

## **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

2008-2009

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

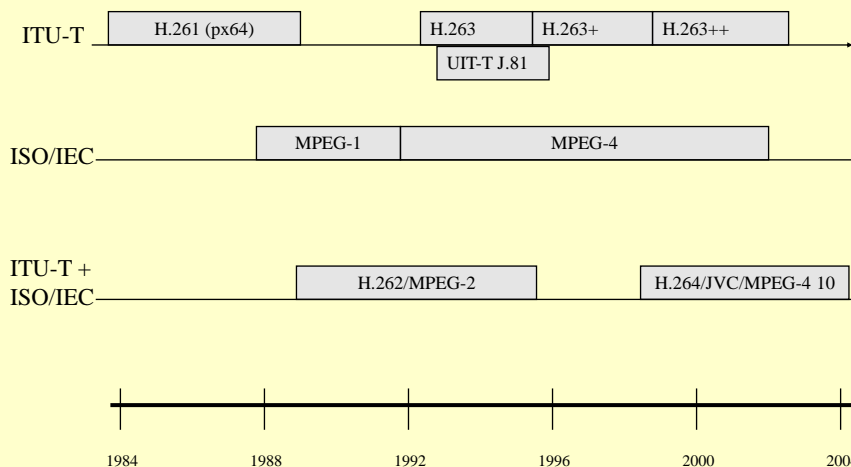
- Introducción
- Descripción del sistema.
  - Codificación de una imagen de la secuencia.
  - Detalle del codificador y del decodificador.
  - Obtención de la imagen predicción.
  - Diagrama de bloques.
- Estrategias de predicción.
  - Predicción temporal.
  - Introducción del movimiento en el cálculo 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



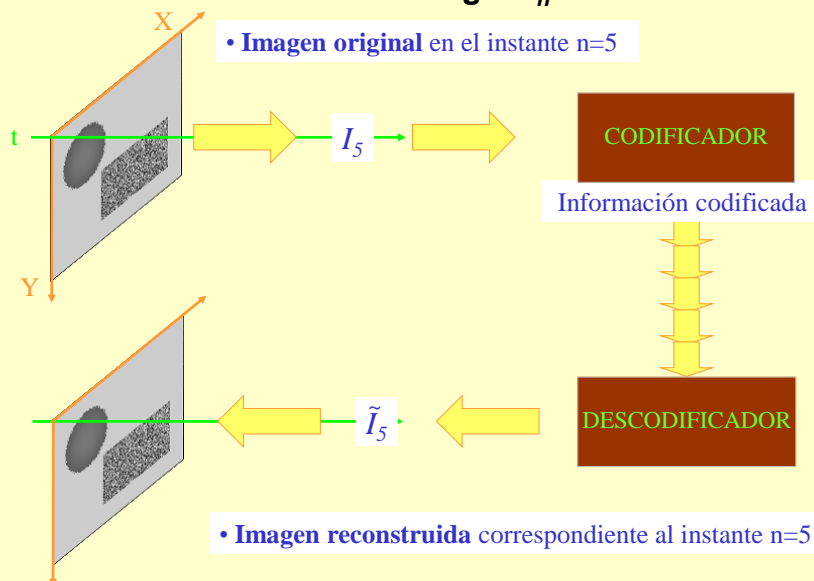
## Representación y Codificación de Señales Audiovisuales en Televisión Digital

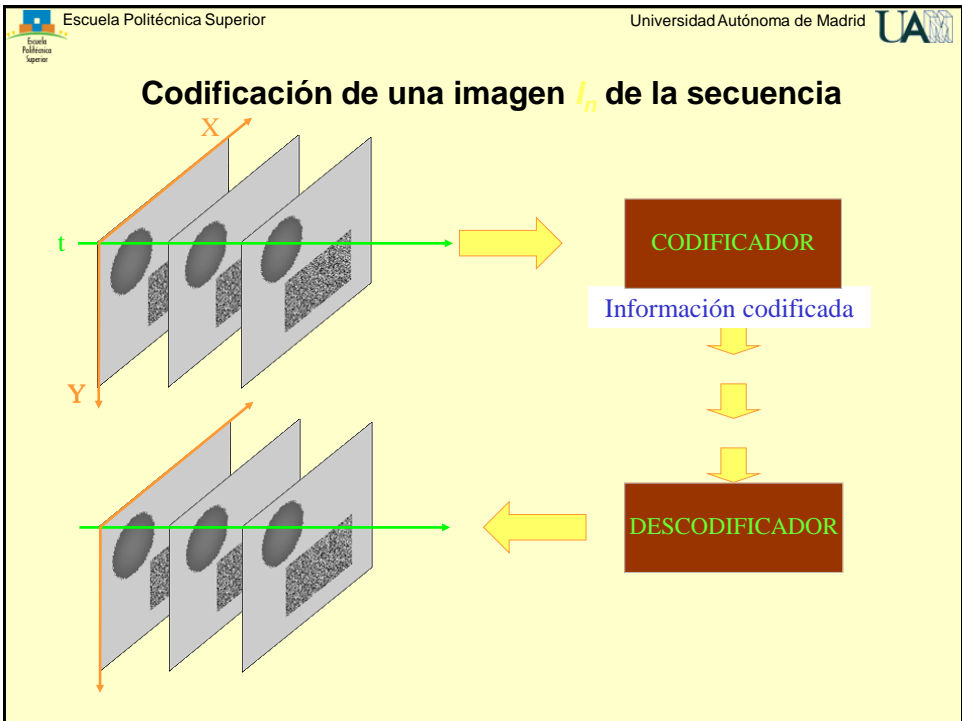
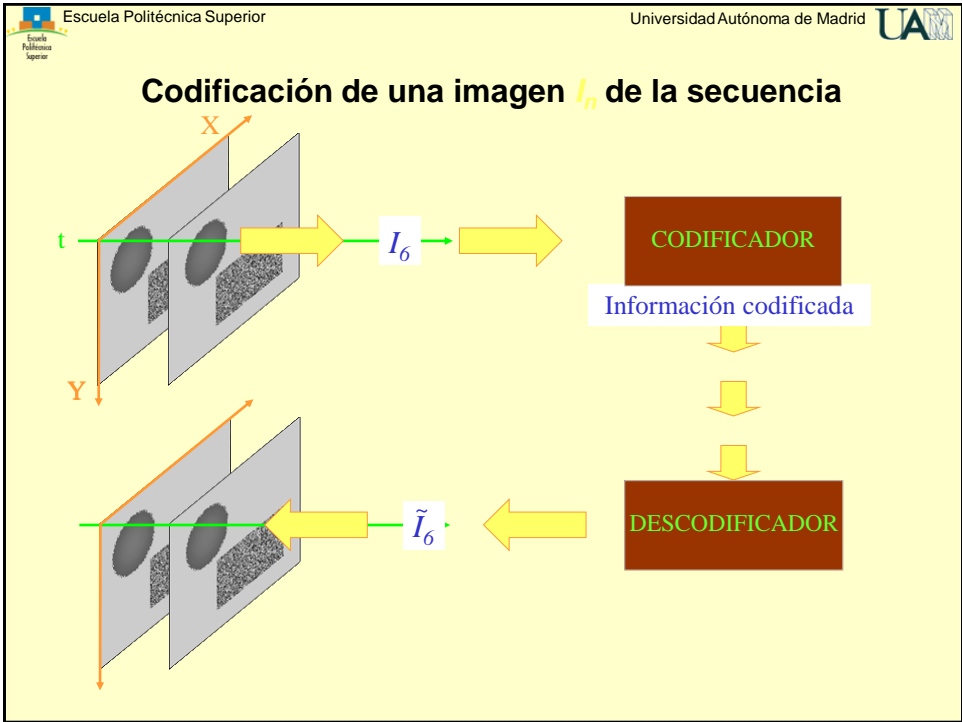
### Compresión de vídeo híbrido transformacional

- *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

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional (5)*

### Codificación de una imagen $I_n$ de la secuencia






Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Representación y Codificación de Señales Audiovisuales en Televisión Digital

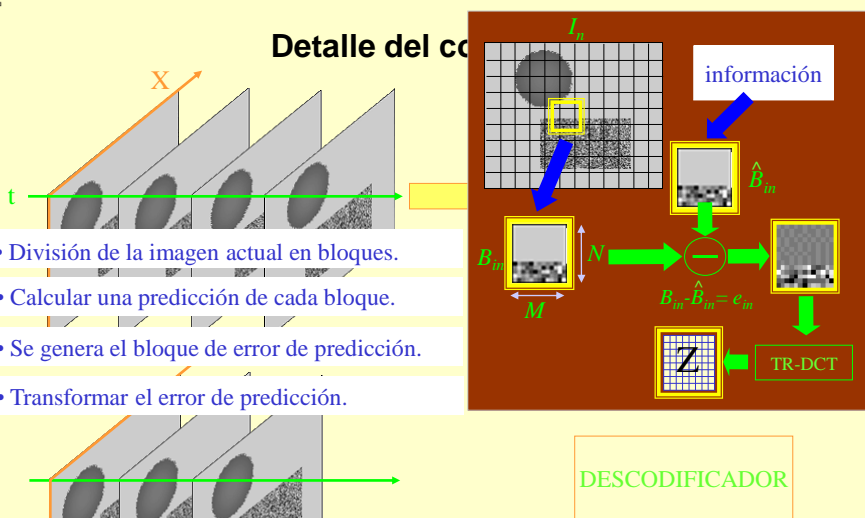
### Compresión de vídeo híbrido transformacional

- *Introducción*
- **Descripción del sistema.**
  - *Codificación de una imagen de la secuencia.*
  - **Detalle del codificador y del 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, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional* (9)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

### Detalle del codificador



- División de la imagen actual en bloques.
- Calcular una predicción de cada bloque.
- Se genera el bloque de error de predicción.
- Transformar el error de predicción.

DESCODIFICADOR

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Detalle del codificador

The diagram illustrates the encoder process. On the left, a 3D stack of image blocks is shown with axes X, Y, and t. A green arrow points from the blocks to a 10x10 prediction matrix. A blue cloud contains the question "¿Qué tipo?". Below the matrix is a list of five bullet points:
 

- División de la imagen actual en bloques.
- Calcular una predicción de cada bloque.
- Se genera el bloque de error de predicción.
- Transformar el error de predicción.
- Cuantificar los coeficientes de la transformación.

 To the right, a 10x10 grid shows the prediction matrix. Below it, two matrices are shown: Q (Quantization) and Z (Coefficients). A box labeled "DESCODIFICADOR" has an arrow pointing to the Z matrix. A "TR-DCT" block is shown with an arrow pointing to the Z matrix. A "Q-SVH" block is shown with an arrow pointing to the Q matrix.

