

TELEVISIÓN DIGITAL 2011/12	
Práctica 3: Análisis de flujos MPEG-2	
Fecha de realización:	Fecha tope de entrega de la memoria:

1 Introducción

El campo principal de aplicación del estándar MPEG-2 Video (ISO/IEC 13818) es la TVD (contribución y distribución) y los DVDs, pero su carácter genérico lo hace útil en diversos entornos. Dada su complejidad, desde el punto de vista industrial, se implementan dispositivos con un subconjunto de herramientas (perfiles) y niveles de complejidad (niveles). Para TVD se utiliza el MainProfile@MainLevel.

MPEG-2 Video sigue un esquema híbrido predictivo: separación en MB/Bs, Predicción temporal (estimación movimiento a nivel MB), DCT de la señal residual (a nivel B), recorrido en zig-zag (diferente según Y o C), cuantificación mediante matriz de visibilidad, VLC, y control de buffer.

El estándar define las siguientes estructuras de datos organizadas jerárquicamente y directamente relacionadas con la estructuración del flujo binario resultante de la codificación de la secuencia:

- **Secuencia:** conjunto de cuadros consecutivos en el tiempo. Sintaxis: Código de cabecera e Información general: dimensión horizontal y vertical de cuadros, Relación de aspecto, Número de cuadros por segundo, Velocidad de canal fija o variable, tamaño del buffer, tablas empleadas para la cuantificación
- **Grupo de Cuadros (Group of Pictures –GOP-):** secuencia de cuadros correspondiente a un periodo de la estructura de modos de procesamiento. Su función principal es permitir el acceso aleatorio a partes de la secuencia. Sintaxis: Código de cabecera y código de tiempos correspondiente al primer cuadro del GOP
- **Cuadro:** una componente de luminancia y dos de crominancia. Unidad primaria de codificación. Sintaxis: Código de cabecera e Información particular: referencia temporal del cuadro, modo de codificación, dimensiones máximas de los vectores de movimiento, etc. Si el modo es entrelazado un cuadro tiene dos campos
- **Tira:** conjunto de macrobloques consecutivos (de izquierda a derecha y de arriba abajo) en la imagen. La definición de la longitud y número de tiras en

la imagen es libre, pudiendo llegarse a una tira por imagen (todos sus MBs). La tira es la unidad de resincronización en el caso de errores. Sintaxis: Código de cabecera e información particular: posición del primer MB de la tira en la imagen, parámetros globales de cuantificación de los MBs de la tira

- **Macrobloque:** mínimo conjunto de bloques enteros tanto Y como Cr, Cb que ocupan la misma posición espacial. Unidad de estimación y compensación de movimiento. Sintaxis: información sobre el tipo de predicción empleado, el vector de movimiento (en su caso) o el parámetro que controla el nivel de cuantificación de los coeficientes transformados. 16x16 en Y (en CrCb depende de la estructura de muestreo).
- **Bloque:** grupo de 8x8 muestras (Y, Cr o Cb) que constituyen la unidad mínima de tratamiento. Unidad de codificación DCT. Sintaxis: información codificada de los coeficiente cuantificados

Si la señal es entrelazada se decide a priori si en la codificación de los cuadros de la secuencia se van a tratar los campos que los constituyen ...

- ... por separado (desentrelazándolos): **modo campo (field mode)**. Habrá un MB para cada uno de los dos campos. Solamente es posible si la secuencia de entrada es entrelazada.
- ... juntos (manteniendo el entrelazado): **modo cuadro (frame mode)**. El MB tiene muestras que pertenecen a ambos campos. Es posible para secuencias de entrada entrelazadas y progresivas

Existen 4 modos de predicción (1 más que en MPEG-1)

