

HERRAMIENTA PARA LA EXPERIMENTACIÓN DE LA CODIFICACIÓN, COMPRESIÓN Y TRANSMISIÓN DE VÍDEO

Jose Oliver¹, M.P. Malumbres²

¹Becario F.P.I. ²Titular de Universidad de la U.P.V.
Departamento de Informática de Sistemas y Computadores
e-mail: [joliver, mperez}@gap.upv.es](mailto:{joliver, mperez}@gap.upv.es)

RESUMEN: Esta ponencia muestra una herramienta desarrollada para la experimentación en la codificación y transmisión de secuencias de vídeo, temática abordada en la asignatura Redes Avanzadas de la Facultad de Informática de la UPV. Además de realizar una breve descripción de la asignatura, se mostrarán las características de la herramienta así como una serie de ejercicios propuestos que van desde la propia codificación hasta los aspectos más complejos de la compresión de vídeo.

1.- INTRODUCCIÓN.

Redes Avanzadas es una asignatura que se encuentra ubicada en el último semestre de la titulación de Ingeniero en Informática de la Facultad de Informática de la Universidad Politécnica de Valencia. Su carga docente es de 6 créditos (3 de teoría y 3 de laboratorio) y su carácter es optativo. Sus contenidos están diseñados con el fin de ofrecer materias avanzadas y complementarias a las estudiadas en las asignaturas troncales de Redes de Computadores de la titulación. Por ello, es una asignatura que cada año necesita actualizar sus contenidos para mostrar las tecnologías que están saliendo al mercado o que en un futuro inmediato estarán disponibles.

La asignatura se divide en dos grandes bloques temáticos:

- El primero trata sobre aspectos avanzados de las tecnologías LAN y WAN y sus perspectivas de futuro. Aproximadamente requiere el 40% de las sesiones de teoría.
- Y el segundo, aborda el tema de redes multimedia (*Multimedia Networking*), con objeto de estudiar las peculiaridades que deben tener las redes para poder transportar adecuadamente el tráfico multimedia. Para ello, se introducen los medios continuos más habituales en las aplicaciones multimedia (audio y vídeo), la necesidad de utilizar compresores describiendo los estándares más habituales y por último los protocolos y arquitecturas propuestas para poder transportar el tráfico multimedia en Internet con una cierta calidad de servicio.

En cuanto a las sesiones de laboratorio (véase la Tabla 1), están centradas en el segundo bloque temático, ya que el primero tiene un carácter muy teórico y son tecnologías que, en algunos casos, no están disponibles o bien las herramientas de simulación son muy complejas y/o costosas.

P1: API multimedia de Windows: Audio.	4 h.	14.29%
P2: API multimedia de Windows: Vídeo.	4 h.	14.29%
P3: Algoritmos de compresión de audio.	4 h.	14.29%
P4: Algoritmos de compresión de vídeo.	10 h.	35.71%
P5: Transporte de audio y vídeo en una LAN	6 h.	21.43%

Tabla 1: Programa de prácticas de Redes Avanzadas.

Para la realización de las prácticas se requieren el software y hardware siguiente: El entorno de programación Borland Builder Visual C++, una tarjeta de sonido, auriculares, micrófono y cámara de vídeo. También se dispone de un servidor [1] en el que se encuentra el material de prácticas (fuentes incompletos, las soluciones en forma de ejecutable, RFCs, artículos, FAQs, fuentes de compresores, estándares, presentaciones, etc.) de esta asignatura.

Las dos primeras prácticas son introductorias. En ellas se muestra la forma de acceder al API multimedia de Windows, con objeto de capturar y reproducir tanto audio como vídeo. Esto es indispensable para poder llevar a cabo el resto de prácticas, que se basan en ésta. Las dos siguientes versan sobre las técnicas de codificación y compresión de audio y vídeo respectivamente y por último se realiza una práctica en donde se experimentan transmisiones de audio y vídeo en una LAN utilizando un sencillo protocolo.

La herramienta que se presenta aquí se utiliza en la práctica número cuatro², a la que se dedica más de la tercera parte de las sesiones de laboratorio y, junto con la anterior, es la que más posibilidades tiene para la experimentación, ya que las técnicas recogidas en ellas se pueden “oir” y “ver”. Esto suele ser muy estimulante para los alumnos ya que su motivación se ve incrementada con el único propósito de que sus algoritmos sean capaces de realizar codificaciones y/o compresiones que ofrezcan resultados palpables, y no simples resultados numéricos.