61	0	7	-1	1	0	0	0	0	-33	-18					
4	0	-6	1	-2	0	0	0	52	-41	-44	43	43			
-4	0	3	0	0	0	0	0	-59	-7	49	15	30	4	-20	-34
3	1	0	0	0	0	0	0	48	31	-13	-7	-8	11	4	39
-2	-2	0	0	0	0	0	0	-38	-49	-19	4	-5	-15	17	-45
0	0	0	0	0	0	0	0	23	61	38	3	21	33	-21	29
0	0	0	0	0	0	0	0	-26	-47	-49	22	-18	-34	30	-35
0	0	0	0	0	0	0	0	2	33	26	21	27	26	0	17

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Detalle del decodificador

The diagram illustrates the decoder process. On the left, a 3D stack of image blocks is shown with axes X, Y, and t. A yellow arrow points from the blocks to a box labeled "CODIFICADOR". Above the box is a blue cloud with "información" and "¿Qué tipo?". Below the box is "Información codificada". Three yellow arrows point down to a box labeled "DESCODIFICADOR".

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Detalle del decodificador

información

¿Qué tipo?

• Reconstruir los coeficientes de la transform.

$Q \rightarrow Q^{-1} \cdot SVH \rightarrow \tilde{Z}$

990	9	72	-16	35	51	-33	-18	976	0	70	-16	24	0	0	0
52	-9	-88	26	-41	-44	43	43	48	0	-84	19	-52	0	0	0
-59	-7	49	15	30	4	-20	-34	-56	0	48	0	0	0	0	0
-48	31	-13	-7	-8	11	4	39	42	17	0	0	0	0	0	0
-38	-49	-19	4	-5	-15	17	-45	-36	-44	0	0	0	0	0	0
23	61	38	3	21	53	-21	29	0	0	0	0	0	0	0	0
-26	-47	-49	22	-18	-34	30	-35	0	0	0	0	0	0	0	0
2	33	28	21	27	26	0	17	0	0	0	0	0	0	0	0

Descodificador

Escuela Politécnica Superior Universidad Autónoma de Madrid

### Detalle del decodificador

información

¿Qué tipo?

• Reconstruir los coeficientes de la transform.

• Transformada inversa del error de predicción.

• Se determina la predicción del bloque.

• Generación del bloque reconstruido.

$Q \rightarrow Q^{-1} \cdot SVH \rightarrow \tilde{Z} \rightarrow TR^{-1}$

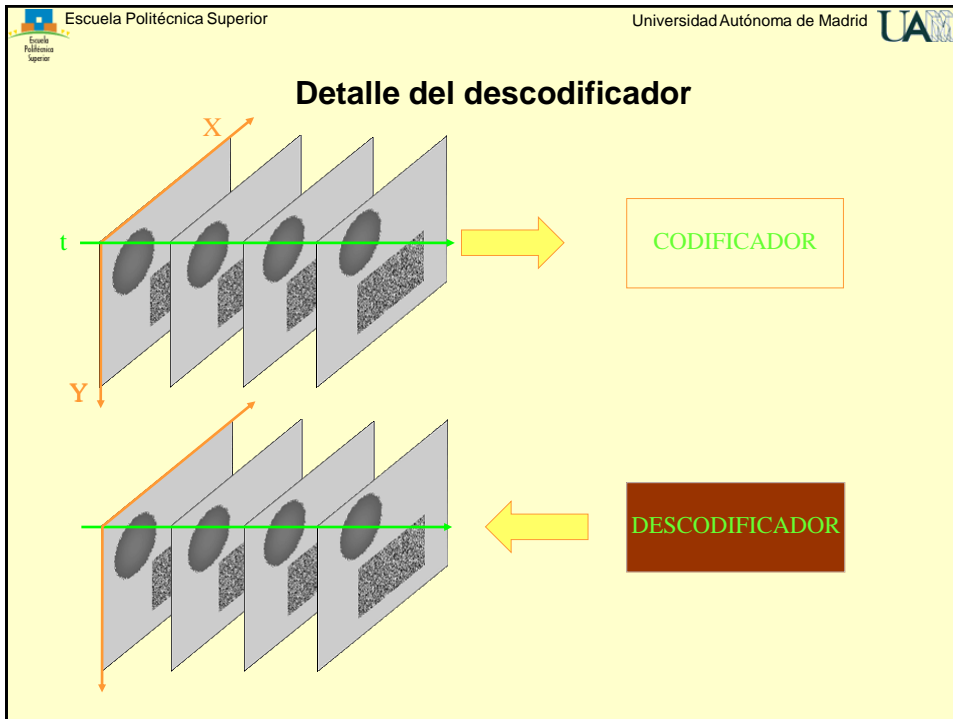
$\tilde{B}_m = \tilde{z}_m + \hat{B}_m$

$\tilde{z}_m$

$\hat{B}_m$

Información disponible

Descodificador



Escuela Politécnica Superior Universidad Autónoma de Madrid

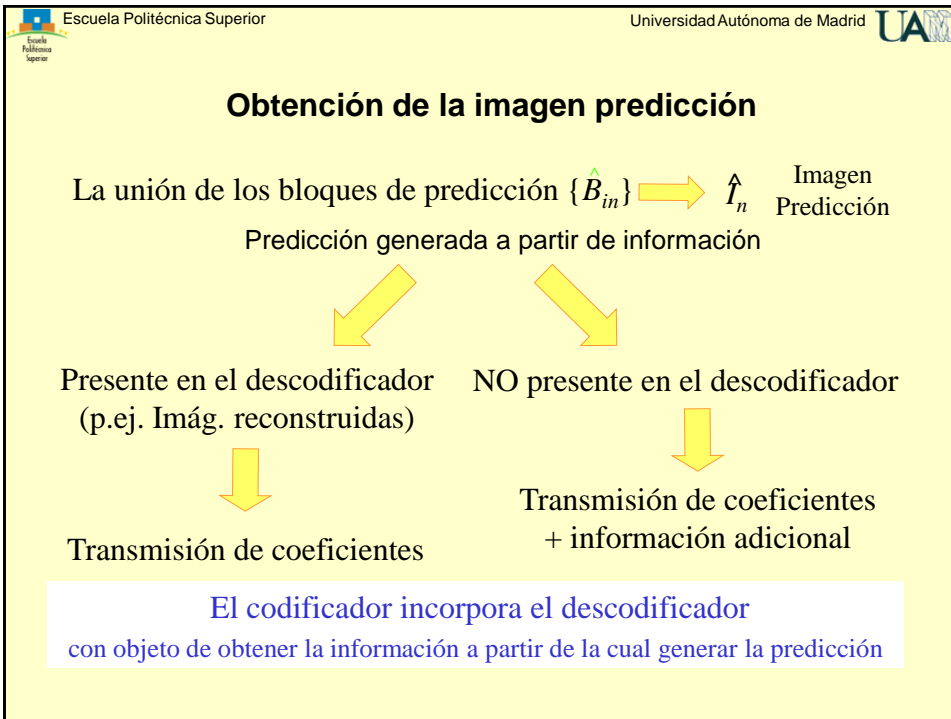
## Representación y Codificación de Señales Audiovisuales en Televisión Digital

### Compresión de vídeo híbrido transformacional

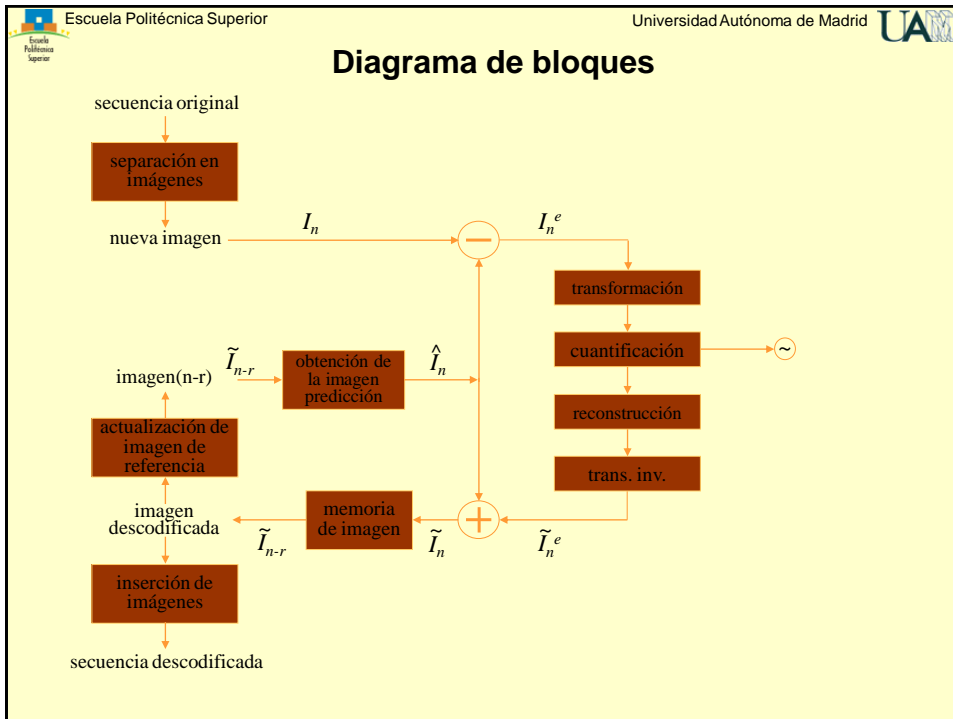
- *Introducción*
- **Descripción del sistema.**
  - *Codificación de una imagen de la secuencia.*
  - *Detalle del codificador y del decodificador.*
  - **Obtención de la imagen predicción.**
  - Diagrama de bloques.
- Estrategias de predicción.
  - Predicción temporal.
  - Introducción del movimiento 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (16)





- 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 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (18)



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, 2008-2009) 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**

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Predicción temporal

Extensión directa al tratamiento de secuencias

píxeles próximos en el tiempo → píxeles similares

The diagram illustrates a sequence of four frames in a 3D perspective. The vertical axis is labeled 'Y' and the horizontal axis is labeled 'X'. A green arrow points from left to right across the frames, indicating the direction of temporal correlation. The frames show a dark circular shape and a textured rectangular area that remain relatively stationary across the sequence.

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Predicción temporal

Extensión directa al tratamiento de secuencias

píxeles próximos en el tiempo → píxeles similares

The diagram illustrates a sequence of four frames in a 3D perspective. The vertical axis is labeled 'Y' and the horizontal axis is labeled 'X'. A blue arrow points from left to right across the frames, indicating the direction of movement. A yellow arrow points from left to right across the frames, indicating the direction of temporal correlation. The frames show a dark circular shape and a textured rectangular area that move across the sequence.

**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.

Dirección || eje de tiempos

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Predicción temporal

Extensión directa al tratamiento de secuencias

secuencia original

píxeles próximos

¿Se podría mejorar la predicción introduciendo información del movimiento presente en la escena?

Dirección || eje de tiempos

...es similares

**Conclusiones:**

- 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.

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 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (26)

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Introducción del movimiento en la predicción

Codificación sin emplear movto.

$I_n$   $\hat{I}_n = \tilde{I}_{n-1}$

↓ error de predicción ↑ error de predicción

TR-DCT

analicemos la secuencia sintética

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Introducción del movimiento en la predicción


Codificación sin emplear movto.