- **Intracuadro:** no se hace uso de predicción. El MB se codifica directamente mediante la DCT
- **Predictivo:** un MB se codifica tras realizar una predicción obtenida a partir de otro cuadro o campo de la secuencia.
 - Si es anterior: predicción hacia delante (forward): La referencia (campo/cuadro) ya pasada en el tiempo
 - Si es posterior: predicción hacia atrás (backward): La referencia (campo/cuadro) “todavía no ha llegado”
- **Interpolado:** el MB se codifica en modo predictivo pero haciendo uso de predicciones de un cuadro o campo anterior y otro posterior de la secuencia
- **Dual Prime:** es una variación del interpolado, donde los dos MBs a interpolar se corresponden a predicciones forward. Para vídeo entrelazado genera cuadros P con la eficiencia de cuadros B, pero sin su retardo

En MPEG-2 existen tres modos de codificación a nivel de cuadro

- Modo Intracuadro: **cuadros I**
 - No se hacen uso de predicciones.
 - Todos los MBs van con modo de predicción intracuadro
- Modo Predictivo: **cuadros P**
 - Se codifican haciendo uso de predicciones de un cuadro anterior de tipo I o P.
 - Los MBs pueden predecirse en modo intracuadro (MB intra), modo predictivo con el último I o P anterior (MB FW). Adicionalmente, en modo cambio, puede predecirse también en modo Dual Prime.
- Modo Interpolado Bidireccional: **cuadros B**
 - Se codificación haciendo uso de predicciones tomadas del cuadro anterior I o P y/o del cuadro posterior I o P.
 - Los MBs pueden predecirse en modo intracuadro (MB intra), modo predictivo con el último I o P anterior (MB FW), modo predictivo con el siguiente I o P posterior (MB BW), modo interpolado (calculando el promedio de una predicción del cuadro anterior I o P y del cuadro posterior I o P – MB interpolado-)

1.1 Objetivos

Esta práctica persigue los siguientes objetivos:

1. **Analizar visualmente la codificación MPEG-2** mediante el programa **VISUALmpeg** de BeSchuur.
2. **Transcodificar secuencias de vídeo** mediante el programa **Reference** de MainConcept.

El enunciado de esta práctica será complementado con explicaciones sobre esos tres programas, y en particular sobre VISUALmpeg, puesto que se trata de descubrir sus posibilidades, y comparar su valor didáctico con el del codificador MPEG-2-like ha desarrollar en las siguientes prácticas.

1.2 Memoria

La memoria será en formato libre y se realizará en Word. Se aconseja ir haciéndola según se desarrolla la práctica (en el laboratorio está instalado el Word).

Cada grupo deberá redactar una memoria explicativa del trabajo realizado en el laboratorio. Dicha memoria tendrá formato libre pero deberá ser concisa y clara, y

responder a todas las cuestiones planteadas por el enunciado correspondiente, dando razonamientos convenientes para cada respuesta u observación. Además se valorará la descripción de las dudas, problemas y soluciones que hayan surgido a lo largo de la práctica. Se podrán incluir capturas de pantalla.

Uno de los objetivos de las memorias en formato libre es que l@s alumn@s logren diferenciar lo importante de lo que lo es menos, y lo expresen con **explicaciones razonadas y conclusiones propias**, como se debe poder esperar de l@s ingenier@s potenciales que son, en este último curso de la carrera.

Salvo indicación expresa en contra al final del enunciado de la práctica, las memorias tendrán un **máximo de cuatro páginas** DIN A4 (tamaño de letra mínimo de 10 pt) y no contendrán ni hojas en blanco ni portada independiente: l@s autor@s de la memoria y el número *M* y título de la práctica se indicarán al principio de la primera página.

Este número máximo de páginas ha de ser tomado como una seria advertencia contra los feos vicios de arrasar bosques, por un lado, y de desbarrar sobre vida, obra y milagros de cada función del Prolink, por otro. Es mucho más interesante **resaltar los resultados pedidos en el enunciado y comentarlos, dando explicaciones propias** de por qué se ha preferido ciertos parámetros a otros, o de cómo se relacionan los experimentos llevados a cabo con lo aprendido en las clases teóricas.

