

Informática 4º ESO

Video digital

1 El video

Los inicios del video hay que remontarlos al 28 de diciembre de 1895, fecha en que los hermanos Lumière realizaron la primera proyección pública de un invento que ellos denominaron cinematógrafo.

El cinematógrafo introducía por primera vez la técnica para crear imágenes en movimiento que se ha seguido utilizando desde entonces: la de proyectar imágenes ligeramente distintas en secuencia rápida cómo mínimo 10 imágenes por segundo. La retina humana mantiene cada imagen durante un corto periodo de tiempo, al mismo tiempo que se recibe la siguiente. El cerebro "mezcla" las imágenes superpuestas creando la ilusión de movimiento.

Los sistemas actuales de video, cine y televisión, aplican el mismo principio, aunque utilizando distinta velocidad de proyección de imágenes. Así el cine convencional utiliza 24 imágenes por segundo mientras la televisión en Europa utiliza 50 y en Estados Unidos utiliza casi 60. En la literatura se denomina a este valor *frames per second* o *fps*.

Por lo tanto ya sabemos que el video está formado por imágenes completas que se proyectan en rápida sucesión. Los antiguos sistemas analógicos utilizaban distintas técnicas para almacenar las imágenes, las cuales no vamos a comentar en detalle. Sólo indicar que estas técnicas adolecían de los mismo problemas que aparecían en los sistemas de grabación analógica de sonido. Sucesivas copias de una grabación van degradando la calidad de ésta hasta que se vuelve inservible.

La aparición de los modernos sistemas de procesamiento digital que permiten el mover gran cantidad de datos ha provocado la introducción paulatina de sistemas en los que se permite producir, transmitir y visualizar video en formato digital. Hoy en día estamos inmersos en una competición a gran escala entre los distintos fabricantes por ofrecer video digital de cada vez más calidad.

2 Video digital

Los sistemas de video digital siguen el mismo principio básico de los sistemas de video analógicos convencionales, esto es, generar rápidamente imágenes distintas a fin de que se cree la ilusión de movimiento. La diferencia con los sistemas analógicos es que la señal digital permite la producción, transmisión y reproducción del video de una forma más fiel del original, permite un número ilimitado de copias idénticas de un material y un procesamiento del mismo que era muy difícil o incluso imposible en algunos casos de realizar con los sistemas analógicos.

2.1 Características del video digital

Cualquier sistema de video digital se puede distinguir de otro por una serie de características que definimos a continuación.

2.1.1 Velocidad de refresco

También llamada velocidad de reproducción o frecuencia de refresco, este valor indica cuantas imágenes diferentes se muestran cada segundo. Un mayor valor de este parámetro generalmente ofrece una mejor calidad visual así como menos cansancio para el espectador. Como se comentó al inicio de este tema, la frecuencia de refresco se mide en cuadros por segundo (*frames per second* ó *fps*). El valor mínimo obtenido por experimentación es de 10

aunque usualmente se utilizan valores bastante más altos: 24 para el cine, 25 para la televisión europea, 29,97 para la americana e incluso hay sistemas de video profesional y doméstico que utilizan valores de 50 fps. ofreciendo una calidad muy buena.

2.1.2 Proporción de aspecto

La proporción de aspecto (aspect ratio en inglés) define las proporciones del ancho y del alto de la imagen. Normalmente se expresa como dos números separados por dos puntos, por ejemplo 4:3.

La interpretación de la proporción de aspecto es sencilla. En una imagen con proporción 4:3, el número de pixels de ancho (de cada línea) de la imagen dividido por el número de pixels de alto (número de líneas) da el mismo valor que 4 dividido entre 3.

Por ejemplo, cada imagen de un vídeo grabado en un VideoCD de venta en España tiene una anchura de 352 pixels (cada línea de la imagen tiene 352 pixels) y una altura de 264 pixels (hay 264 líneas). Si dividimos 352 entre 264 obtenemos 1,333.... Lo mismo que si dividimos 4 entre 3.

Cuanto mayor es la proporción entre el ancho y el alto más alargada horizontalmente se muestra la imagen, obteniendo mayor sensación de profundidad porque el ojo captura la imagen de forma periférica.