$I_n$   $\hat{I}_n = \tilde{I}_{n-1}$

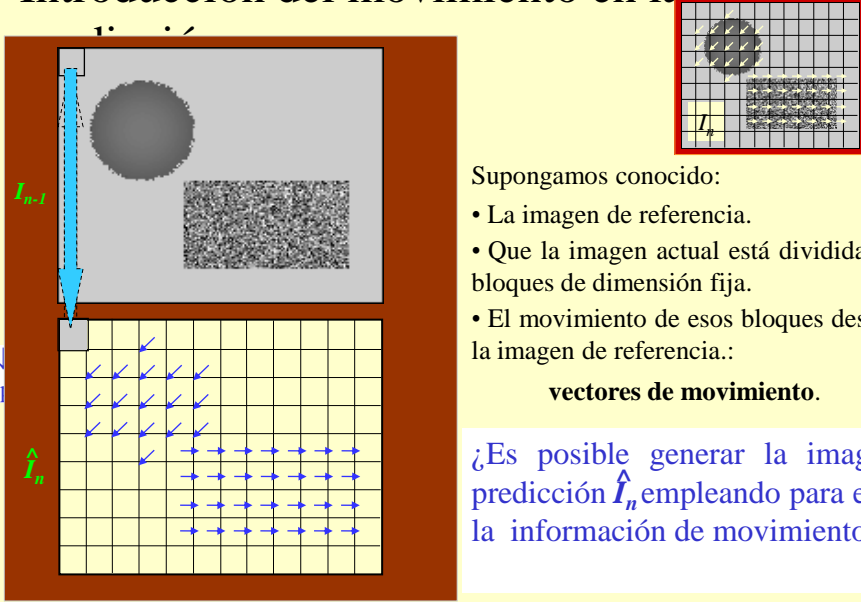
↓ error de predicción ↑ error de predicción

TR-DCT

Esta es la imagen que empleamos como referencia para analizar el movimiento...

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Introducción del movimiento en la




Supongamos conocido:

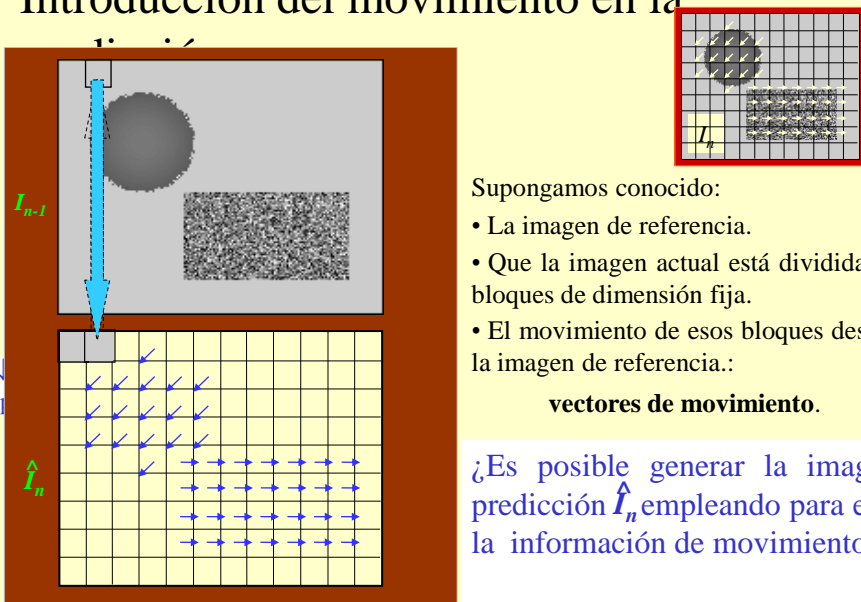
- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

**vectores de movimiento.**

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

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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?

Escuela Politécnica Superior Universidad Autónoma de Madrid

## 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?

Escuela Politécnica Superior Universidad Autónoma de Madrid

## 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.:

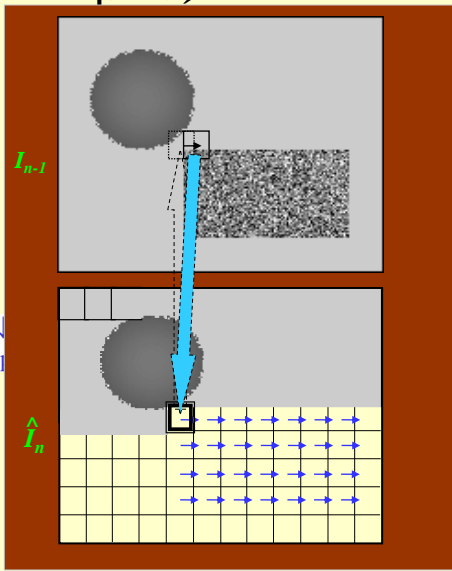
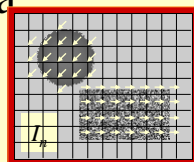
**vectores de movimiento.**

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



Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Introducción del movimiento en la





Supongamos conocido:

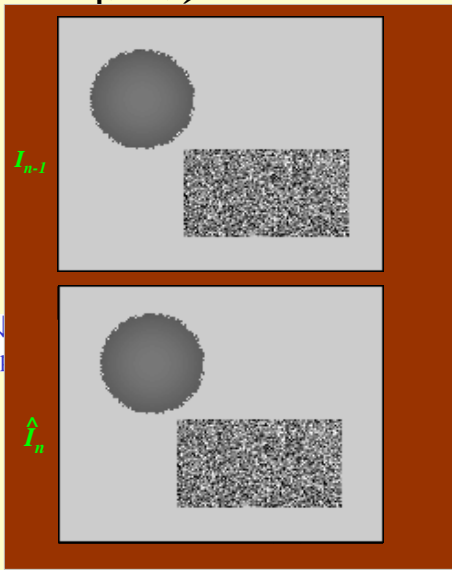
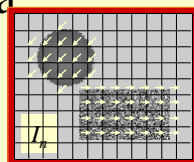
- La imagen de referencia.
- Que la imagen actual está dividida en bloques de dimensión fija.
- El movimiento de esos bloques desde la imagen de referencia.:

**vectores de movimiento.**

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

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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?

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 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (35)

Escuela Politécnica Superior Universidad Autónoma de Madrid

## 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.

$\hat{I}_n = I_{n-1} + VM$

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Estimación y compensación

**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.

!!!situación óptima!!!  
error de predicción  
NULO...

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 descodificador pueda generar la imagen predicción

- Información de movimiento.

Para que la imagen predicción sea idéntica en el codificador y en el descodificador

**La imagen o imágenes de referencia deben ser imágenes reconstruidas.**

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 a current frame  $I_n$  and a reference frame  $\tilde{I}_{n-1}$ . A motion vector  $VM$  is estimated from the two frames. This vector is used for motion compensation to produce a predicted frame  $\hat{I}_n = I_{n-1} + VM$ . The prediction error is calculated as  $I_n - \hat{I}_n$ . This error is then processed by a TR-DCT block.

Escuela Politécnica Superior Universidad Autónoma de Madrid

## Estimación y compensación

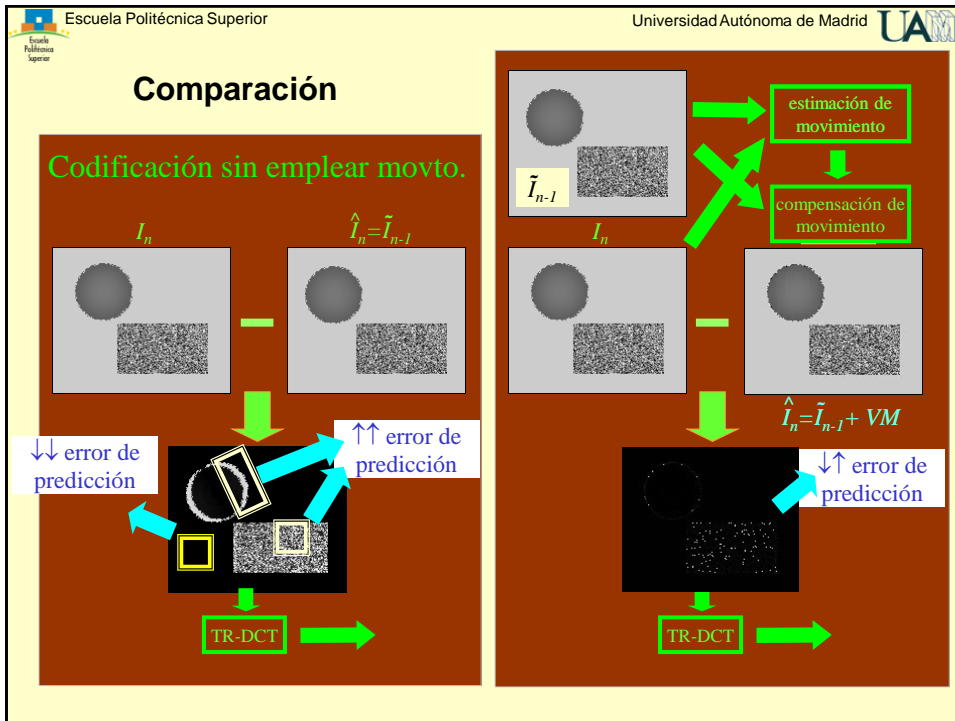
**¿Qué habría que transmitir?**  
 El error de predicción y la información necesaria para que el decodificador pueda generar la imagen predicción

- 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 a current frame  $I_n$  and a reference frame  $\tilde{I}_{n-1}$ . A motion vector  $VM$  is estimated from the two frames. This vector is used for motion compensation to produce a predicted frame  $\hat{I}_n = \tilde{I}_{n-1} + VM$ . The prediction error is calculated as  $I_n - \hat{I}_n$ . This error is then processed by a TR-DCT block.