2.- CARACTERÍSTICAS DE LA HERRAMIENTA.

Para el análisis de la codificación, compresión y transmisión de vídeo hemos desarrollado una completa plataforma, de manera que el alumno puede experimentar fácilmente sobre varios de los algoritmos de codificación y compresión de vídeo, variando los parámetros de configuración de los mismos, y observando inmediatamente los resultados obtenidos.

Esta herramienta está basada en los protocolos TCP/IP y UDP/IP, siguiendo la clásica arquitectura cliente-servidor y usando el API multimedia que proporciona Windows 9x para la captura de imágenes y procesado de ficheros de intercambio AVI. Los codificadores de vídeo empleados son una modificación de la implementación oficial del estándar MPEG-2 [2] [3], usando las recomendaciones H.261 [4] para obtener un enfoque orientado hacia la videoconferencia.

² Aunque dispone de módulos que permiten la transmisión de vídeo en una LAN, pudiendo utilizarse también en la última práctica de la asignatura.

A la hora de realizar la transmisión, uno de los parámetros que se puede seleccionar es la fuente de vídeo, que podrá ser tanto una vídeo cámara de sobremesa como un fichero AVI. Trabajar con una misma secuencia fuente en distintas pruebas es fundamental para que el alumno pueda obtener resultados que únicamente dependan de los algoritmos y parámetros de compresión seleccionados, y no de la secuencia que registre en ese momento la cámara. De esta forma, para asegurarse de que distintas pruebas se puedan realizar con una misma secuencia fuente usaremos vídeos almacenados en disco.

Otro de los parámetros de transmisión que se puede configurar es el tipo de protocolo que se usará para realizar la transmisión de las imágenes, eligiendo entre el protocolo UDP y el TCP. De esta manera, el alumno puede observar como, cuando se transmite un gran volumen de información, el protocolo UDP pierde más paquetes (y por tanto imágenes) que TCP, ya que éste último realiza más control sobre la transmisión, lo que también se traduce en una mayor sobrecarga computacional.

Sobre el tipo de codificación que el alumno puede elegir y sus parámetros, tenemos:

- *Codificación RGB.* La imagen se transmite sin compresión (24 bpp).
- *Codificación YUV.* Aplicamos una codificación YUV con una reducción de color 4:1:1, y con la posibilidad de transmitir sólo información de luminancia (imágenes en B/N)
- *DCT trucando altas frecuencias.* Se calcula la DCT sobre cada uno de los bloques de 8x8 (de luminancia o crominancia), pudiendo seleccionar el alumno el tamaño del bloque final. Cuanto menor es el tamaño final, más rápida se efectúa la DCT pero de peor calidad es la imagen resultante, ya que perderemos los componentes de más alta frecuencia.
- *MJPEG.* Consiste en aplicar JPEG [5] de forma sucesiva a cada cuadro de la secuencia, de manera que únicamente se consigue compresión espacial. Para este algoritmo podremos seleccionar el tipo de recorrido que se realiza para almacenar los coeficientes cuantizados (ZIG-ZAG o scan alternativo) y el propio parámetro de cuantización. El alumno observa fácilmente que con mayores cuantizaciones se alcanza un mayor nivel de compresión, a costa de una menor calidad de las imágenes.
- *H.261.* En este caso se puede seleccionar el tamaño de la ventana de búsqueda de bloques similares. Cuanto mayor es la ventana de búsqueda, mayor será el coste computacional, pero también la probabilidad de encontrar un bloque de referencia que permita minimizar el error residual aumenta. Otro parámetro que se puede seleccionar es la frecuencia con la que se codifica un cuadro I que, al ser totalmente autocontenidos, permitirán recuperarse frente a errores.

El alumno, para realizar las pruebas, deberá fijar en el servidor los parámetros deseados, de forma que éste tome las imágenes de la fuente especificada, las comprima y las transmita al cliente, que las recogerá y las descomprimirá para finalmente imprimirlas en una ventana gráfica. Después de una sesión de transmisión de vídeo, el alumno puede obtener información sobre cómo ha resultado la compresión y la transmisión, visualizando una pantalla de estadísticas similar a la que se muestra en la figura 1. Las estadísticas que se facilitan las podemos clasificar en:

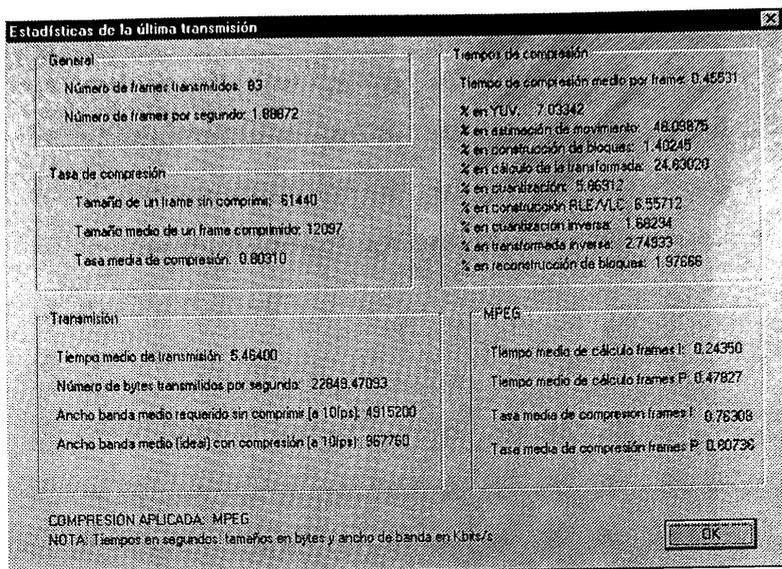


Figura 1: Estadísticas obtenidas

- Estadísticas generales. Incluye el número de cuadros transmitidos y los cuadros por segundo.
- Tasa de compresión. En este apartado se muestra cuantos bytes ocupa una imagen original, y el tamaño medio de una imagen comprimida junto con la tasa media de compresión alcanzada.
- En el apartado de transmisión se indica el tiempo medio necesario para transmitir una imagen, el número de bytes que se han transmitido por segundo y el ancho de banda que se requeriría para obtener una tasa de inyección de 10 cuadros/segundo.
- Tiempos de compresión. En esta sección se muestra el tiempo medio necesario para comprimir un cuadro y se realiza un desglose porcentual de la proporción del tiempo total que se ha empleado en cada fase del algoritmo.
- En el caso de H.261, se indica la tasa media de compresión obtenida en imágenes codificadas como I frente a las P, y el tiempo medio usado en cada caso. El alumno observará como el nivel de compresión obtenido en los cuadros I es inferior pero también el tiempo de cálculo es menor, sobre todo con grandes ventanas de búsqueda.
- Uno de los mejores métodos para determinar la calidad de las imágenes es la observación de las mismas por parte de los alumnos. Sin embargo, a veces resulta interesante tener algún valor numérico que muestre la similitud entre la imagen original frente a la comprimida, para esto se facilita el Error Cuadrático Medio (MSE, *Mean Square Error*), la Relación Señal Ruido (SNR, *Signal to Noise Ratio*) y la Relación Señal Ruido Pico (PSNR, *Peak SNR*).

3.- ALGUNAS EXPERIENCIAS PROPUESTAS A LOS ALUMNOS.

Una vez el alumno ha estudiado los principales algoritmos de compresión de imagen, y se ha familiarizado con la herramienta descrita en el punto anterior, se proponen diversas experiencias prácticas.

Uno de los parámetros que determina de forma más clara la calidad de la imagen y la tasa de compresión obtenida es el parámetro de cuantización. Esta es la primera variable que estudia el alumno, al que se le facilita una típica secuencia AVI de videoconferencia (160x120 píxels, 24 bpp, 10 cuadros/seg. y 8.4 seg.) que codifica y transmite en una red local Ethernet 10BaseT. Se pide al alumno que rellene una tabla semejante a la tabla 2, donde se observa el nivel de compresión obtenido con los distintos niveles de cuantización, el tiempo necesario para realizar la compresión y el ancho de banda requerido usando H.261. También se pide al alumno que contemple la calidad visual de las imágenes (ver figura 2) y observe su relación con el PSNR.