La proporción 4:3 ha sido la más utilizada en televisión hasta hace unos pocos años, siendo sustituida más tarde por la más utilizada actualmente que es de 16:9, habiendo ahora en el mercado sistemas que ofrecen una proporción de 21:9.

2.1.3 Resolución de la imagen

Cada una de las imágenes que componen una secuencia de video (también llamadas cuadros o fotogramas) están compuestas por pixels, al igual que las imágenes fijas o estáticas que se estudian en el tema de tratamiento de imágenes fijas, no existiendo un sistema "vectorial" de video de amplio uso, sólo sistemas experimentales. Como ya se estudió en dicho tema, una imagen con mayor número de puntos (o puntos más pequeños) ofrece una mayor calidad y definición de la misma. A este parámetro se le denomina resolución de la imagen.

Al contrario que en los sistemas de imágenes fijas, en los sistemas de videos digitales sólo se indica la altura en pixels de la imagen (el número de líneas), calculándose la anchura a partir de éste valor y la proporción de aspecto.

Así, la televisión analógica en España tenía un formato de 625 líneas, la misma aproximadamente que un DVD de video.

Los sistemas de video que se están extendiendo actualmente utilizan 720 ó 1080 líneas, aunque ya se están empezando a comercializar equipos de los llamados 4K que tienen una resolución *horizontal* de 4096 pixels, estando la vertical definida a través de la proporción de aspecto.

2.1.4 Entrelazado

El sistema de entrelazado se desarrolló en los inicios de la televisión para intentar mejorar la calidad de las imágenes en escenas en las que había movimientos rápidos.

En un sistema entrelazado, cada vez que se dibuja una nueva imagen no se hace de forma completa sino sólo se dibujan líneas alternas, esto es, en un dibujo de imagen se refrescan sólo las líneas pares y en el siguiente las impares, repitiéndose este proceso indefinidamente, de forma que la imagen se refresca realmente el doble de veces de lo que indica la frecuencia de refresco, esto es, en el sistema español la imagen se refrescaría 50 veces por segundo, pero como en cada refresco sólo se actualiza la mitad de la imagen, la frecuencia "real" sería de 25 fps.

Los sistemas actuales permiten el uso o no del entrelazado, habiendo sistemas que refrescan la imagen de forma entrelazada en cada dibujo y sistemas que refrescan la imagen completa

en cada uno. A los primeros se le llaman *sistemas entrelazados* (y se denominan i) y a los segundos *sistemas progresivos* (que se denominan p). Así, un sistema 720i tiene una resolución vertical de 720 líneas y un refresco entrelazado mientras que uno 1080p tiene una resolución vertical de 1080 líneas y refresco progresivo.

En cuanto a calidad, los sistemas progresivos dan mayor reposo a la visión pues la imagen se actualiza de forma completa aunque los entrelazados dan mayor realismo a las escenas con movimientos rápidos. El cine siempre utiliza sistemas progresivos.

2.2 Sonido

Un sistema de video digital no está completo sin sonido. Esto es así porque es habitual que las producciones de video estén acompañadas de sonido, ya sea conversación, sonido ambiente o una banda sonora de fondo. El sonido que acompaña a los sistemas de vídeo digitales está también almacenado en formato digital, por supuesto, y sus formatos, características, etc. son los mismos que utilizan los sistemas de sonido digital, que son objeto de estudio en otro tema.

2.3 Subtítulos

Los sistemas de video digital modernos también hacen hueco para la adición de subtítulos en las producciones. Hay muchos formatos de subtítulos, pero en la mayoría de los casos una secuencia de subtítulos consta de una lista de los mismos en la que se indica el momento que debe aparecer cada uno, el tiempo que debe permanecer en pantalla y el texto, junto con su formato, si lo tuviera. Distintos formatos de subtítulos añaden funcionalidades extra como cambiar la posición de los subtítulos en la pantalla (colocarlos en la parte superior, por ejemplo), colores de letra, negrita, cursiva, etc.

2.4 Streams

Una producción de video digital completa consta usualmente de la secuencia de imágenes (lo que se conoce como el vídeo propiamente dicho), una o más bandas sonoras (por ejemplo la misma banda sonora en distintos idiomas, comentarios del director o actores, etc.) y una o más secuencias de subtítulos. A cada una de estas partes se las conoce en la jerga como *streams*, estando todas presentes en los archivos o la transmisión de vídeo.