Escuela Politécnica Superior Universidad Autónoma de Madrid

## Pero, ¿Existen desventajas?

**El codificador**

- ★ Incorpora la **estimación de movimiento**
  - ↑ Complejidad
  - ↑ Carga computacional

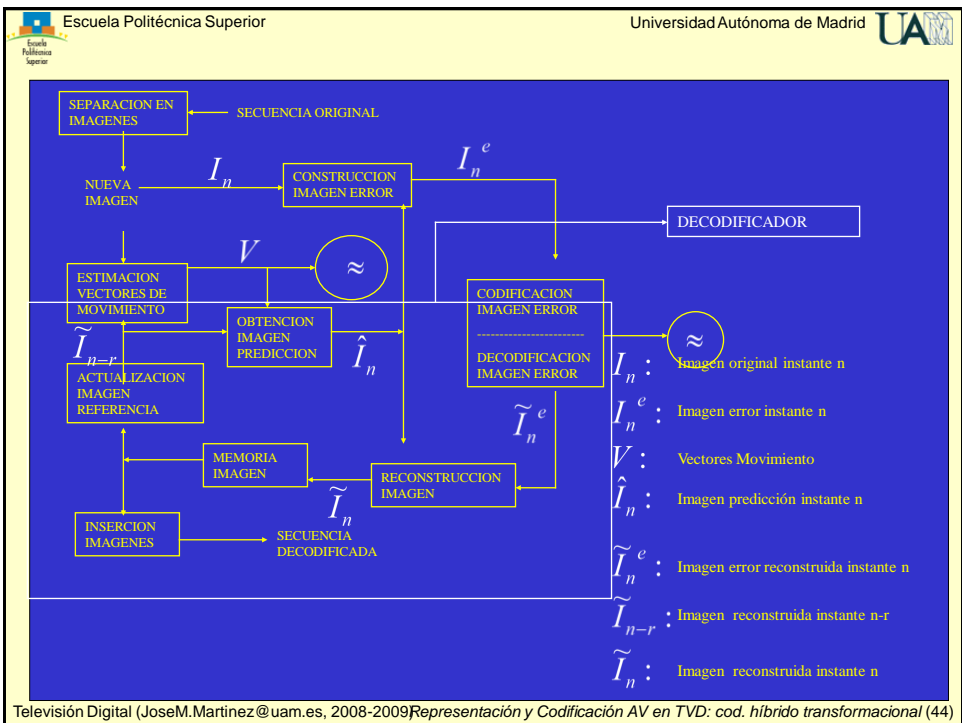
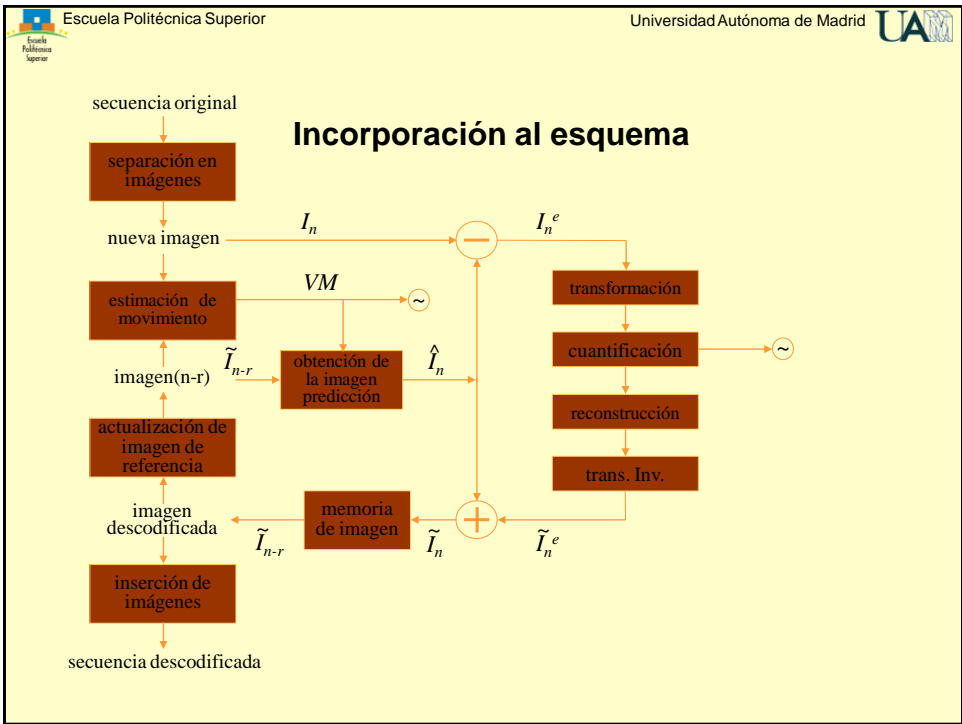
↓

↑ **Precio** : función de lo eficiente que sea la estimación


- Tiene que realizar la compensación, igual que el descodificador...

¿Por qué se utiliza?

En presencia de movimiento permite **reducir** de un modo muy significativo el **ancho de banda** requerido **para la transmisión**



Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (44)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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)

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional (45)*

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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 que coincidir
    - Cambio de iluminación genera un flujo óptico
    - Esfera uniforme que rota sobre su eje no genera un flujo óptico (en condiciones de iluminación homogénea)

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional (46)*

## 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (47)

## 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.

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (48)



## Estimación para codificación

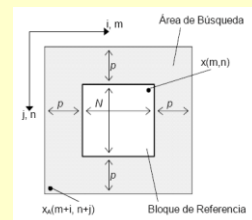
- el objetivo es reducir el error de predicción
- se trabaja por bloques
  - Más robusto frente al ruido
  - Menos vectores a transmitir
  - Consistente con la codificación por macrobloques DCT
- Se busca para la imagen a codificar el bloque que más se parece en la imagen de referencia (en el caso más 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

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

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Ajuste de bloques (II)

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

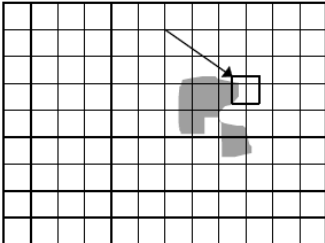


Imagen previa

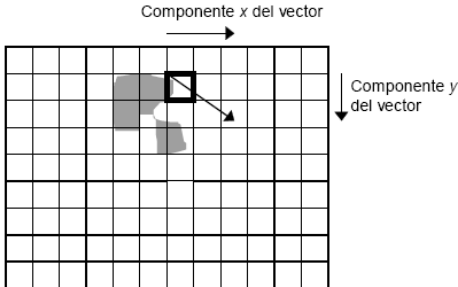



Imagen actual

Demostración

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (51)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (52)

## 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-)

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

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

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

- 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
  - o 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.
  - o Three Step Search (TSS)
 

"Motion compensated interframe coding for video conferencing". T. Koga, K. Inuma, A. Hirano, Y. Iijima, T. Ishiguro. Proc. Nat. Telecommun. Conf. New Orleans, L.A. Nov 1981.
  - o 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  $p/2$ .

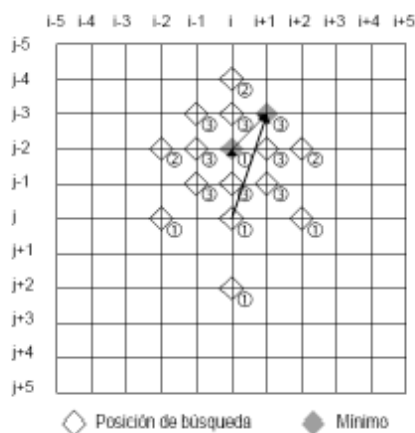
En la segunda iteración se calculan tres posiciones más con el mismo paso y tomando como posición central la del mínimo obtenido en la búsqueda anterior.

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.

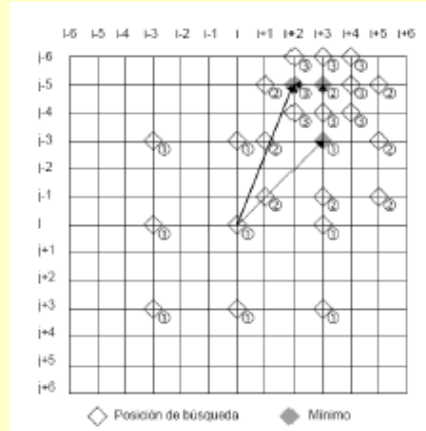
El mínimo de esta última búsqueda se toma como posición óptima.



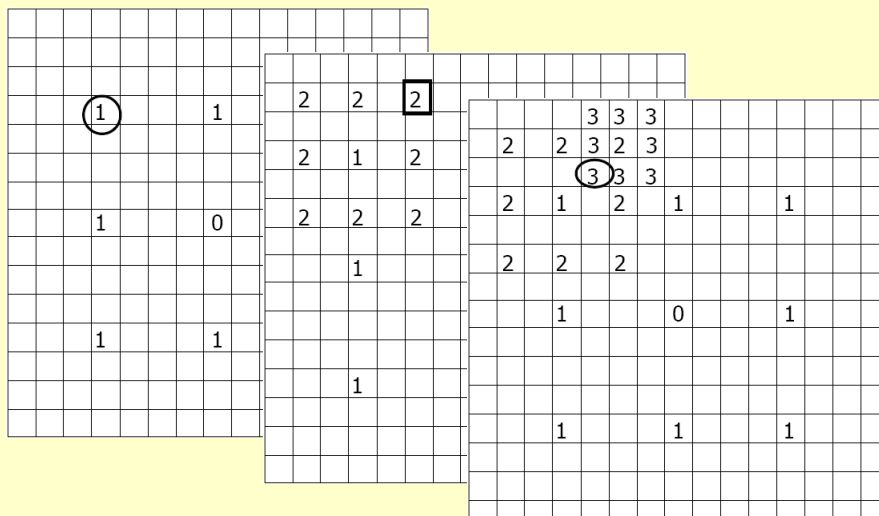
## Algoritmos de ajuste de bloques Búsqueda logarítmica (III) (TSS)

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 pixel ( $kp/q$  indica el número de iteraciones)
- TSS ( $p=q=7 \Rightarrow k=3$ )



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



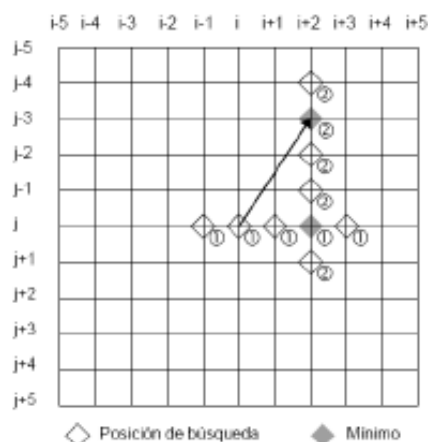
## Algoritmos de ajuste de bloques Búsqueda logarítmica (V) (TSS)


- Siendo  $p=q$ 
  - o Número de comparaciones por bloque  $(8k+1)$   
vs  $(2p+1)(2q+1)$
  - o Píxeles del bloque  $MN$
  - o  $(8k+1)MN$  evaluaciones de la función de coste
  - 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,  $F \times C$  píxeles
  - o Operaciones/seg=  $\{(FCT)/MN\} 3 (8k+1)MN = FCT 3 (8k+1)$
  - o  $T=25$ ,  $C=720$ ,  $F= 576$ ,  $p=q=15 \Rightarrow 1,03$  G OPS ( $p=q=7 \Rightarrow 0,77$  G OPS)  
vs  $29,98$  G OPS ( $p=q=7 \Rightarrow 6,99$  G OPS)

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

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

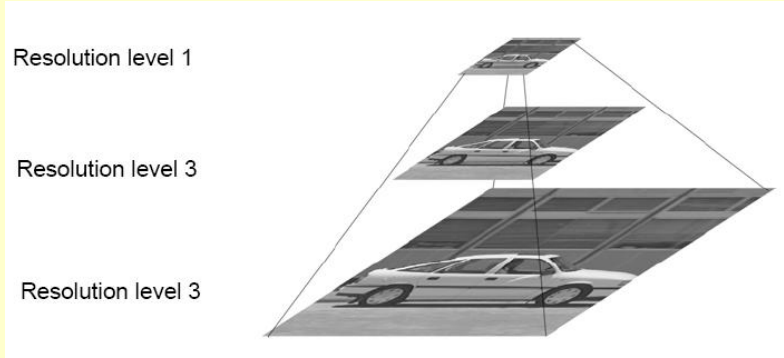
Se comienza la búsqueda en la dirección horizontal (eje  $i$  en la figura) con paso de un píxel comenzando por la posición central y las dos adyacentes y continuando en el sentido de la que produzca menor error hasta que se encuentre un mínimo. A partir de dicho mínimo, se continúa la búsqueda en dirección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




Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Algoritmos de ajuste de bloques Búsqueda Piramidal (método multiresolución) (I)

Exhaustiva sobre imagen reducida con bloque reducido



Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (61)

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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.

