M_L = M/2^L$, $N_L = N/2^L$, $p_L = p/2^L$, $q_L = q/2^L \Rightarrow u_L = u/2^L$, $v_L = v/2^L$
- Para el nivel L-1, se selecciona el área de búsqueda en el entorno de $(x_{L-1}, y_{L-1}) = (2x^L + 2u^L, 2y^L + 2v^L)$
- Se aplica búsqueda exhaustiva ($p=q=1$), de forma que se refina el vector anterior en +1,0,-1 (\Rightarrow 9 bloques)
- Y se repite hasta L=0

Requiere más memoria, coste adicional de filtrado, y pérdida de detalles que reducen la fiabilidad y que no se puede corregir debido a que tras el nivel L, solamente existe libertad de movimiento en 9 bloques.

Algoritmos de ajuste de bloques

Búsqueda Piramidal (método multiresolución) (V)

Coste (nivel L búsqueda exhaustiva)

- Nivel L
 - Número de comparaciones nivel L : $(2p_L+1)(2q_L+1)$
 - Píxeles del bloque nivel L : $M_L N_L$
 - Píxeles de imagen nivel L : $F_L C_L$
 - Cada píxel: una diferencia, un cuadrado (MSE) o un valor absoluto (MAE), una suma (más el tiempo de acceso a píxel)
 - Asumiendo un vídeo de T cuadros/seg, FxC píxeles
 - Nivel L: Operaciones/seg= $F_L C_L T^3 (2p_L+1)(2q_L+1)$
- Resto niveles (L-j, j = 1..L)
 - Número de bloques nivel L-j : 9
 - Resto igual
 - Nivel L-j: Operaciones/seg= $27 F_{L-j} C_{L-j} T$
- T=25, C=720, F= 576, p=q=15, L=2 (p/2=q/2=8 por exceso) => 507,38 M OPS (p=q=7 => 396,09 M OPS)
 - vs TSS: 1,03 G OPS (p=q=7 => 0,77 G OPS)
 - vs exhaustiva: 29,98 G OPS (p=q=7 => 6,99 G OPS)

Algoritmos de ajuste de bloques

Combinación de algoritmos de búsqueda rápida (I) (*)

Por supuesto, resulta muy fácil emplear alguno de los métodos de búsqueda rápida que hemos visto en el apartado anterior para encontrar los vectores de desplazamiento de alguno (o todos) los niveles de resolución. Tendríamos entonces un método de búsqueda híbrido, dando lugar a tantas posibilidades como combinaciones posibles existen entre métodos.

Representación y Codificación de Señales Audiovisuales en Televisión Digital

Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- **Estrategias de predicción.**
 - *Introducción*
 - *Predicción temporal.*
 - *Introducción del movimiento en el cálculo de la predicción.*
 - *Estimación y compensación de movimiento.*
 - *Métodos de Estimación de movimiento*
 - **Métodos de Compensación de movimiento**
- **Sistemas de contención.**
- **Errores de transmisión y necesidad de refresco.**
- **Líneas de mejora**

Compensación de movimiento

Tras la obtención de los vectores de movimiento se “genera” una imagen de predicción con los bloques de la de referencia desplazados

- Compensación de bloques independientemente
- Efectos de bloque

Métodos avanzados de compensación (*)

- Modelado de movimientos diferentes al puramente translacional
- Movimiento continuo en la imagen
- Más costosos
 - Compensación por solapamiento de bloques: peso en función de la distancia
 - Compensación por interpolación de la rejilla de control: basada en *wrapping* (solamente se dan vectores de movimiento de los puntos de control y el resto se interpola).



Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- *Estrategias de predicción.*
 - *Introducción*
 - *Predicción temporal.*
 - *Introducción del movimiento en el cálculo de la predicción.*
 - *Estimación y compensación de movimiento.*
 - *Métodos de Estimación de movimiento*
 - *Métodos de Compensación de movimiento*
- **Sistemas de contención.**
- Errores de transmisión y necesidad de refresco.
- Líneas de mejora



Sistemas de contención

Para cada imagen de la secuencia se transmite:

Canales de velocidad de transmisión fija



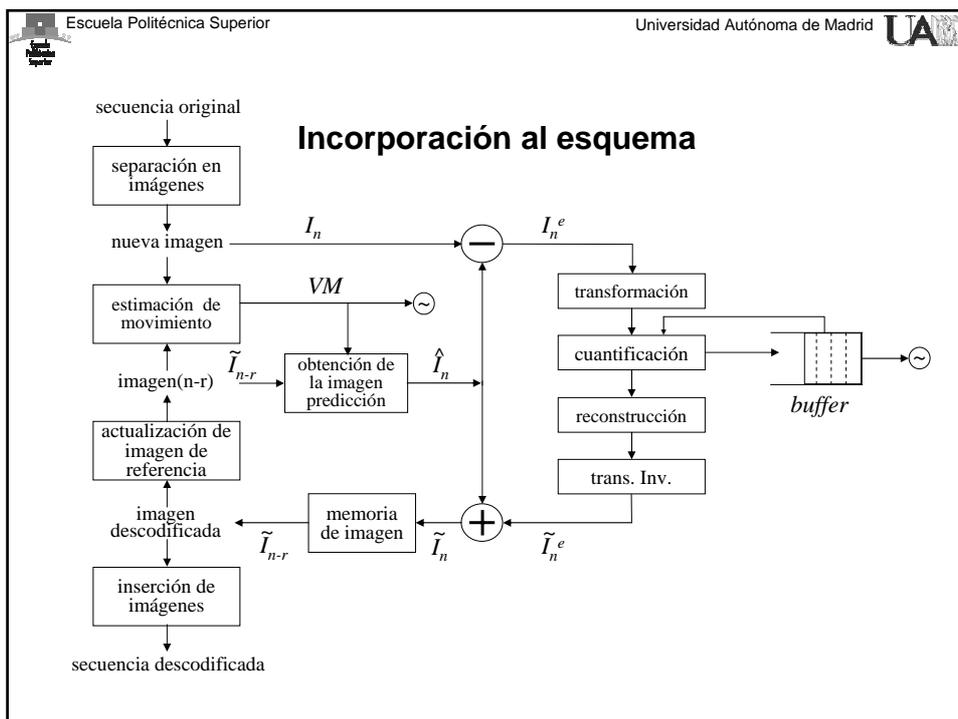
~~velocidad binaria fija~~



Necesidad de sistemas de contención (*buffers*: memorias FIFO)
que adapten la velocidad binaria del codificador al canal de transmisión