3 Codificación del video digital

Al igual que con las imágenes fijas o el sonido digital, el video digital es una secuencia de datos numéricos (con un volumen considerable).

Para hacernos una idea del volumen de datos que puede mover un clip de video digital vamos a hacer algunos cálculos.

Supongamos que tenemos un clip de video en formato DVD Video y que dura un minuto. Cada imagen tiene una resolución de 720 x 576 pixels y hay 25 imágenes por segundo. Si tomamos que cada pixel ocupa 3 bytes, tenemos que una sola imagen ocupa $720 \times 576 \times 3 = 1.244.160$ bytes (algo más de un megabyte). Un segundo consta de 25 de estas imágenes, por lo que un segundo ocupa $1.244.160 \times 25 = 31.104.000$ bytes (aproximadamente unos 30 megabytes). Un minuto, que contiene 60 segundos necesitaría $31.104.000 \times 60 = 1.866.240.000$ bytes (aproximadamente 1,7 gigabytes) de almacenamiento.

Como se puede ver, el almacenar o transmitir este volumen de información (sin contar la información del sonido, subtítulos, etc.) requiere un gran gasto en ancho de banda o dispositivos de almacenamiento, por lo que se hace necesario encontrar alguna manera de reducir dicho coste.

La primera y más básica forma consiste en comprimir cada una de las imágenes utilizando cualquiera de las técnicas ya desarrolladas para comprimir las imágenes fijas. Este es el sistema utilizado en algunos sistemas de video de alta calidad, como por ejemplo el utilizado en las cámaras DV (y miniDV). En estos sistemas cada imagen se almacena utilizando el sistema de compresión JPEG (el mismo utilizado para comprimir las imágenes fijas). De esta

forma se reduce en un 20% de media el espacio necesario para almacenar el vídeo.

Los sistemas modernos de codificación de video utilizan técnicas más sofisticadas y que requieren mayor potencia de procesamiento para comprimir el video. Para ello se utilizan las propiedades específicas de las imágenes de video.

En una secuencia de imágenes de vídeo, normalmente la diferencia entre una imagen y la siguiente es mínima. Por ejemplo, si una escena muestra una conversación entre dos personas, el fondo de la imagen varía poco o nada de una escena a otra, centrándose el cambio de una imagen a otra en la cara de las personas que hablan o en sus gestos. Los sistemas de compresión de video modernos analizan las diferencias entre una imagen y la anterior y sólo almacenan los cambios producidos entre las dos, ahorrando de media un 50% o más de espacio, dependiendo de la cantidad de cambios que se produzcan entre imágenes, esto es, se comprimirá más cuando la escena sea más estática y menos cuando en la escena se produzca mucho movimiento.

3.1 Formatos de video actuales

Los formatos más utilizados actualmente para la codificación de video son los siguientes:

- M-JPEG (Motion JPEG). Es el formato utilizado en los sistemas DV o miniDV. Codifican cada imagen como una imagen JPEG. Dan buena calidad pero ocupan mucho espacio.
- MPEG. (Motion Pictures Expert Group). Es un estándar internacional de codificación de video. Hay varias versiones del mismo:
 - MPEG-1. Es el primer formato de video que produjo esta organización. Se utilizó en los LaserDisc y VideoCD. Daba una calidad similar al video VHS.
 - MPEG-2. Es el formato utilizado en los DVD de Video. De una calidad bastante buena. También es el formato utilizado en la TDT que no es de alta resolución (HD).
 - MPEG-4. El formato MPEG-3 no se llegó a lanzar al mercado. El nuevo MPEG-4 da una calidad similar al MPEG-2 pero utilizando mucha menos capacidad de almacenamiento. También es conocido como el formato DivX o Xvid.
- H.264. Es el formato más avanzado actualmente de codificación de video. Es el que se utiliza para los contenidos en alta definición (HD), por ejemplo los discos BluRay utilizan este formato para almacenar el video. Permite una relación excelente entre calidad de imagen y capacidad de almacenamiento. Está sujeto a patentes si se quiere utilizar.
- WMV (Windows Media Video). Formato de video desarrollador por Microsoft. Da una buena relación calidad / almacenamiento pero tiene el inconveniente de que es propietario de Microsoft y trabajar con él es complicado fuera de entornos Windows.
- Theora. Es un formato libre desarrollado por la misma organización que el formato Vorbis de audio. Además de ser libre da una excelente calidad, aunque según los expertos ésta es inferior a la de H.264.
- VP8. Formato originalmente desarrollado por On2 y posteriormente comprado por Google y liberado. Ofrece un nivel de calidad similar al H.264 pero es totalmente libre. Cualquiera puede crear video en este formato sin pagar licencias y crear un reproductor que lo soporte sin tener que pagar tampoco. Desde 2014 es el formato por defecto de los vídeos de YouTube.
- Quicktime. Es un formato propietario de Apple que tiene una buena relación calidad / almacenamiento. El problema con él es que hay que pagar licencia a Apple si se quiere utilizar, lo que ha hecho que no esté muy extendido fuera del ecosistema Apple.
- RealVideo. Este formato se creó originalmente para la distribución de video en streaming desde Internet, esto es, para reproducción de video mientras se descarga, sin necesidad de esperar a que esté descargado el contenido completo para iniciar la reproducción. Ofrece una compresión muy buena aunque su calidad no es alta.