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (62)

Escuela Politécnica Superior Universidad Autónoma de Madrid

### 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.

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (63)

Escuela Politécnica Superior Universidad Autónoma de Madrid

### 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  $M_L = M/2^L$ ,  $N_L = N/2^L$ ,  $p_L = p/2^L$ ,  $q_L = q/2^L$  =>  $u_L = u/2^L$ ,  $v_L = v/2^L$
- 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 (=> 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.

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (64)



## Algoritmos de ajuste de bloques

### Búsqueda Piramidal (método multiresolución) (V)

Coste (nivel L búsqueda exhaustiva)

- Nivel L
  - o Número de comparaciones nivel L :  $(2p_L+1)(2q_L+1)$
  - o Píxeles del bloque nivel L :  $M_L N_L$
  - o Píxeles de imagen nivel L :  $F_L C_L$
  - 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 Nivel L: Operaciones/seg =  $F_L C_L T^3 (2p_L+1)(2q_L+1)$
- Resto niveles (L-j, j = 1..L)
  - o Número de bloques nivel L-j : 9
  - o Resto igual
  - o 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)
  - o vs TSS: 1,03 G OPS (p=q=7 => 0,77 G OPS)
  - o 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



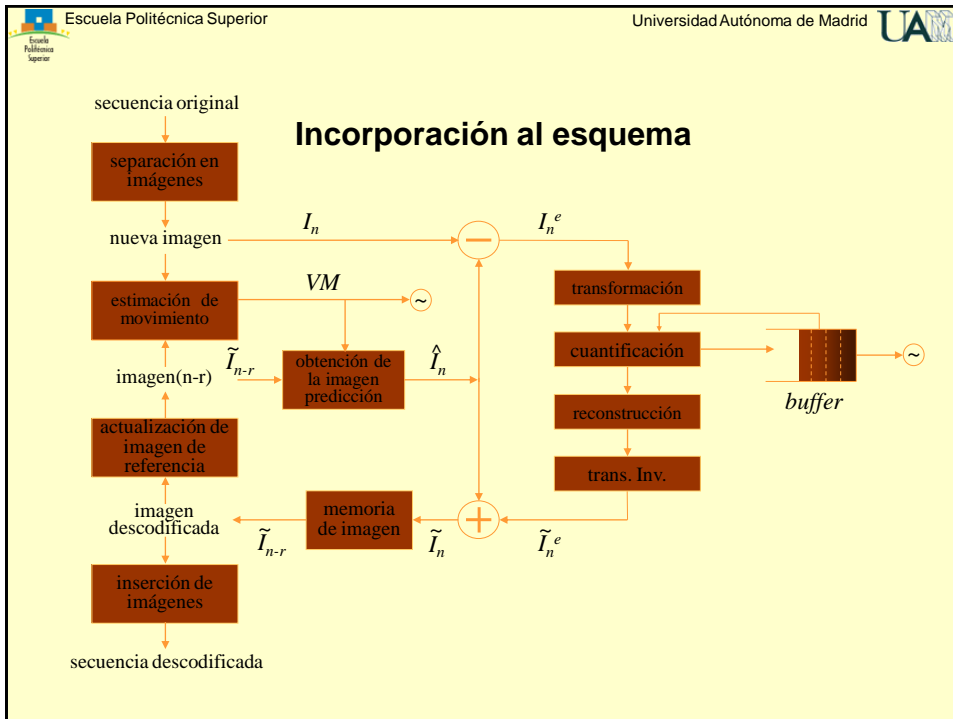
~~velocidad binaria fija~~

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

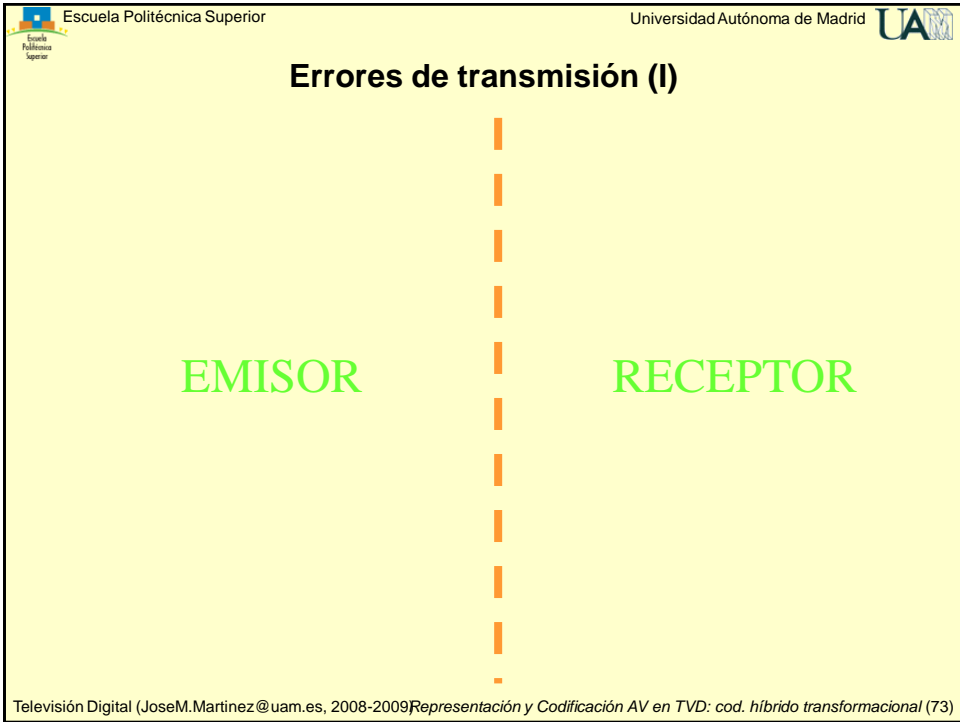
### Compresión de vídeo híbrido transformacional

- *Introducción*
- *Descripción del sistema.*
  - *Codificación de una imagen de la secuencia.*
  - *Detalle del codificador y del 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, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (T2)


Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Errores de transmisión (I)

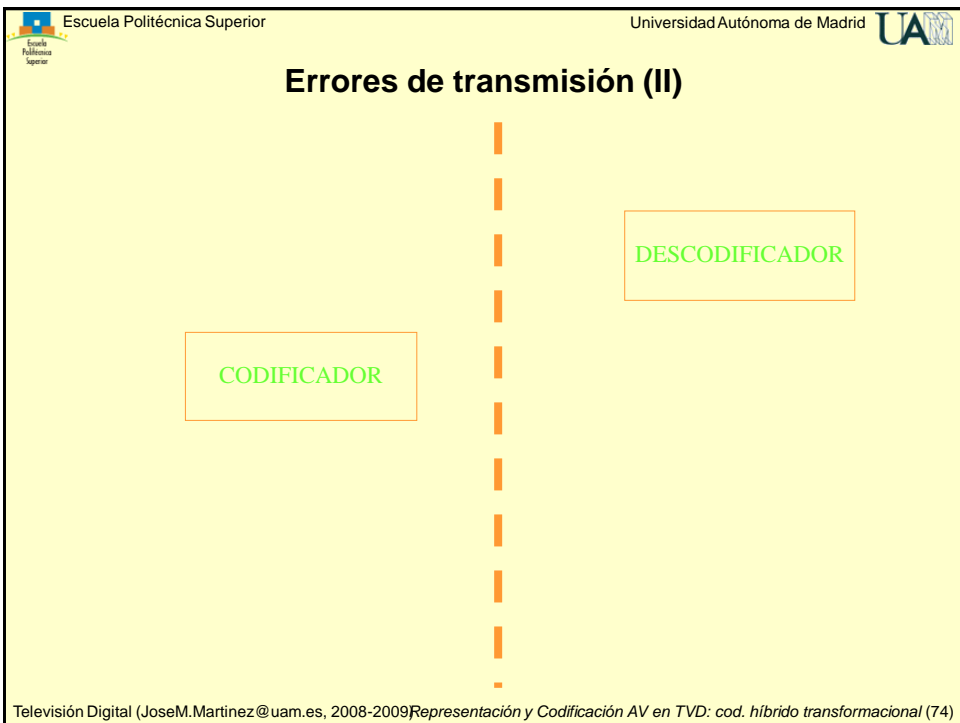


EMISOR RECEPTOR

Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (73)