Y aseguren que no se pierde información: no se desborda el *buffer*

¿Cómo se puede adaptar la velocidad binaria?
Controlando la **etapa de cuantificación** ⇒ alterar la calidad de las imágenes
reconstruidas



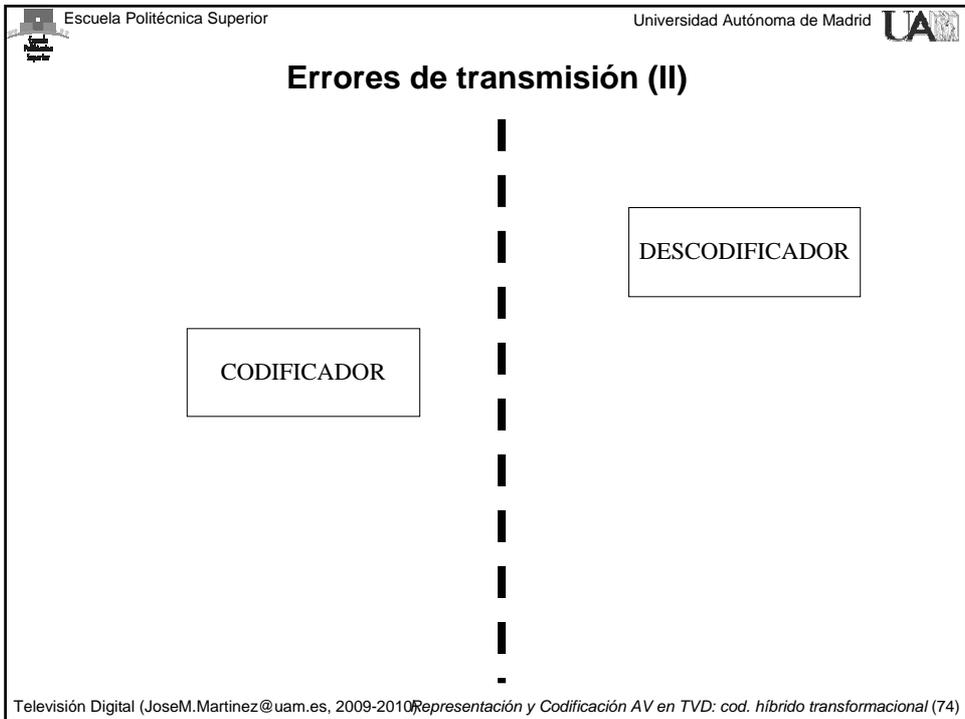
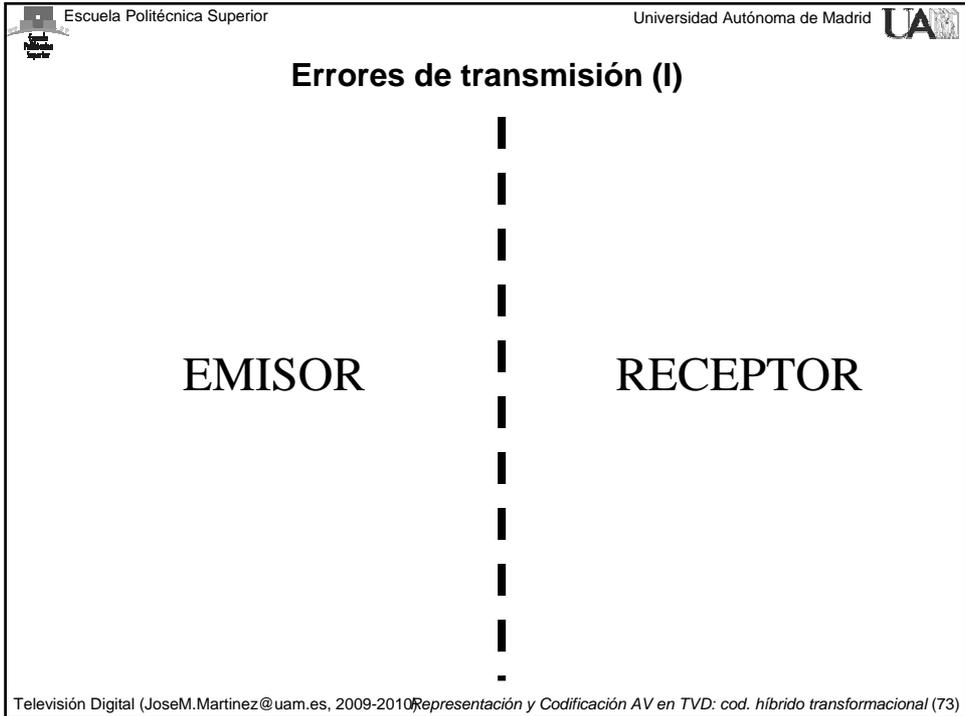
Escuela Politécnica Superior Universidad Autónoma de Madrid