Parámetro de cuantización	Tasa de Compresión	Tiempo de compresión (seg)	PSNR (dB)	Ancho de banda requerido (Kbps)
1	80.209 %	0.24386	42.21	313
8	97.527 %	0.21212	35.13	54
15	98.668 %	0.20796	31.57	30
22	99.051 %	0.20658	29.66	21
30	99.238 %	0.20588	28.51	17

Tabla 2: Estadísticas según la cuantización aplicada

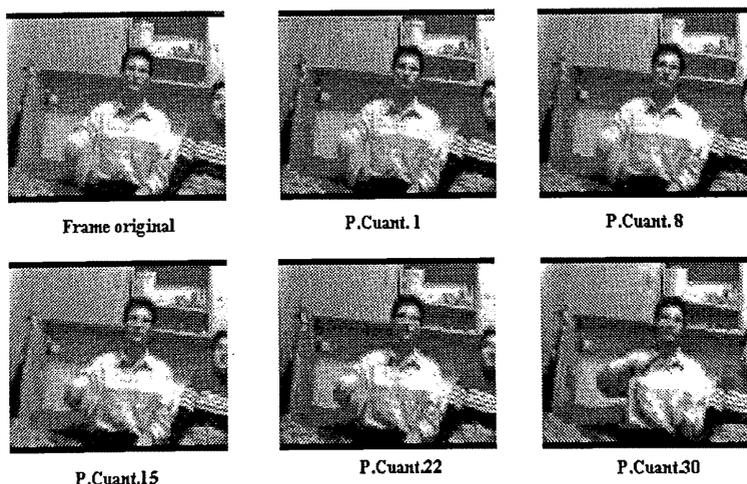


Figura 2: Cuadros obtenidos con distintos parámetros de cuantización

Una segunda experiencia práctica se realiza con el tamaño de los bloques con los que se opera al calcular la DCT, de manera que el alumno, a parte de rellenar una tabla similar a la anterior, debe observar como, con bloques más pequeños, el cálculo de la DCT pasa de ser una de las fases más costosas a tener unos costes computacionales mucho más equilibrados.

Otra experiencia propuesta es variar el tamaño de la ventana de búsqueda de H.261. En las anteriores pruebas no se realizaba una búsqueda de macrobloques en el cuadro de referencia (la ventana era de 0x0). Cuando se amplía el tamaño de la ventana de búsqueda, la fase de estimación de movimiento se convierte en una de las más costosas, lo que el alumno deberá notar en sus experiencias.

4.- ACTIVIDADES ADICIONALES.

El alumno tiene disponible un libro [6] que servirá como soporte teórico y guión práctico para la parte de la asignatura que aborda la codificación y compresión de vídeo. Una vez el alumno ha afianzado los conocimientos teóricos con las experiencias prácticas descritas en el punto anterior, deberá resolver, durante sucesivas sesiones prácticas, una serie de ejercicios propuestos en el libro, que pueden tener carácter obligatorio o voluntario.

Algunos de los ejercicios propuestos requieren la implementación de varios algoritmos de compresión, como es la programación de un codificador RGB/YUV, la implementación de diversas versiones de la DCT optimizando su coste computacional, o la construcción de un codec simplificado JPEG y H.261. Otros ejercicios deberán ser resueltos analíticamente, como es la codificación de un bloque en zig-zag usando las tablas para RLE/VLC o el cálculo del diagrama inverso de Vetterli y Ligtenberg para acelerar el procesamiento de la IDCT.

5.- CONCLUSIONES.

En esta ponencia hemos presentado una herramienta que sirve de apoyo para las prácticas de laboratorio de la asignatura Redes Avanzadas. Una de las características más sobresalientes de ésta es que ofrece a los alumnos un entorno donde se pueden experimentar de forma sencilla y cómoda las distintas técnicas de codificación y compresión de vídeo. Además es una herramienta que permite añadir nuevos algoritmos y técnicas para su evaluación y estudio, por lo que no sólo permite experimentar con los algoritmos que lleva integrados, sino que es posible diseñar y evaluar nuevos algoritmos.

REFERENCIAS.

- [1] <http://www.redes.upv.es/~mperez/rav>. Servidor WWW de la asignatura Redes Avanzadas.
- [2] MPEG Software Simulation Group. *MPEG-2 Codec v1.2*, 1996.
- [3] Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s. ISO/IEC 11172-2: Video, 1993.
- [4] *Video Codec for Audiovisual Services at px64 Kbit/s. CCITT Recommendation H.261*. International Telegraph and Telephone Consultative Committee (CCITT), 1990.
- [5] Digital Compression and Coding of Continuous Tone Images (Part 1: Requirements and Guidelines). ISO/IEC 10918-1, 1992.
- [6] J. Oliver, M.P. Malumbres. *Compresión de imágenes: fundamentos teóricos y aspectos prácticos*. Pendiente de publicación por la U.P.V.