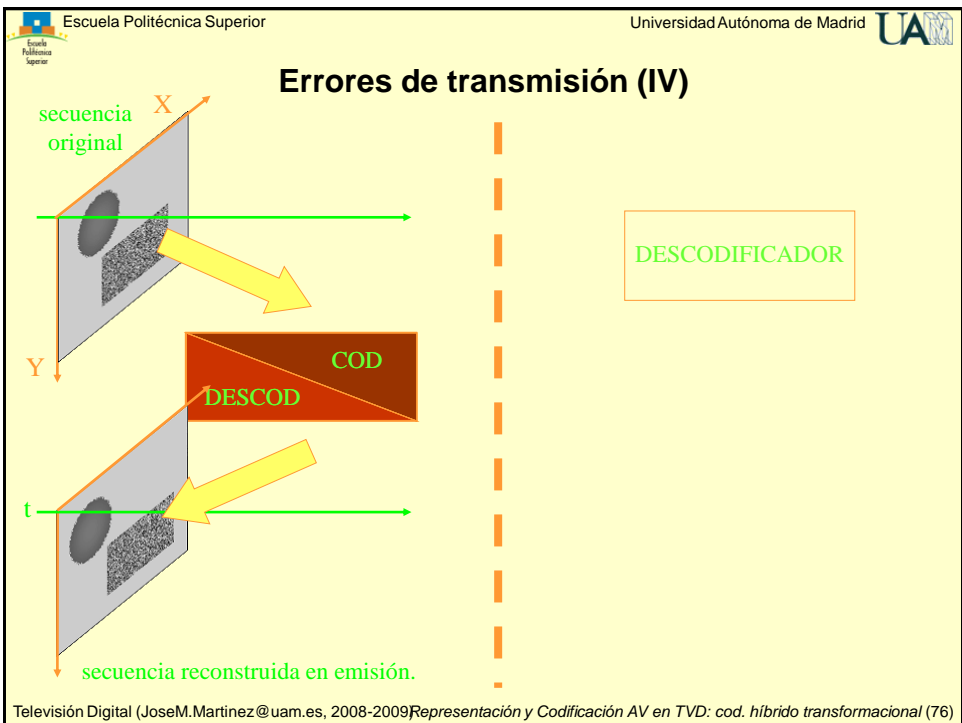
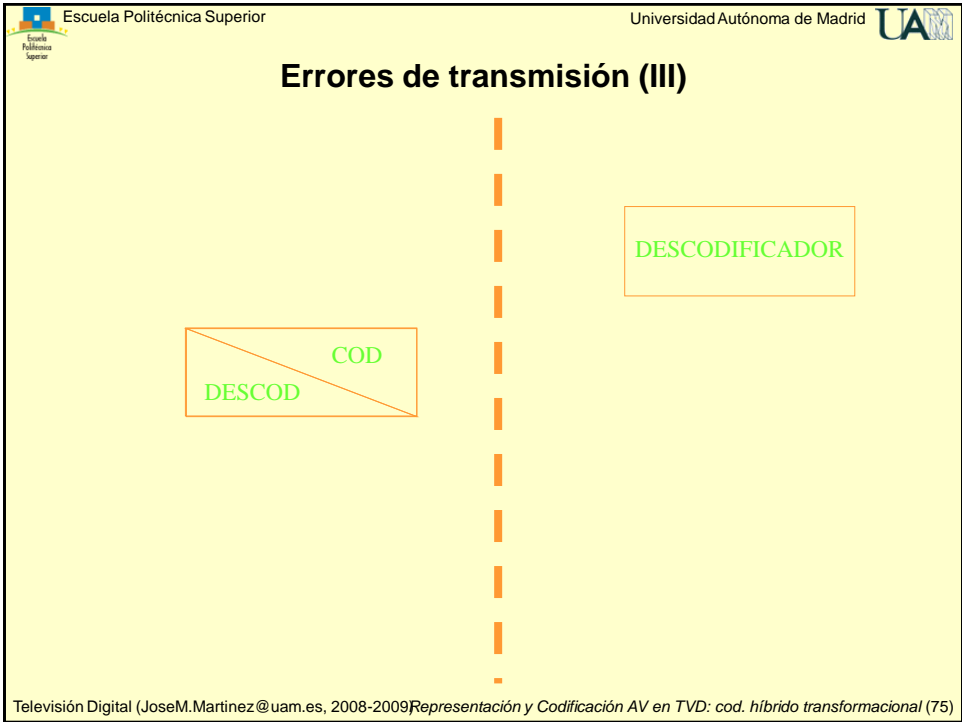
Escuela Politécnica Superior Universidad Autónoma de Madrid 

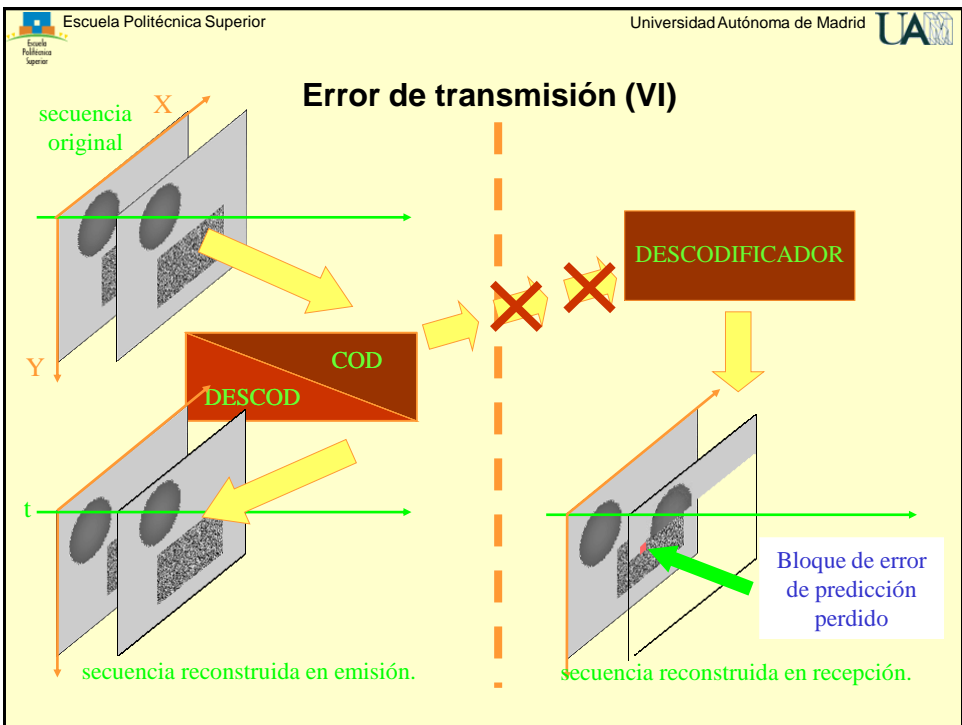
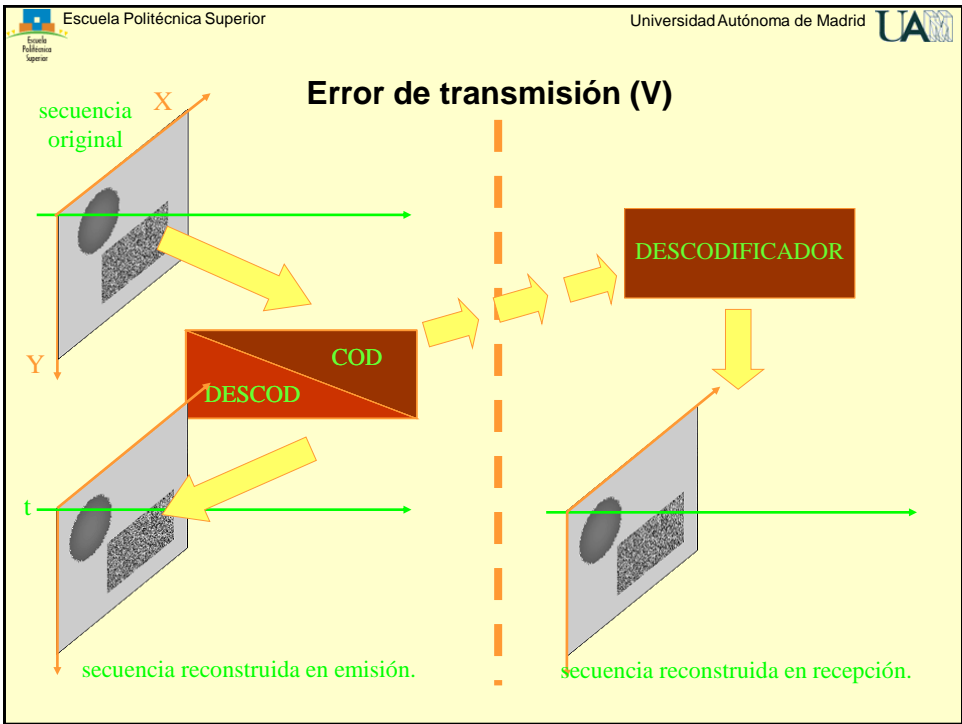
## Errores de transmisión (II)

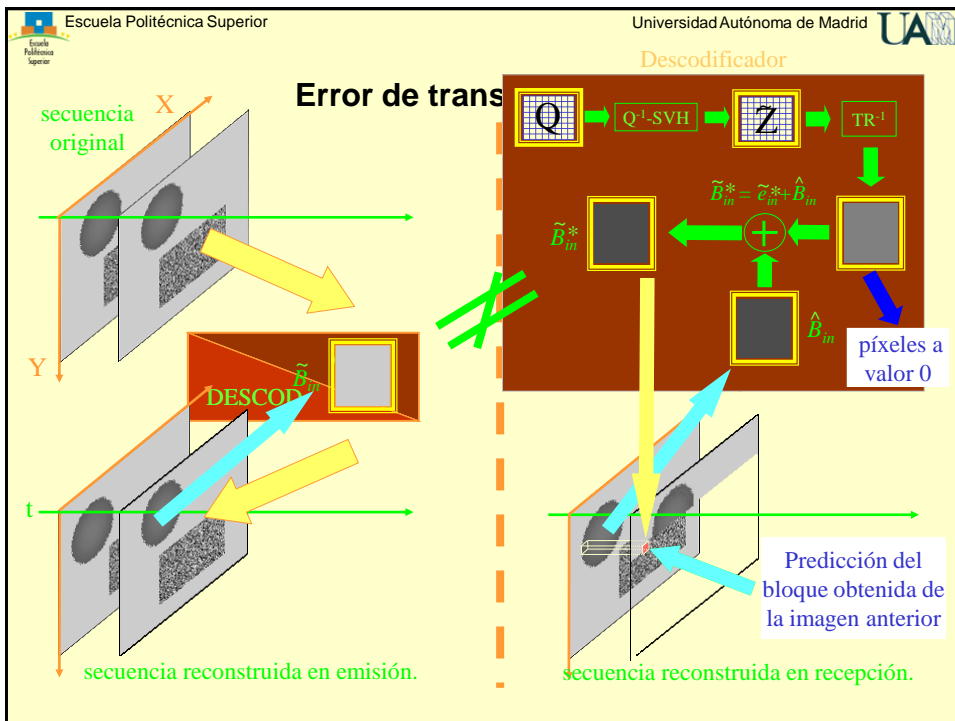
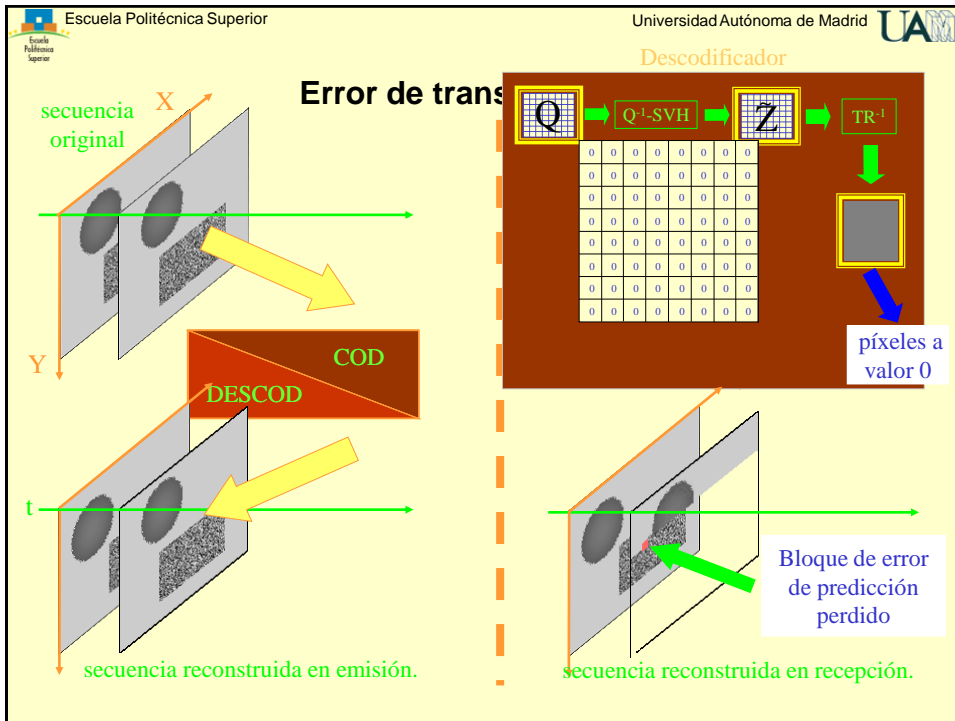