Tampoco hay que caer en el extremo opuesto, poniendo en la memoria una mera descripción de los pasos seguidos y de los resultados obtenidos, sin ningún tipo de valor añadido. De hecho, en cada memoria habrá obligatoriamente un apartado final de **Conclusiones**, que será una síntesis de lo aprendido en la práctica correspondiente y valdrá hasta dos puntos.

La memoria (fichero .doc autocontenido producido con M\$-Word) se deberá entregar, **por correo electrónico** (JoseM.Martinez@uam.es) y **comprimidos con WinZip**. *Para esta práctica, se podrá exceder el límite de 4 páginas, siempre que la calidad de la extensión lo merezca.*

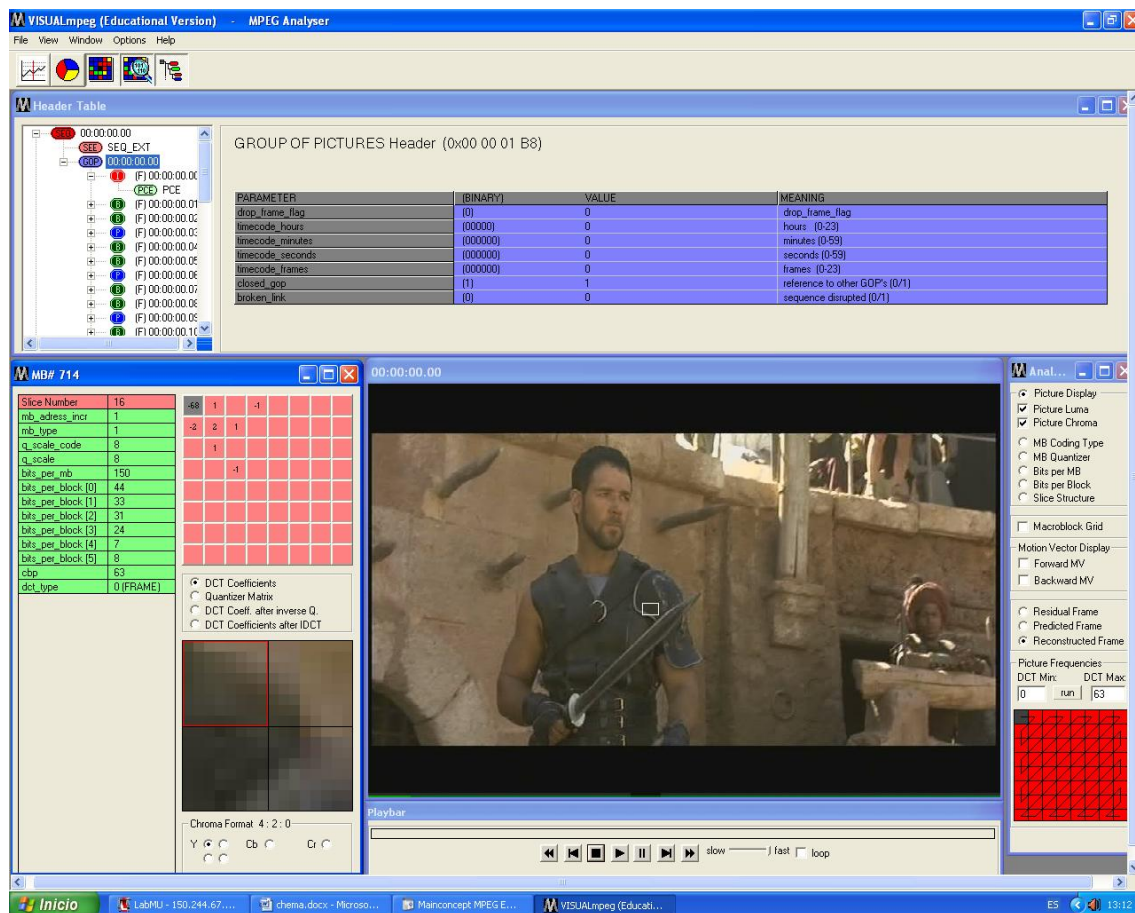
2 Desarrollo de la práctica

2.1 VISUALmpeg (BeSchuur)

En los PCs del LTVD hay instaladas dos versiones de VISUALmpeg: la normal (o Edu) sirve para analizar ficheros de vídeo codificado mediante MPEG-{1,2}, y se dispone de diez licencias para ejecutarla simultáneamente en diez PCs clientes; la avanzada (o AVC), permite analizar vídeo codificado mediante la norma AVC, también conocida como H.264 o MPEG-4 Parte 10, pero se dispone sólo de una licencia de ella, por lo que no se utilizará en las prácticas del LTVD.

Al lanzar el programa VISUALmpegEdu y abrir la secuencia que se debe usar como primer ejemplo, disponible en el directorio (montado mediante samba)

ltvd<x>:\streams\Gladiator@4Mbps.mpv, aparece la ventana principal (siempre maximizada), que tiene el siguiente aspecto:



En la barra de botones, a la derecha de los típicos de crear un fichero nuevo y abrir uno ya existente, están los siguientes, por orden (de izquierda a derecha):