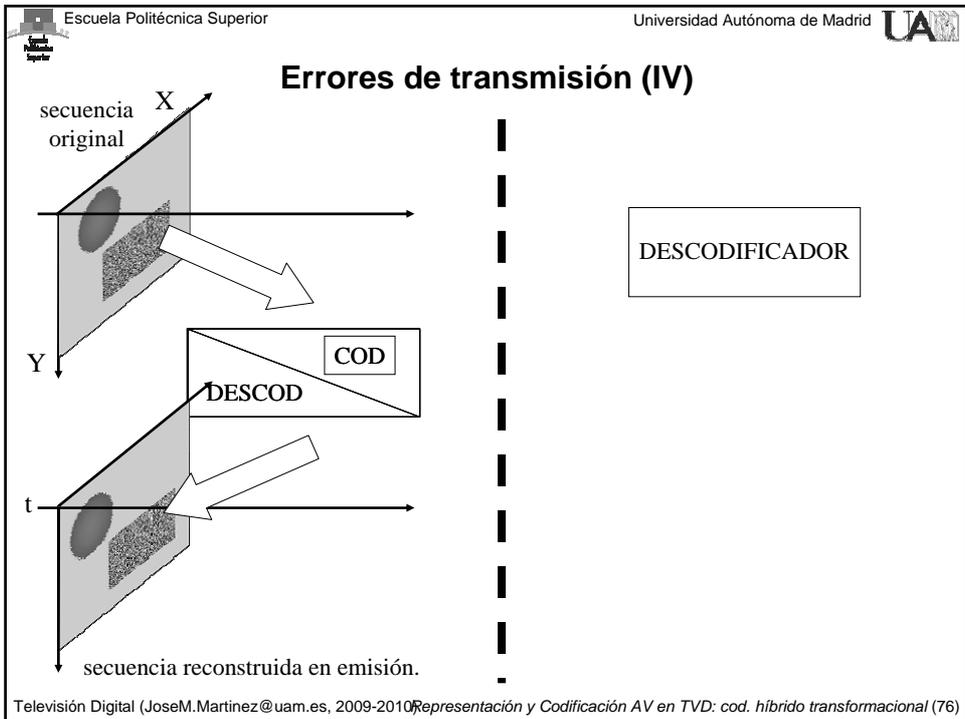
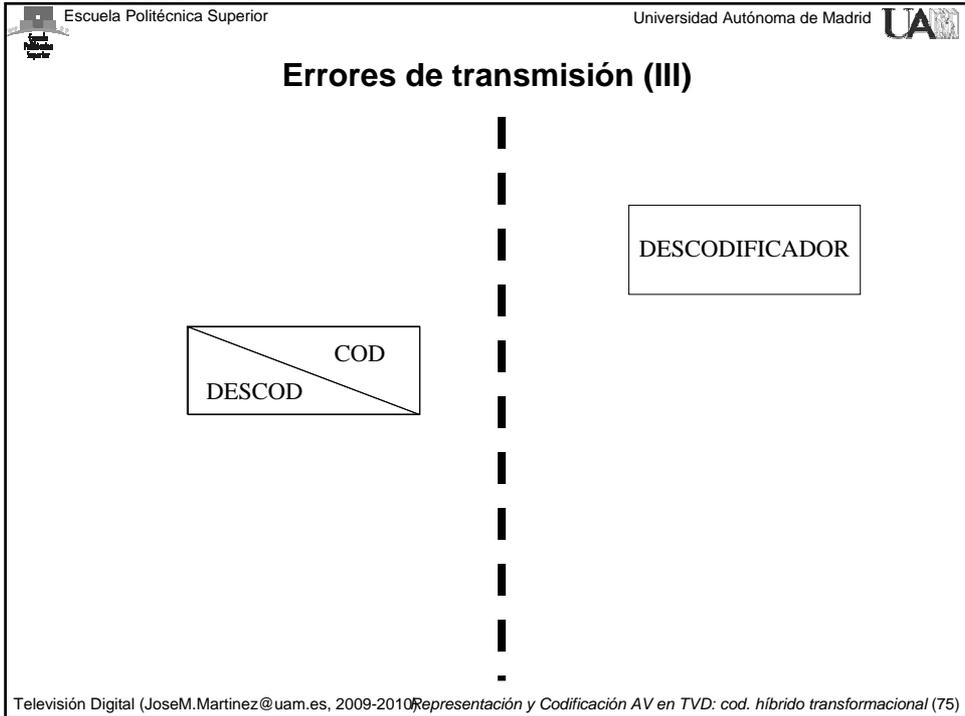
Representación y Codificación de Señales Audiovisuales en Televisión Digital

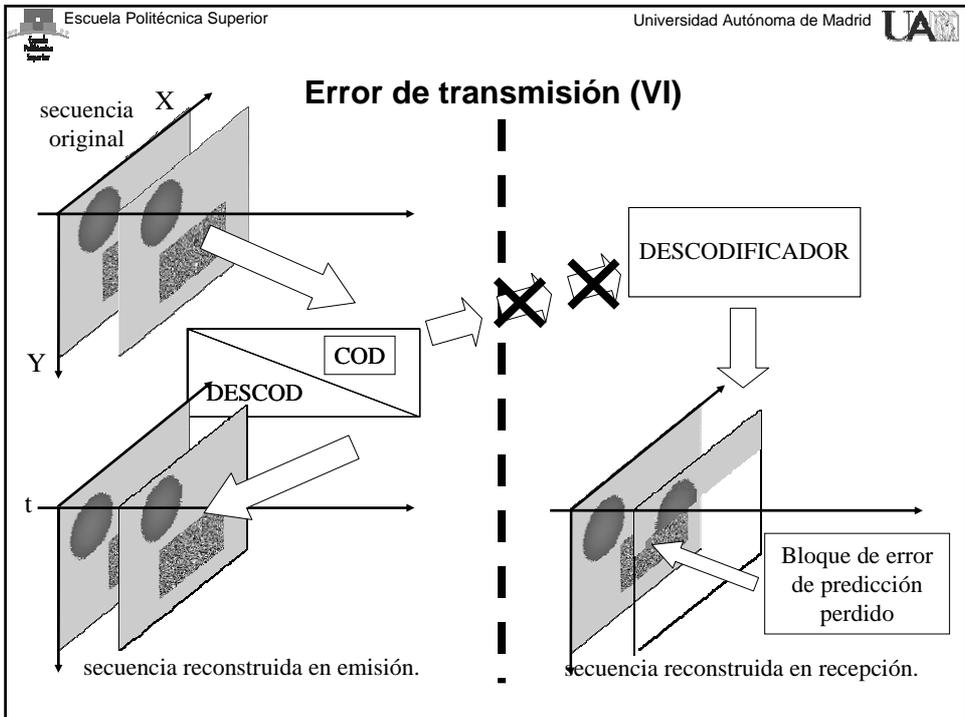
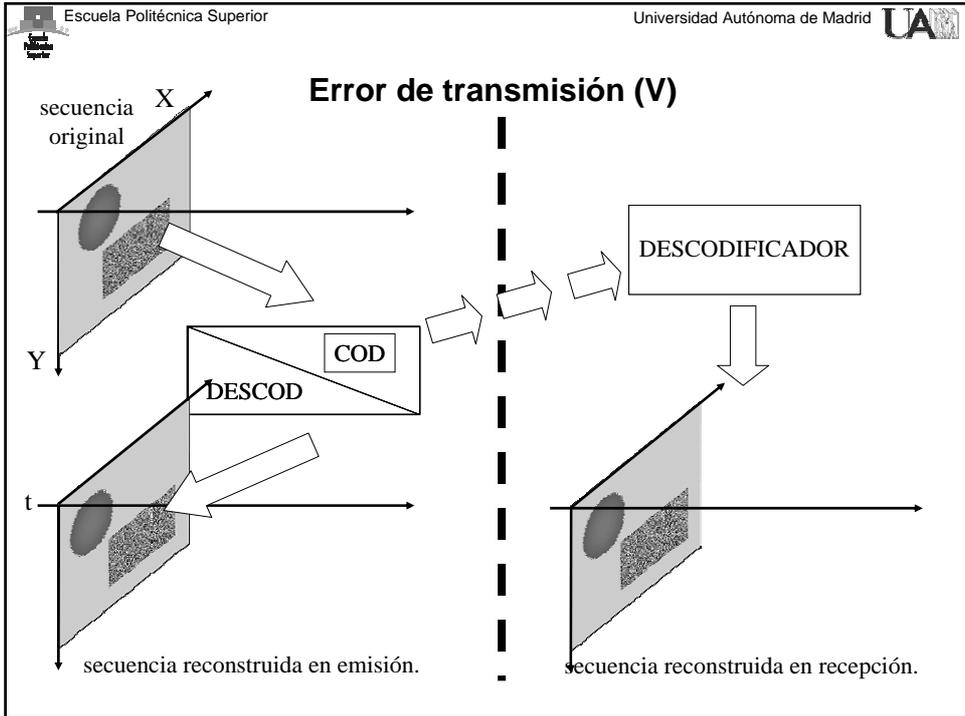
Compresión de vídeo híbrido transformacional