CODIFICADOR DESCODIFICADOR

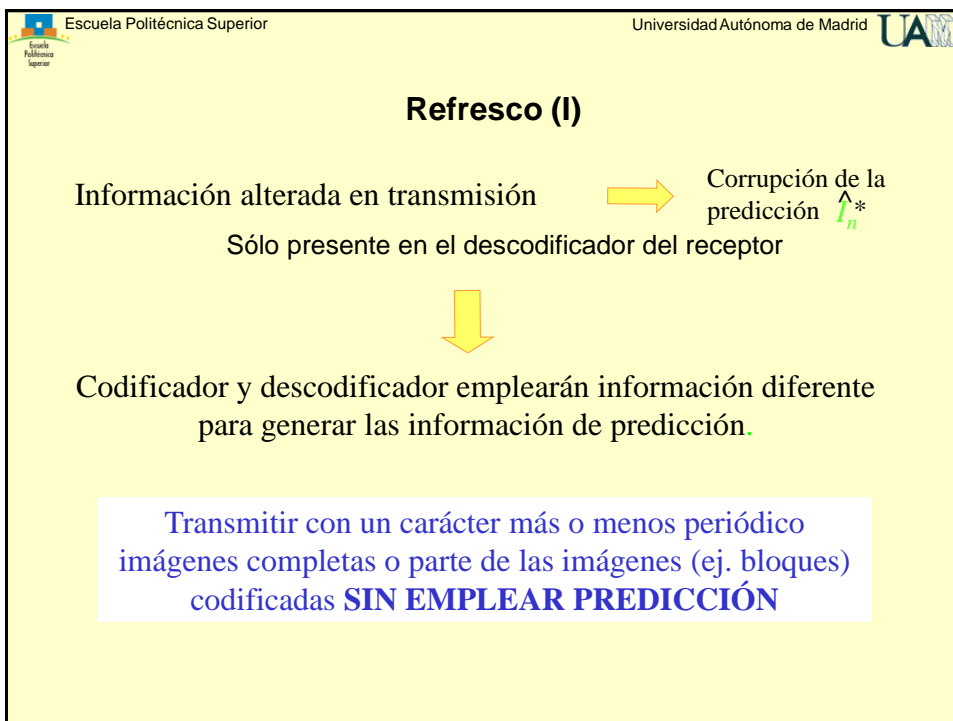
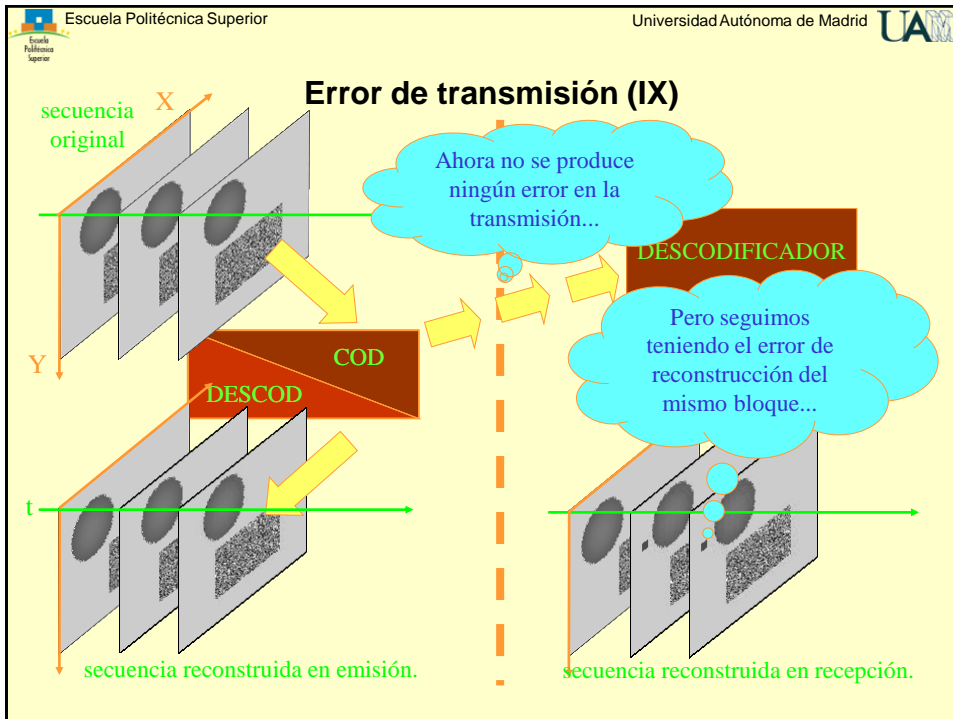
Televisión Digital (JoseM.Martinez@uam.es, 2008-2009) Representación y Codificación AV en TVD: cod. híbrido transformacional (74)

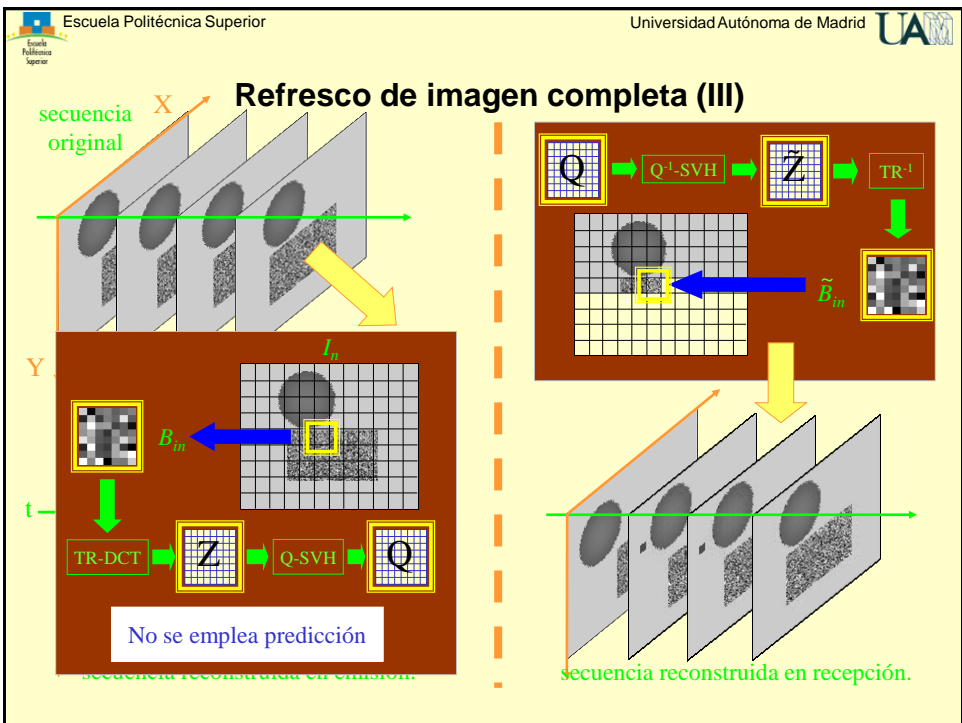
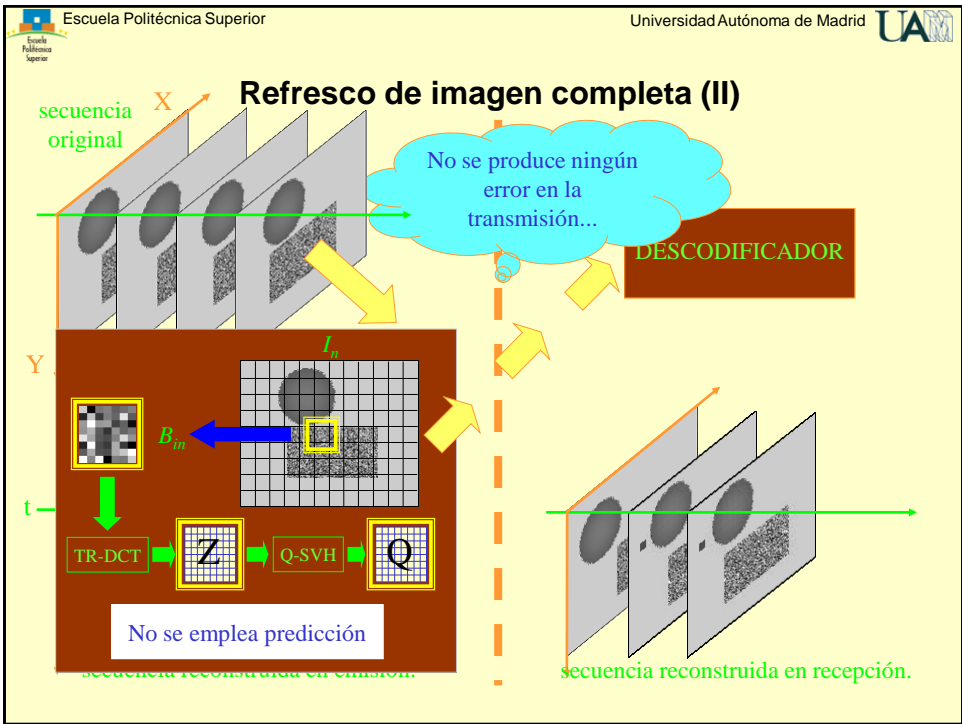