- Cuatro que permiten [des]activar la presencia de los siguientes elementos: árbol de cabeceras, miniaturas, panel de opciones y controles de vídeo.
- Dos que permiten que el contenido de la columna central de la ventana conmute entre información visual (imagen) y numérica (datos de cabeceras MPEG-2), y que el contenido de la columna izquierda conmute entre árbol de cabeceras e información detallada de macrobloque.
- Dos que permiten presentar el cuadro en curso ampliado al máximo en el área de imagen o centrado en ella.
- Dos que sirven para analizar calidad y errores de la señal.

Bajo la barra de botones se encuentran:

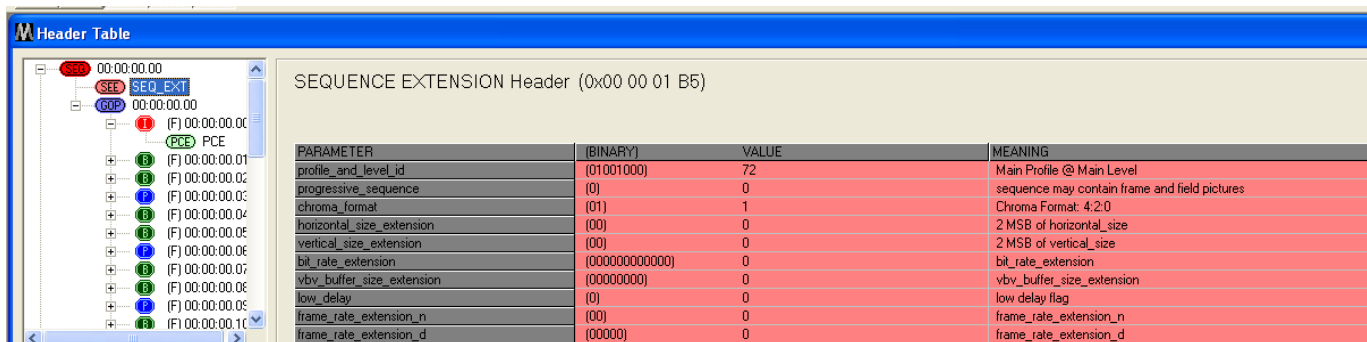
- La columna de la izquierda, que en su recuadro superior muestra la información esencial de la secuencia y el cuadro en curso. Bajo este recuadro está el árbol de cabeceras.
- La columna central, que tiene arriba un área de miniaturas, en el centro el área de imagen, y abajo los controles de vídeo. De éstos, quizá sea preciso aclarar que, además de los botones/símbolos comunes de *play*, *stop* y *pause*, hay otros

que sirven para avanzar/retroceder cuadro a cuadro, o GOP (*Group Of Pictures*) a GOP. En la escala lineal del tiempo, las marcas representan el comienzo de cada GOP y, en la escala numérica, el cuadro se denota en el formato *hh:mm:ss.cc*. Puede ser conveniente avanzar groseramente con esos controles y, una vez alcanzado el punto aproximado deseado, seleccionar el cuadro concreto que interese analizar en el área de miniaturas.