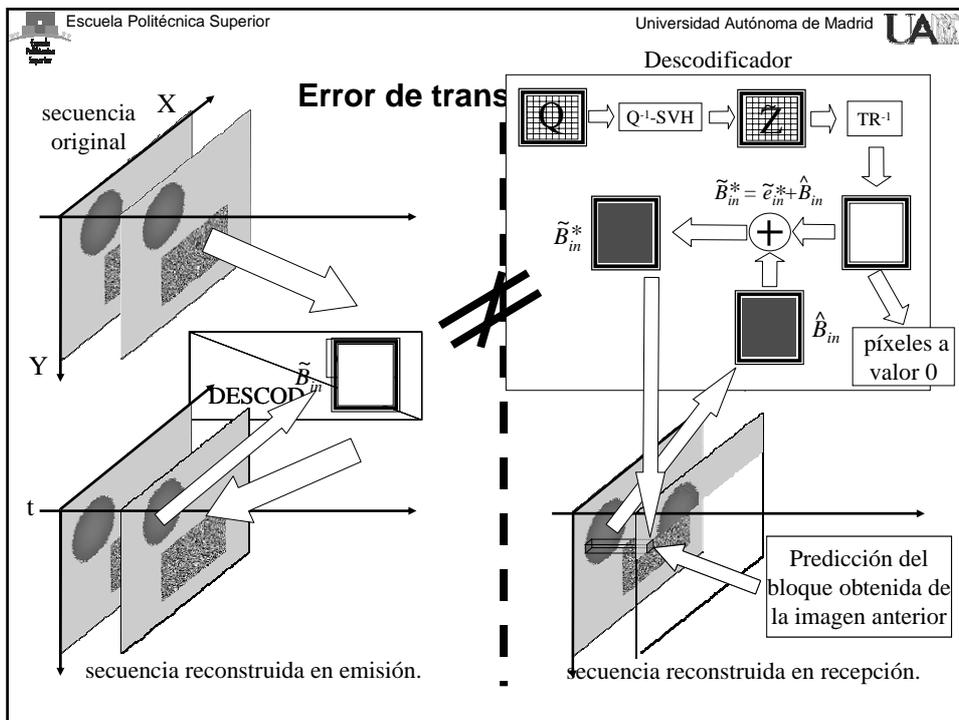
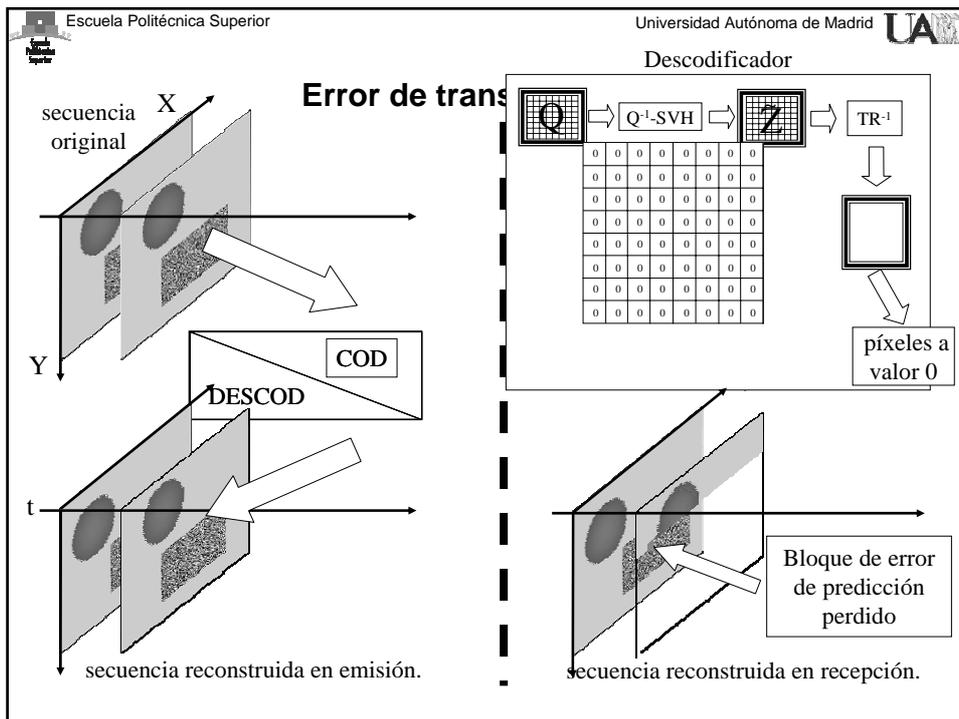
- **Introducción**
- **Descripción del sistema.**
 - Codificación de una imagen de la secuencia.
 - Detalle del codificador y del descodificador.
 - Obtención de la imagen predicción.
 - Diagrama de bloques.
- **Estrategias de predicción.**
 - Introducción
 - Predicción temporal.
 - Introducción del movimiento en el cálculo de la predicción.
 - Estimación y compensación de movimiento.
 - Métodos de Estimación de movimiento
 - Métodos de Compensación de movimiento
- **Sistemas de contención.**
- **Errores de transmisión y necesidad de refresco.**
- **Líneas de mejora**