3.2 Formatos de video que llegan

- H.265. Es el sucesor de H.264. Sigue siendo, como aquel, un formato propietario y sujeto a patentes y promete una mejor calidad de video y un tamaño razonable del video en formato 4K.
- VP9. La respuesta de Google a H.265. La calidad es ligeramente inferior a la de H.265 pero requiere menos potencia de cálculo para reproducir un vídeo. Además está libre de patentes para todo aquel que quiera usarlo.

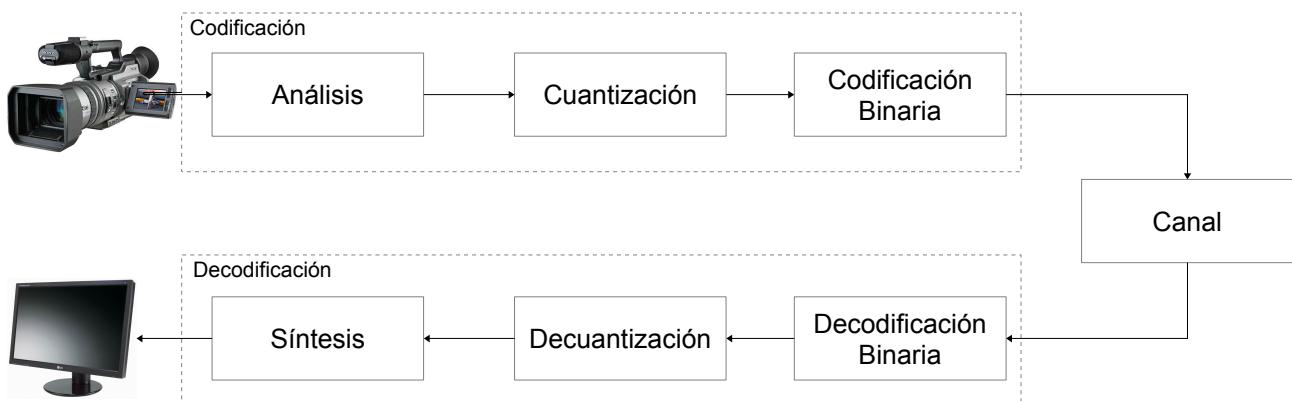
3.3 Funcionamiento de un codec de video

Se llama *codec* de video al software que realiza la *codificación* y *decodificación* de video. Se conoce por *codificación* al proceso que toma una señal de video en bruto y la comprime y convierte en un formato estándar de video. Éste proceso suele requerir una potencia de cálculo más o menos alta, por lo que requiere o bien tiempo en un equipo poco potente o un equipo bastante potente si se quiere realizar en tiempo real, esto es, a la misma velocidad a la que se está generando el video original. El proceso contrario, convertir la imagen comprimida y codificada al formato bruto, listo para visualizar, se llama *decodificación*. Normalmente requiere mucha menos potencia de cálculo para realizarse que el proceso de *codificación*. Ésto es así por diseño ya que la reproducción requiere hacerse en tiempo real y debe de exigir la menos potencia de cálculo posible, a fin de que los equipos reproductores (que son mayoría respecto a los que codifican) sean económicos.