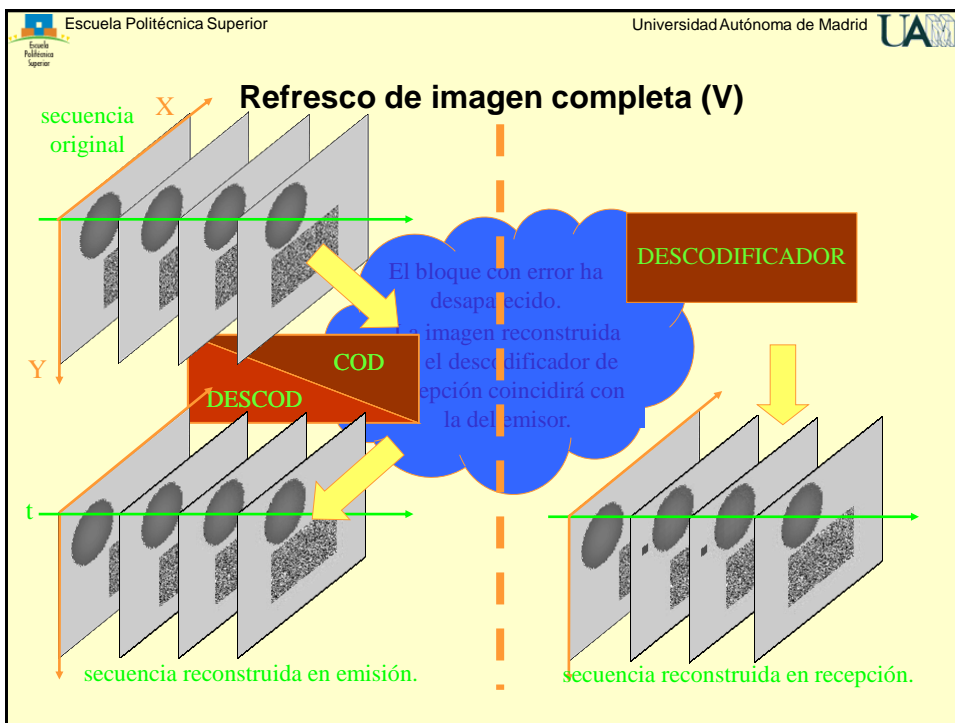
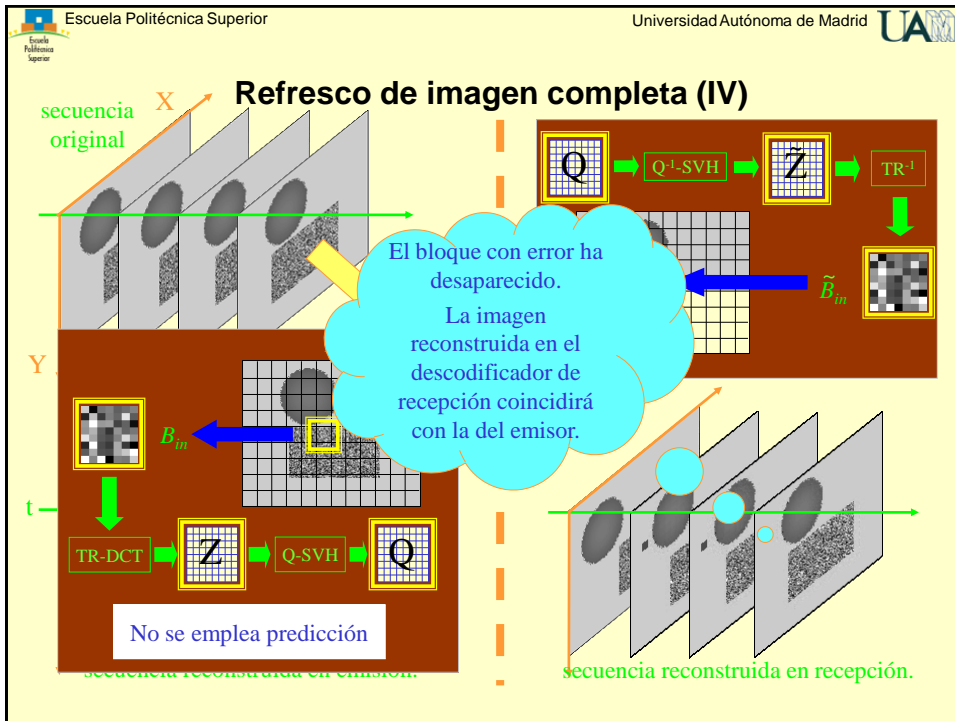













Escuela Politécnica Superior Universidad Autónoma de Madrid 

## 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 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, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional (87)*

Escuela Politécnica Superior Universidad Autónoma de Madrid 

## Líneas de mejora

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

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

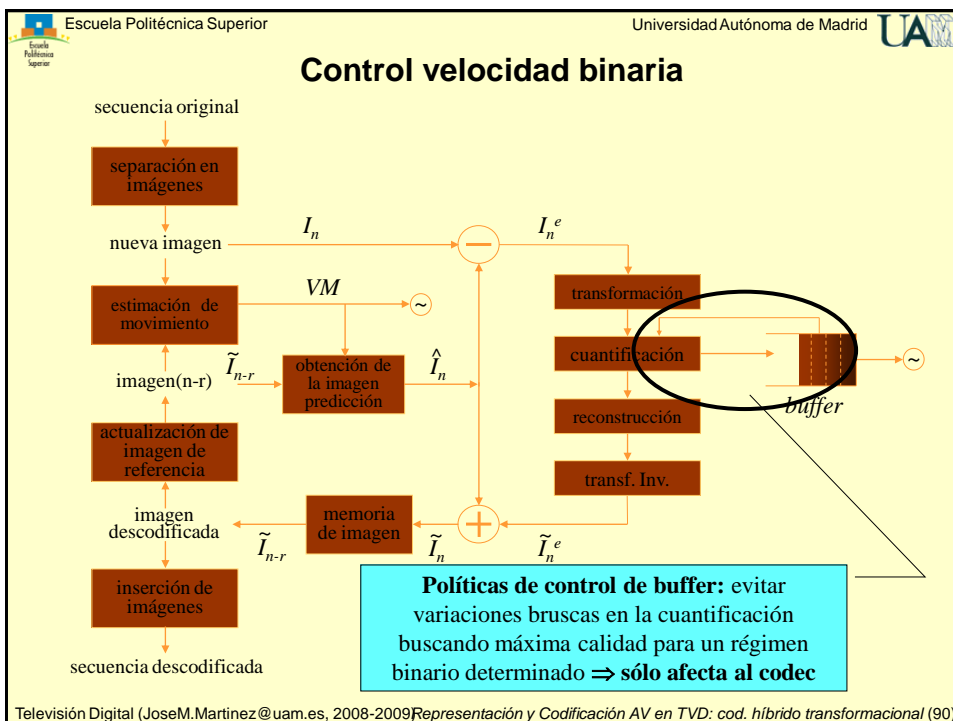
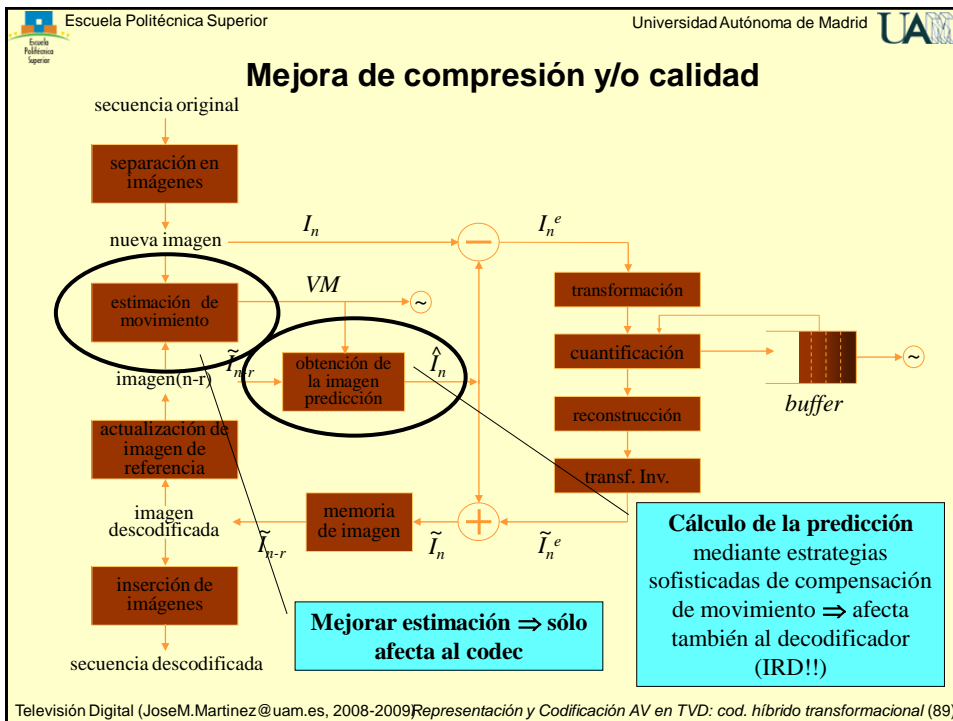
MPEG-4: H.263- pero introduce la codificación de objetos

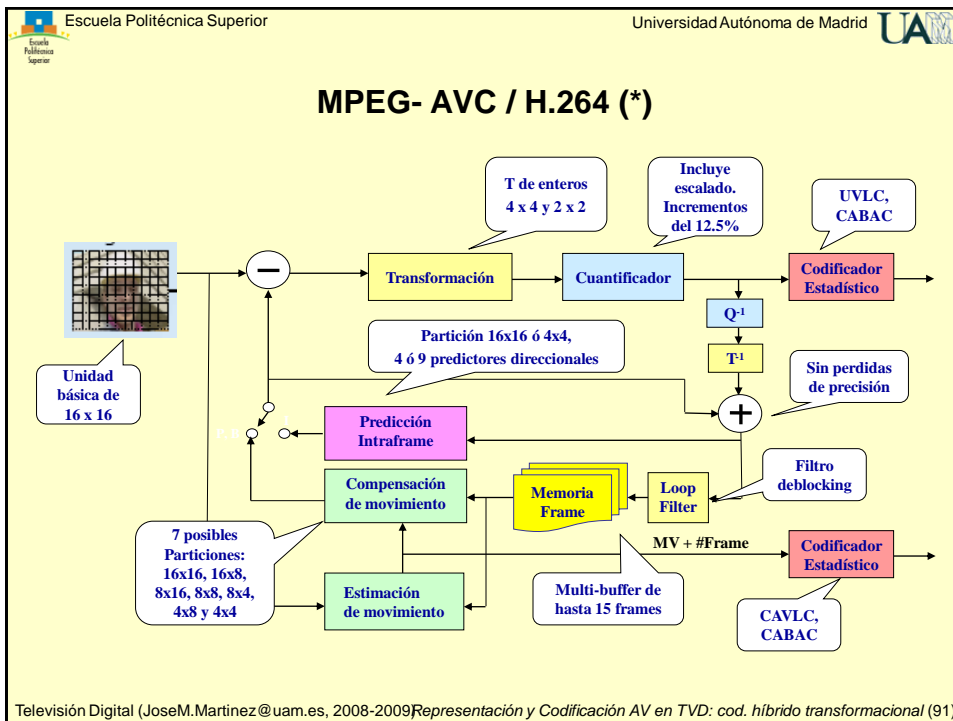
- Necesidad de sincronización y composición en terminal

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, 2008-2009) *Representación y Codificación AV en TVD: cod. híbrido transformacional (88)*





Escuela Politécnica Superior Universidad Autónoma de Madrid

## Conclusiones

- ★ Introducción de un esquema general de codificación.
- ★ Implicaciones de la utilización de estrategias predictivas.
- ★ Incorporación del movimiento al cálculo de la predicción.
- ★ Procesos de estimación y compensación de movimiento.
- ★ Sistemas de contención y control de la velocidad binaria.
- ★ Errores de transmisión: Necesidad de refresco.
- ★ Líneas de mejoras



## 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.