- La columna de la derecha, consistente en un panel de opciones que incluye, en particular, un selector de los coeficientes de la DCT usados en cada momento para la reconstrucción de las imágenes mostradas a la izquierda (obviamente, los cambios introducidos mediante este selector no modifican la secuencia codificada).

2.1.1 Parámetros básicos de la secuencia MPEG-2

Pulsar el botón **H** para cambiar de la información de imagen a la de cabeceras. En la parte izquierda de la pantalla se puede seleccionar la rama del árbol de cabeceras que interese examinar en cada caso, apareciendo los datos en el mismo orden en que se leen del flujo binario MPEG-2. El contenido de la cabecera seleccionada en el árbol se muestra en la parte central, como se ve en la siguiente figura:



PARAMETER	(BINARY)	VALUE	MEANING
profile_and_level_id	(01001000)	72	Main Profile @ Main Level
progressive_sequence	(0)	0	sequence may contain frame and field pictures
chroma_format	(01)	1	Chroma Format: 4:2:0
horizontal_size_extension	(00)	0	2.MSB of horizontal_size
vertical_size_extension	(00)	0	2.MSB of vertical_size
bit_rate_extension	(0000000000000000)	0	bit_rate_extension
vbv_buffer_size_extension	(0000000000)	0	vbv_buffer_size_extension
low_delay	(0)	0	low delay flag
frame_rate_extension_n	(00)	0	frame_rate_extension_n
frame_rate_extension_d	(00000)	0	frame_rate_extension_d

Averiguar los siguientes valores del flujo MPEG-2:

- Tamaño y relación de aspecto de las imágenes, frecuencia de cuadro y velocidad binaria.
- Tamaño de *buffer* mínimo para decodificar.
- Perfil y nivel.
- Formato de muestreo y tipo de secuencia (progresiva/entrelazada).

2.1.2 Estructura del GOP

Recorrer la estructura de las cabeceras y expandir el contenido del primer GOP del fichero (GOP#0). Averiguar tipo y referencia temporal de cada una de sus imágenes en el orden de codificación (en que se transmiten), e.g., **I4 P7 B5**... Hacer lo mismo teniendo en cuenta el orden de presentación. ¿Por qué son diferentes esos dos órdenes?

Recorrer ahora la estructura del segundo GOP (GOP#1) y averiguar tipo y referencia temporal de sus imágenes en ambos órdenes (codificación y presentación). ¿Por qué son distintos esos dos GOPs? ¿De qué tipo son los siguientes GOPs? Visto lo visto, cuál es la longitud del GOP de este flujo?

2.1.3 Parámetros básicos de codificación de imagen

Encontrar la cabecera de extensión de codificación de la primera imagen codificada en el fichero (**I**) y obtener la siguiente información, diciendo qué variables la proporcionan:

- Tipo de imagen (cuadro/campo).
- Orden de los campos.
- Número de intervalos de cuantificación para el coeficiente DC.

2.1.4 Tiras o rebanadas (*slices*) y MBs (MacroBloques)

Pulsar otra vez el botón **H** para volver al «modo visual» y seleccionar, en el panel de opciones (esquina superior derecha), `Slice Structure`. Describir las tiras que componen la imagen.

Seleccionar ahora, en el panel de opciones, `Macroblock Grid` para observar la localización de los MBs en la imagen. ¿Cuántos MBs hay en cada imagen de esta secuencia (i.e., de qué tamaño son)?