Cuando se quiere trabajar con video digital en un ordenador personal, es necesario disponer de los codecs adecuados a los formatos con los que se piensa trabajar. Por ejemplo, si vamos a editar video procedente de una cámara miniDV y queremos generar un DVD de video necesitamos un decodificador del formato MJPEG utilizado en las cámaras miniDV y un codificador del formato MPEG-2 que es el que se utiliza en los DVD de vídeo.

Existen en Internet paquetes ya preparados, conocidos como "codec packs", que contienen gran cantidad de codecs en un sólo paquete. Al instalar uno de ellos se dispone al momento de acceso para decodificar y codificar gran cantidad de formatos de video, incluyendo tanto los más utilizados como otros que, aun siendo menos utilizados, pueden ser necesarios, sobre todo si trabajamos con grabaciones antiguas que utilizan formatos que ya no son tan populares.

Esquemáticamente un codec funciona de la siguiente forma:



3.4 Características principales del vídeo codificado

Las principales características de un video codificado son las siguientes:

- Tasa de transferencia (bitrate). Indica la capacidad que necesita un canal de transmisión para poder transmitir correctamente el video sin que ocurran cortes. Se mide en kbps o mbps (kilobits por segundo o megabits por segundo). Normalmente, aunque no necesariamente, una mayor tasa de transferencia indica una mayor calidad del video. En cualquier caso una mayor tasa indica una mayor necesidad de canal de transmisión o espacio de almacenamiento para poder trabajar de forma correcta con el

video.

- Tasa de cuadros base. Los modernos sistemas de codificación de video comprimen el mismo calculando las diferencias entre una imagen y la anterior y transmitiendo sólo estas diferencias. Debido a que éste sistema introduce errores, cada cierto tiempo o cada cierto número de imágenes, el sistema inserta una imagen completa, de forma que las siguientes se basen en ésta para buscar las diferencias. A éstas imágenes completas se les llama cuadros base (Base frames o B-frames) y a las que se obtiene por diferencia cuadros incrementales (Incremental frames o I-frames). La cuestión fundamental es ¿cada cuantas imágenes se introducen estos cuadros B? Si se insertan muy frecuentemente la calidad del video es mejor pero se ocupa más ancho de banda o espacio. Si se insertan menos frecuentemente desciende la calidad del video pero por otra parte se ocupa menos ancho de banda y espacio. Los modernos sistemas de video permiten calcular esta información tanto automáticamente como forzando por el editor del video una tasa de cuadros base determinada.
- Resolución base. Las imágenes se almacenan con una resolución dada, aunque los sistemas de decodificación permiten modificar el tamaño de la imagen que se mostrará en la pantalla, tanto aumentando la resolución (upscaling) como disminuyéndola (downscaling). Hay que tener especial cuidado con el upscaling ya que produce distorsiones en las imágenes ya que los nuevos puntos hay que "inventarlos". Existen varias técnicas pero cada una tiene sus ventajas y sus inconvenientes. En la práctica lo mejor es almacenar el video a una mayor resolución y realizar downscaling cuando se necesite ya que se obtiene siempre una imagen de mejor calidad.

4 Contenedores de video

Los codecs que hemos visto anteriormente definen como se codifica el video de forma de que ocupen menos espacio a fin de poder transmitirlos o almacenarlos. Pero con esto no basta. Una película, capítulo de una serie o clip de video contiene habitualmente más información que sólo el vídeo. Habitualmente es necesario, al menos, que se incluya un stream de audio junto con el vídeo (a no ser que se trate de un video mudo). Es, por tanto, necesario el "empaquetar" juntos streams de video y audio. Además, hoy en día se suelen incluir otro tipo de streams, como subtítulos, junto con el video y el audio, por lo que es necesario el definir la manera en que se empaquetan juntos los streams. Esta definición se da mediante un formato llamado contenedor de video.

Un contenedor de video define la forma en que se empaquetan y sincronizan distintos streams de video, audio, subtítulos y de otros tipos de forma de que se puedan reproducir tal y como desea la persona que realiza el empaquetado. Existen distintos tipos de formato de contenedor y vamos a comentar a continuación los formatos más empleados.