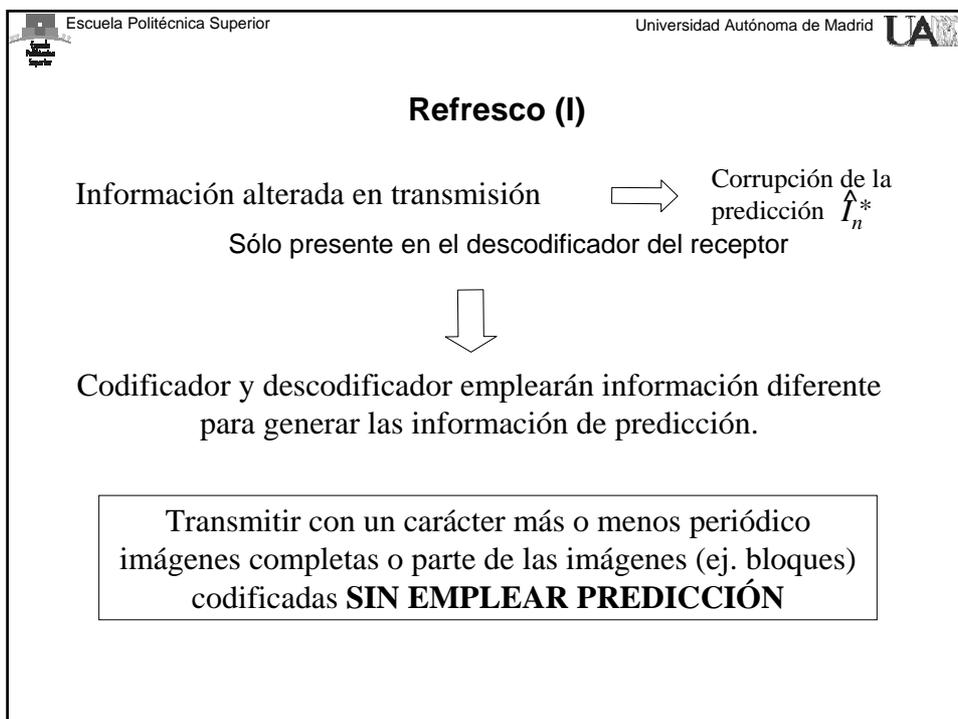
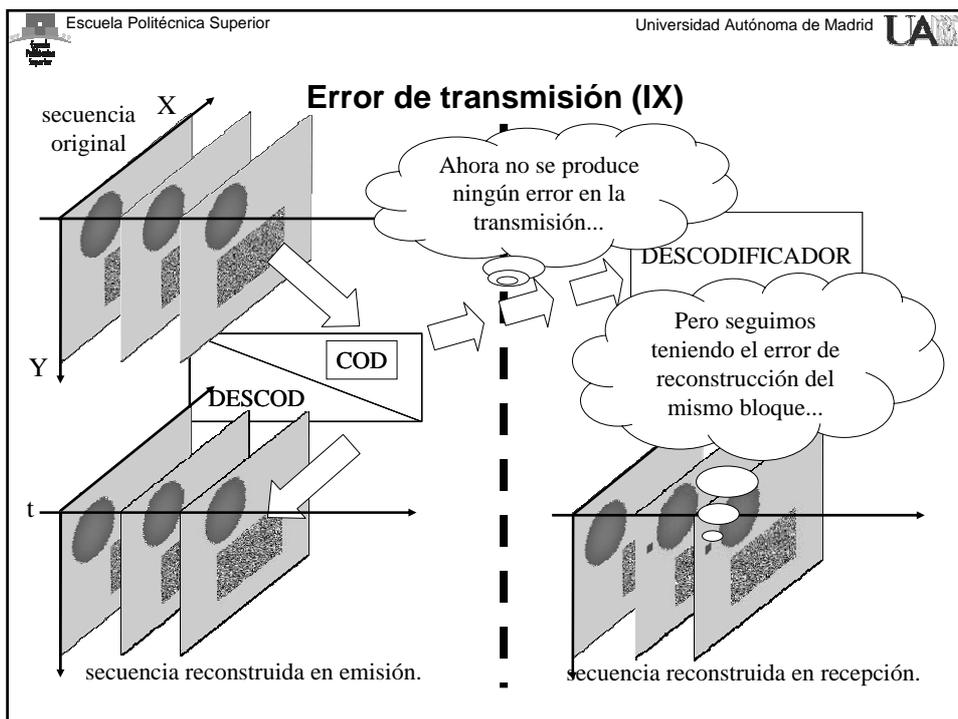
Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: cod. híbrido transformacional (72)