2.1.5 Información de brillo y color

Utilizando los controles de vídeo (borde inferior), reproducir el principio de la secuencia. Desactivar, en el panel de opciones, la reproducción de la crominancia desmarcando la casilla `Picture Chroma` y volver a visualizar el principio de la secuencia, de la que, en este caso, se mostrará sólo la luminancia. Repetir la jugada visualizando sólo su crominancia (al desmarcar `Picture Luma`, el programa muestra un valor medio de gris con la crominancia de la imagen). A la vista de los resultados, ¿qué señal (luminancia o crominancia) lleva la mayor parte de la información visual?

2.1.6 Tipos de imágenes y MBs, y vectores de movimiento

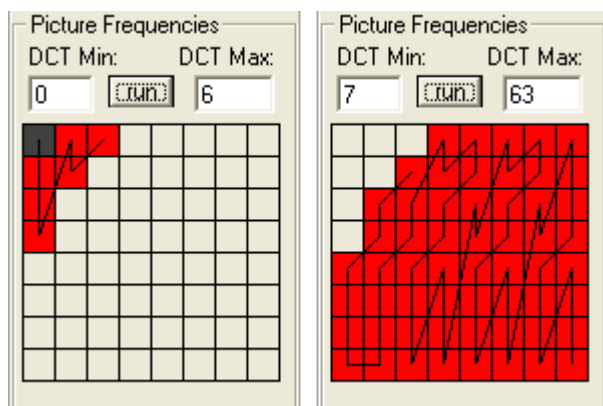
Seleccionar una imagen de uno de los tres tipos (**I**, **P** o **B**) pinchando sobre su miniatura o gracias al árbol de cabeceras, y activar, en el panel de opciones, la visualización de los vectores de movimiento (MVs: *Motion Vectors*) que proceda: `Forward MV` (que se muestran sobre la imagen mediante líneas negras) [y `Backward MV` (líneas blancas)]. ¿Qué tipo de imágenes tienen qué tipo de MVs y por qué?

Activar, en el panel de opciones, la opción `MB Coding Type`. Esto sustituye la crominancia de la imagen, en cada MB, por un color según el tipo de codificación que se usó para ese MB. El código de colores usado aparece entonces en la esquina inferior derecha, aunque el tipo de codificación de cada MB se deduce fácilmente de los MVs usados en la mayoría de los casos. Con la presentación de tipos de MB activa, navegar por los tres tipos de imágenes para averiguar cuáles de los seis posibles tipos de codificación de MB se da en cada tipo de imagen (**I**, **P** y **B**).

2.1.7 Coeficientes de la DCT e información por bloque

En el selector de coeficientes de la DCT, fijar las parejas `DCT Min` y `DCT Max` para los dos casos mostrados en la siguiente figura, y justificar lo observado en cada uno de ellos