- AVI. Son las siglas de **A**udio-**V**ideo **I**nterlaced. Es un formato definido por Microsoft, que lo creó para su uso en la plataforma Windows. Ha sido un formato muy popular durante muchos años aunque tiene carencias que han hecho que no sea adecuado para las necesidades actuales. No puede, por ejemplo, incluir subtítulos y no puede tener un tamaño superior a 2 GB sin necesidad de recurrir a "trucos".
- MKV. Son las siglas de **M**atroska **V**ideo. Es un formato abierto y muy flexible que permite combinar tantos streams como se quiera y de cualquier tipo (incluso streams que no se han inventado aun) y permite archivos de gran tamaño. Es muy utilizado aunque hay reproductores que no lo soportan o que no lo soportan completamente.
- MP4. Es un formato definido en la especificación MPEG-4 y que permite, al igual que MKV, combinar los tipos de stream que se quiera en cualquier número. Permite asimismo la difusión a través de red. Es uno de los formatos más utilizados actualmente en la difusión de videos por Internet. Existe una versión, M4P, que permite la difusión de contenidos protegidos, los cuales no pueden ser reproducidos si no se dispone de las claves o permisos adecuados.
- 3GP. Es un formato creado específicamente para teléfonos móviles. Permite ahorro de espacio y es bastante versátil.

5 Edición de video

La edición de video consiste en manipular la grabación en bruto para obtener un producto terminado. Existen dos tipos de edición de video: lineal y no lineal.

5.1 Edición de video lineal

La edición de video lineal es la forma más básica de edición del mismo. Consiste en editar un clip entero, o una secuencia de éstos, aplicando alguna transformación en el camino.

La edición de video lineal se utiliza principalmente para unir varios clips de video en uno sólo o para aplicar filtros a un video completo para eliminar ruidos, corregir colores o sobreexposiciones, etc. o añadir / quitar pistas de audio, etc.

Ejemplos de aplicaciones populares de edición lineal son Avidemux para Windows o Linux y VirtualDub para Windows.

5.2 Edición de video no lineal

En la edición de video no lineal se colocan los trozos o clips de video sobre una línea temporal, lo que permite crear una secuencia propia, moviendo, cortando, eliminando o copiando trozos de video. Además la mayoría de aplicaciones de edición de video permiten añadir otros efectos muy necesarios en la edición de video como créditos, títulos, transiciones entre clips.

Otra aplicación que contienen los programas de edición no lineales es la de aplicar efectos especiales a determinados clips o escenas, como cambio de color, paso a blanco y negro o sepia, sobreimpresión de textos, etc.

Por último, las aplicaciones de edición de video no lineal también permiten añadir clips de audio para reformar diálogos, colocar música o sonidos de fondo, etc.

La forma de trabajar de todos estos programas es similar. Se crea una librería con los clips de video en bruto que después se van colocando sobre una línea de tiempo (timeline) que constituye la secuencia que tendrá el video ya editado. Sobre esta línea se pueden cortar los clips para eliminar las partes no deseadas o se pueden insertar títulos, imágenes estáticas, etc. Además a cada clip se le puede silenciar el audio, aplicar algún efecto especial, etc.

Asimismo, en el timeline se pueden añadir pistas de sonido que se pueden cortar como los clips, dando un control preciso sobre qué suena en cada momento y cómo se mezclan los sonidos.

Un panel adicional tiene una pequeña vista previa para que la persona que edita puede previsualizar en cada momento cómo quedará el video ya editado de forma que pueda realizar los cambios oportunos para adecuar el resultado a lo que se desea.

Una vez finalizada la edición se pasa a la fase de producción (en inglés rendering) en la que se genera un video en uno de los formatos que soporte la aplicación con todas las ediciones ya realizadas, o sea, el producto final.

Los distintos programas de edición no lineal se diferencian unos de otros en la facilidad de uso, el número de pistas que son capaces de manipular, los formatos de vídeo que soportan, el número de efectos especiales que son capaces de aplicar, la flexibilidad y potencia de la edición de títulos.

Los programas de edición no lineal más conocidos son: Adobe Premiere (Windows), Sony Vegas (Windows), Final Cut (MacOS), Windows Movie Maker (Windows, viene incluido en Windows XP SP2 y posteriores), OpenShot (Linux), Lightworks (Windows / Linux).