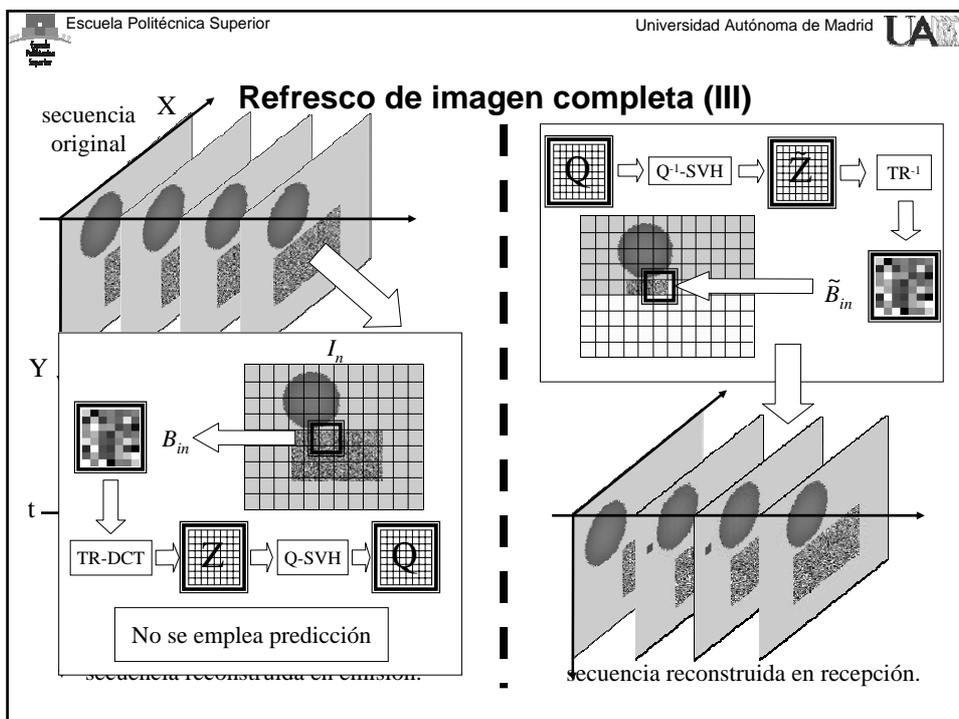
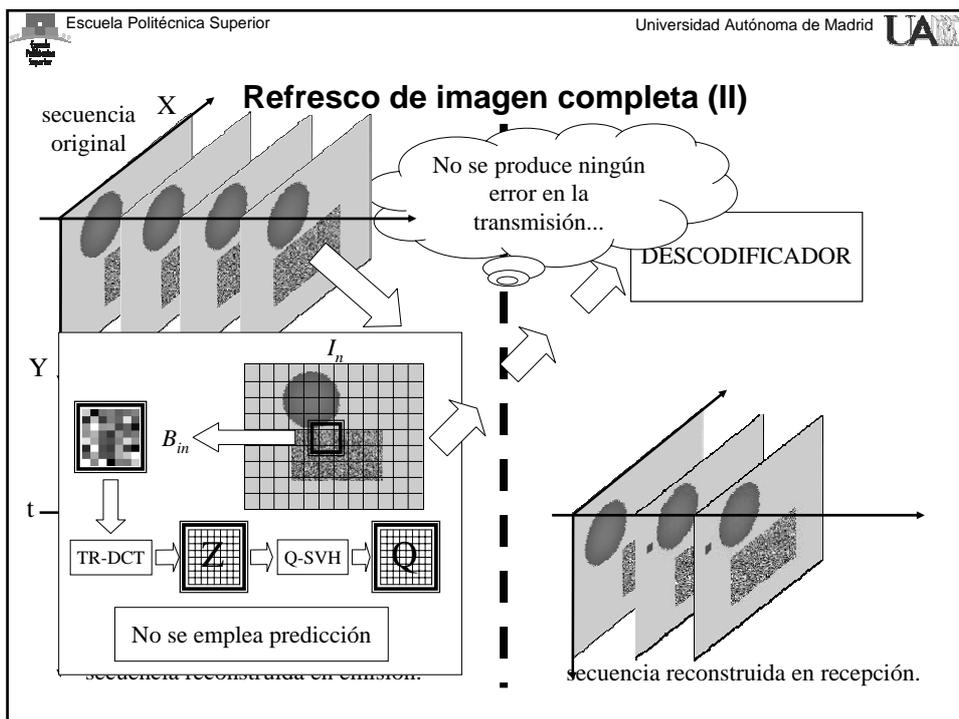


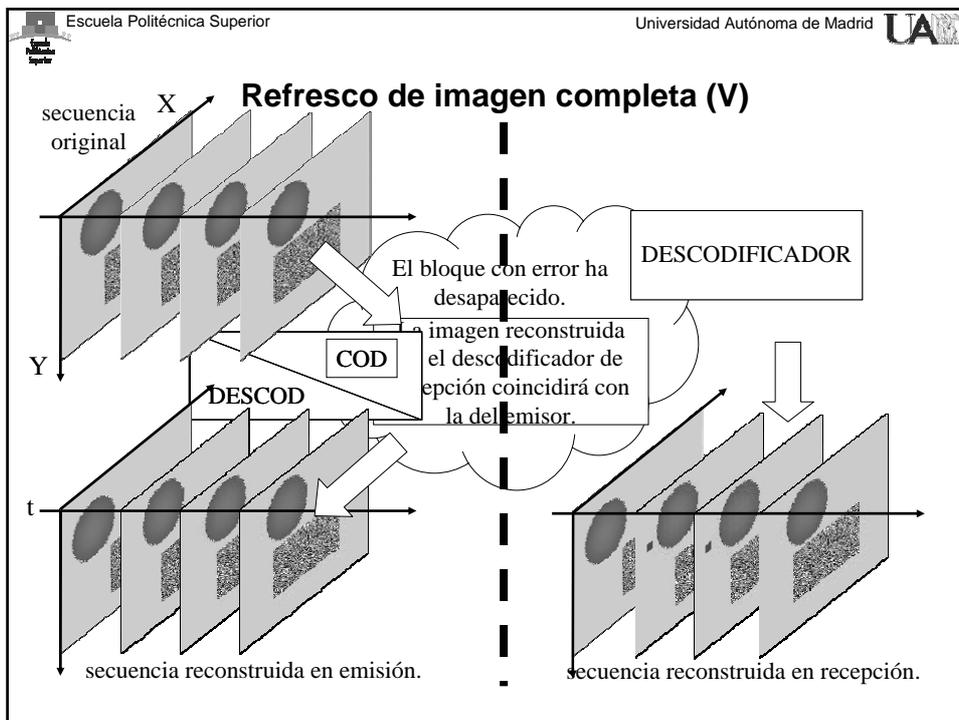
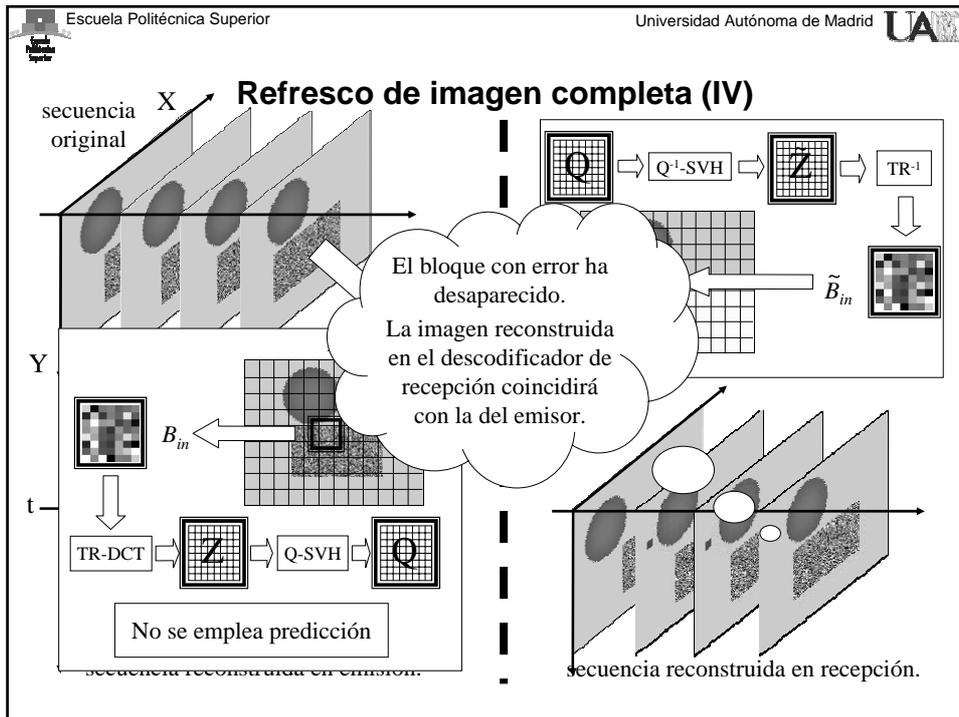