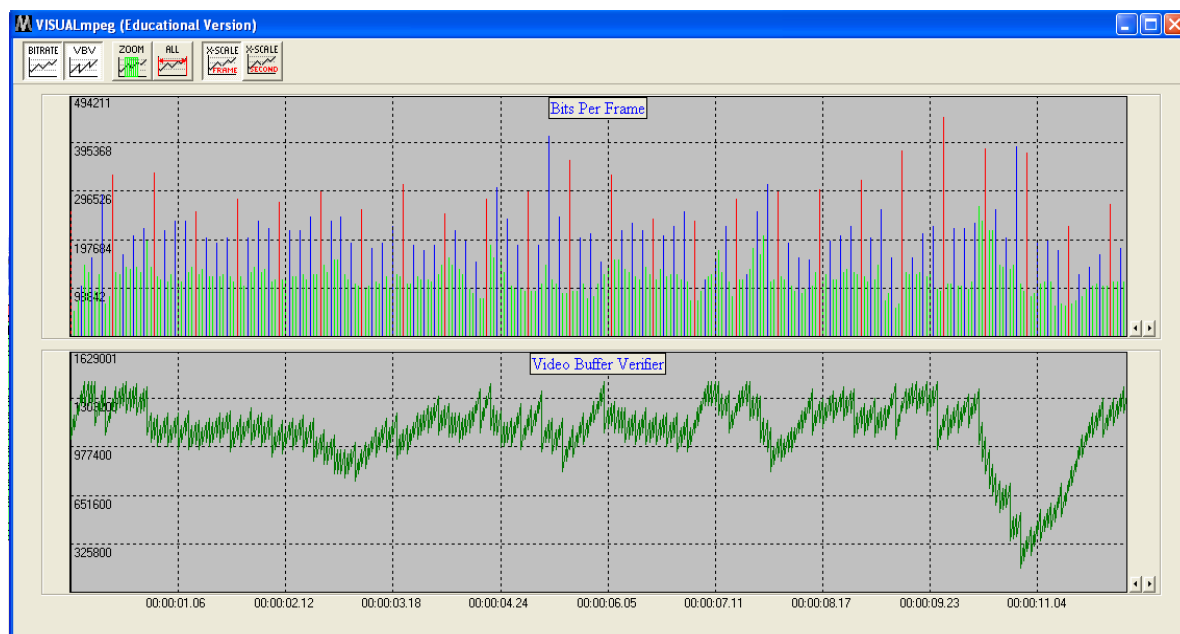
al visualizar la secuencia reconstruida con el correspondiente rango de coeficientes de la DCT.



Restablecer DCT Min y DCT Max a sus valores iniciales y marcar la opción Bits per Block para activar una representación visual de la contribución de cada bloque al presupuesto binario total de la imagen: el brillo de cada bloque es proporcional a la cantidad de bits que se gastan en codificarlo, según el código de colores (escala de grises) mostrado en la esquina inferior derecha. ¿Dónde se gastan más bits?

2.1.8 Análisis de calidad y velocidad binaria

Lanzar la ventana del analizador de calidad:

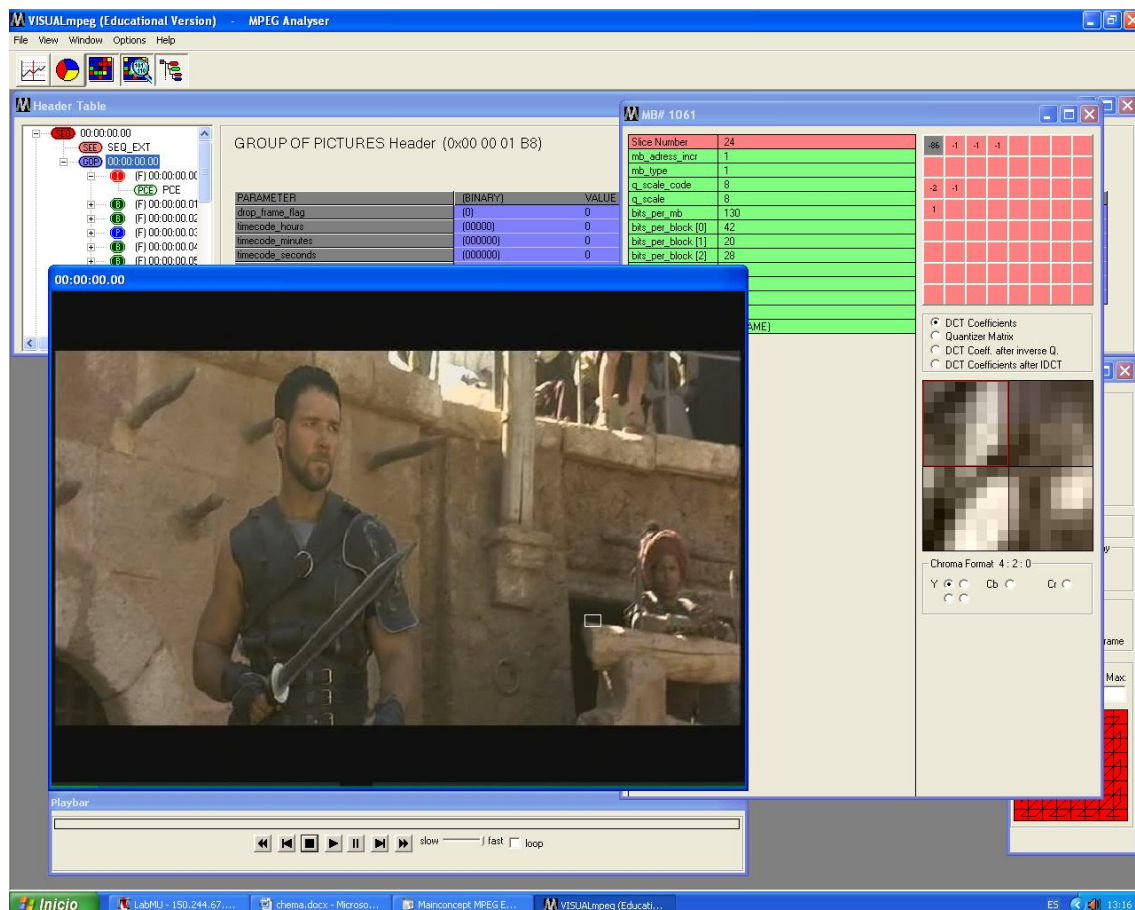


Ordenar los tipos de imágenes (**I**, **P** y **B**) de menor a mayor según la cantidad de información necesaria para codificar cada una, y determinar qué relación hay entre la velocidad binaria [kb/s] en cada instante con el tamaño [kb] mínimo y máximo de las imágenes.

Una vez completado el análisis de calidad, cerrar esta ventana para volver a la principal de VISUALmpeg.

2.1.9 Información detallada de MBs

Visualizar una imagen cualquiera a tamaño real (no maximizada) y recorrerla con el puntero del ratón para obtener en la columna izquierda (donde estaba el árbol de cabeceras) información detallada del MB correspondiente, como se ve en la siguiente figura, y explicar brevemente la información ofrecida por los distintos recuadros.



2.1.10 ¡Vuelta a empezar!

Repetir algunos de los apartados anteriores con los ficheros que hay en directorio (montado mediante samba) `ltvd<x>:\streams\` (u otros grabados en tiempo real mediante el glifo Salidas -> Grabadora) para ver otros ejemplos de secuencias MPEG-2 con distintas estructuras de GOP, diferentes órdenes de recorrido de los coeficientes de la DCT, etc.

2.2 Reference (Main Concept) –opcional-

Esta parte del enunciado es intencionalmente vaga, y la parte correspondiente de la memoria tendrá una estructura completamente libre: se pretende que los distintos grupos del LTVD exploren las posibilidades del codificador profesional de vídeo MainConcept Reference según el tiempo del que dispongan una vez llegados a este punto, y lo bien que hayan entendido los conceptos estudiados en esta práctica y en la(s) anterior(es). Eso sí: se espera, al menos, que dé tiempo a transcodificar varios de los flujos usados anteriormente (TSs o ESs de vídeo) para generar secuencias con distintos tamaños y

calidades. Para ello, se elegirán distintas configuraciones del codificador cambiando el tipo de norma (MPEG-2 o H.264), la estructura del GOP, los parámetros de la DCT o del análisis de movimiento, etc., etc., y en la memoria se detallarán esos cambios y los resultados esperados y obtenidos.

Créditos: Esta práctica está adaptada de las originalmente desarrolladas por Xpertia S.L. como documentación adicional a LabMU y por los profesores de la asignatura Laboratorio de Televisión Digital de la Titulación de Ingeniería de Telecomunicación de la E.T.S.Ing.Telecomunicación de la Universidad Politécnica de Madrid.

José M. Martínez