Representación y Codificación de Señales Audiovisuales en Televisión Digital Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
 - *Codificación de una imagen de la secuencia.*
 - *Detalle del codificador y del decodificador.*
 - *Obtención de la imagen predicción.*
 - *Diagrama de bloques.*
- *Estrategias de predicción.*
 - *Introducción*
 - *Predicción temporal.*
 - *Introducción del movimiento en el cálculo de la predicción.*
 - *Estimación y compensación de movimiento.*
 - *Métodos de Estimación de movimiento*
 - *Métodos de Compensación de movimiento*
- *Sistemas de contención.*
- *Errores de transmisión y necesidad de refresco.*
- **Líneas de mejora**

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: cod. híbrido transformacional (87)

Líneas de mejora

H.263: similar a MPEG-2 (alguna mejora de compresión)

- H.263+, H.263++, H.26L

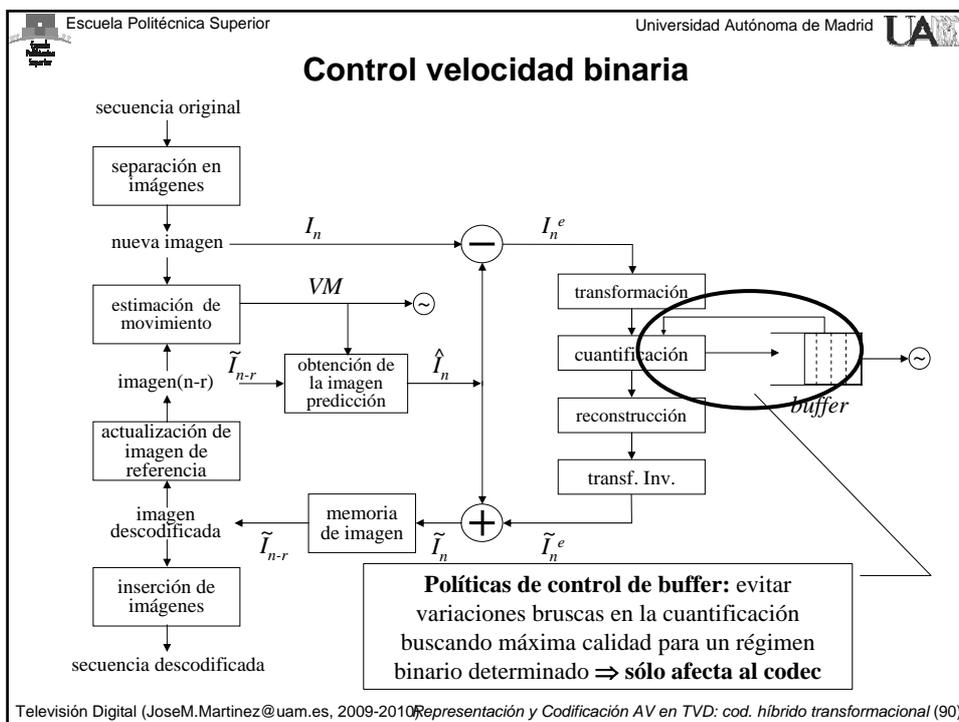
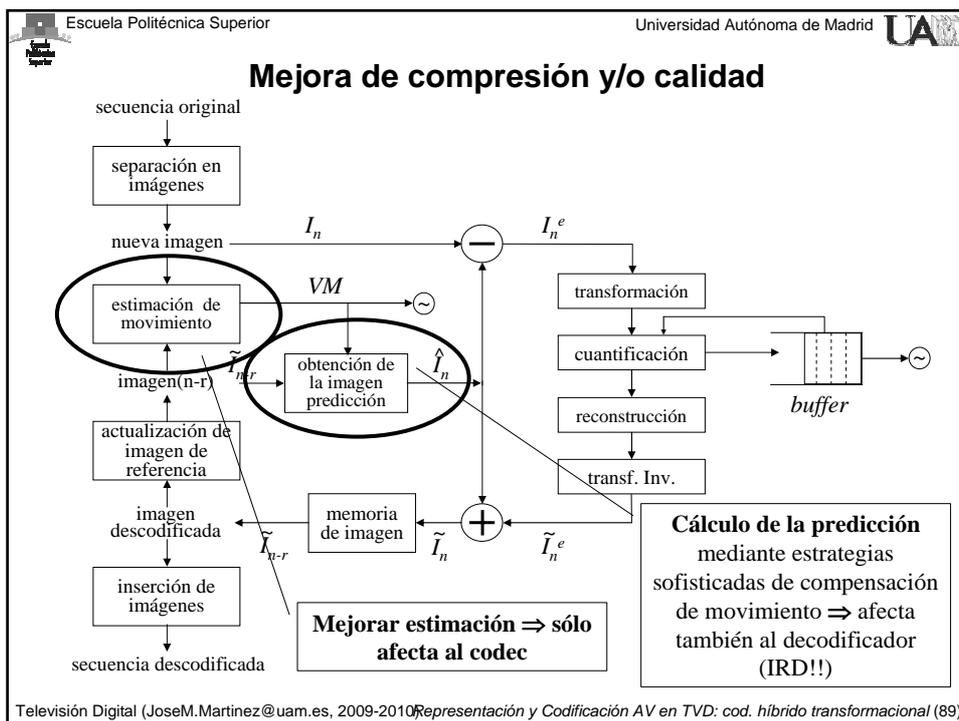
MPEG-4: H.263- pero introduce la codificación de objetos

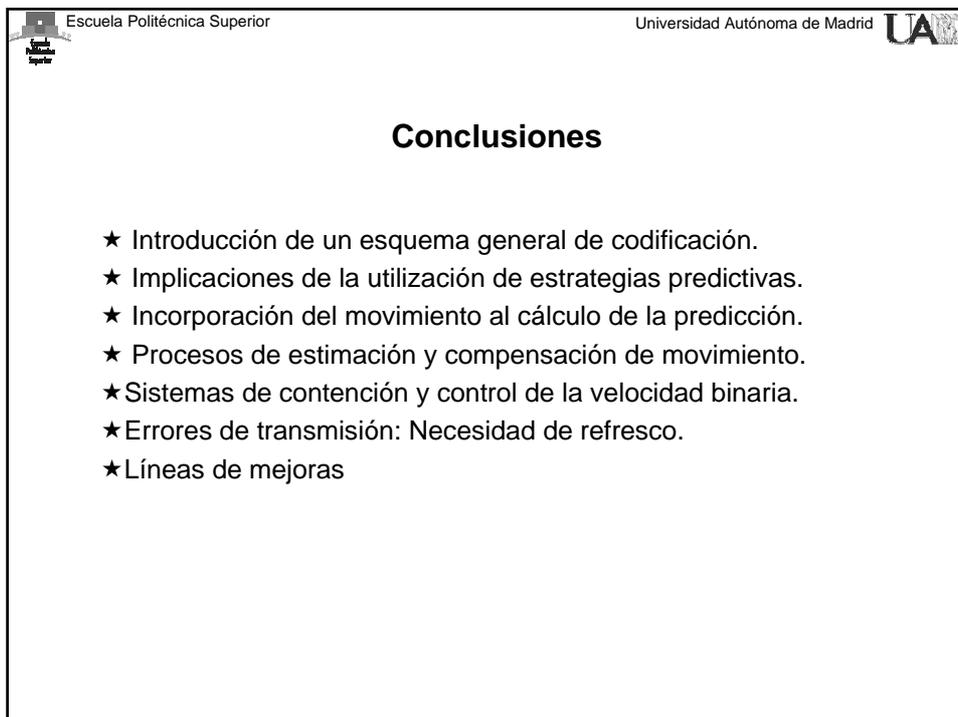
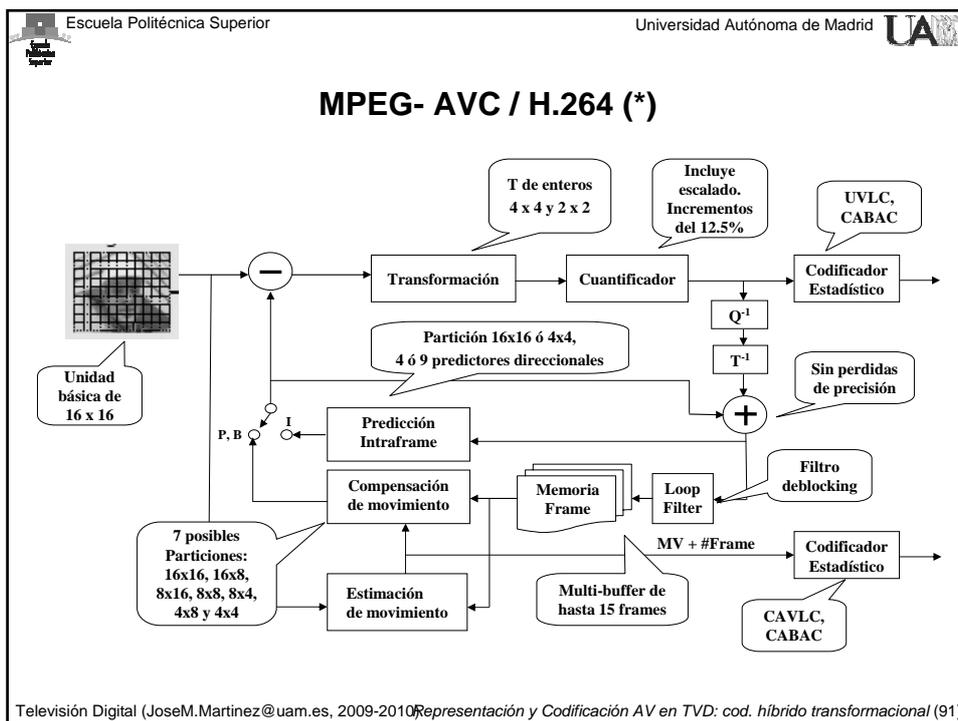
- Necesidad de sincronización y composición en terminal

Avances en nuevos estándares de codificación de vídeo Windows Media 9 y 9+, H.264:

- Consiguen grados de compresión y calidad muy superiores a los ofrecidos por MPEG-2:
 - 40%-50% reducción de velocidad binaria para imágenes de TVD.
- Requieren un aumento muy significativo del procesamiento en el codificador (para un 50% => 80%):
 - Debido principalmente a la complejidad de las tareas de estimación de movimiento.
- Pero también en el decodificador (IRD):
 - La compensación del movimiento es muy sofisticada.
 - Se flexibiliza la dimensión del macrobloque y del bloque para la determinación de la DCT.

Televisión Digital (JoseM.Martinez@uam.es, 2009-2010) Representación y Codificación AV en TVD: cod. híbrido transformacional (88)







Créditos

Estas transparencias están basadas en las elaboradas originalmente por Dr. Luis Salgado Álvarez de Sotomayor, Profesor Titular de la E.T.S.Ing.Telecomunicación de la UPM. Partes de la descripción de ajuste de bloques está basada en un trabajo de Óscar Mateo.