

12. Formatos de vídeo digital: formatos de archivo y *codecs*.

La imagen videográfica y televisiva puede ser analógica o digital. Desde hace años la tecnología digital ha ido sustituyendo a la tecnología analógica, primero en la producción, después en la grabación y reproducción [DVDs comerciales] y, por último en la transmisión [ya sea TDT, televisión digital terrestre o por satélite].

El vídeo digital se graba a menudo en cinta, y después se distribuye en discos ópticos, normalmente DVDs. Hay excepciones, como las cámaras de vídeo que graban directamente en DVD, las videocámaras de Digital8 que codifican el vídeo digital en cintas analógicas convencionales, y otras videocámaras de alto precio que graban vídeo digital en discos duros o memoria flash. El vídeo digital se introdujo por primera vez en 1983 con el formato D-1 de Sony, que grababa una señal no comprimida de vídeo componente de definición estándar en forma digital en vez de en las formas analógicas de banda alta que habían sido frecuentes hasta ahora. Debido al coste, el formato D-1 fue usado principalmente por grandes cadenas de televisión. Finalmente sería reemplazado por sistemas más baratos que usaban datos comprimidos, más notablemente por el Betacam Digital de Sony, aún muy usado como un formato de grabación de campo por productores de televisión profesionales.

Los canales para transmitir la imagen televisiva, tanto analógica como digital, pueden ser los siguientes:

- Transmisión hertziana terrestre, recibida a través de una antena convencional.
- Transmisión por satélite, recibida por una antena parabólica.
- Transmisión por cable.

Normas de emisión televisiva

Durante décadas han existido distintas normas para la transmisión de la televisión en color:

Sistema de televisión	País de origen	Líneas	Imágenes por segundo
NTSC [National Television System Color]	Estados Unidos	525	30
Secam [Sequence a Memoire]	Francia	625	25
PAL [Phase Alternate Line]	Alemania	625	25

La imagen videográfica es una señal del mismo tipo emitida por un reproductor de vídeo o un reproductor de DVD que llega al receptor a través de un cable. La diferencia es que, mientras una cinta de vídeo [VHS, Beta o Vídeo 2000] contiene información analógica sujeta a un continuo deterioro, un DVD contiene información digital inalterable.

La toma de imágenes de vídeo digital. Del mismo modo que en una cámara de fotografía digital, la cámara de vídeo dispone de un panel de elementos sensibles, al modo de los CCDs de un escáner, que cuantifica la información recibida en valores numéricos que se corresponden con información cromática en un modelo de color RGB.

Captura de imágenes

Las cámaras de vídeo digital capturan imágenes de dos maneras: entrelazado y escaneo progresivo.

- a. Las cámaras de toma entrelazada graban la imagen en conjuntos de líneas alternas: primero leen las líneas impares y luego las pares, y así sucesivamente. A un conjunto de líneas par o impar se le llama "campo", y un apareamiento consecutivo de dos campos de paridad opuesta se le llama marco o cuadro.
- b. Una cámara digital de toma progresiva graba cada cuadro como distinto; es decir hace una toma completa de la imagen.

Así, un vídeo entrelazado captura el doble de campos por segundo que el vídeo progresivo cuando ambos funcionan al mismo número de imágenes por segundo. Ésta es una de las razones por la que el vídeo tiene un aspecto particular ya que dibuja una imagen diferente 60 veces por segundo, al contrario que el cine, que lo hace a 24 o 25 imágenes progresivas por segundo.

Resolución de imagen

Del mismo modo que en la fotografía digital la resolución está asociada a la calidad de la imagen. Sin embargo en la imagen en movimiento la tolerancia de los usuarios hacia la calidad de la imagen es mayor que en la fotografía porque el movimiento y el sonido compensa la pobre calidad de las imágenes.

La calidad de emisión "broadcast" para la televisión convencional es de 720 x 576 píxeles, en un ratio de pantalla de 3 x 4. Es la más frecuente en las actuales formas de emisión televisiva y en los DVDs comerciales más difundidos pero puede variar ligeramente en cualquiera de las dimensiones. En la actualidad la calidad más alta conseguida de forma experimental es de 7.680 x 4.320 píxeles a 60 imágenes por segundo.

Cada segundo vemos 25 [en PAL o SECAM] o 30 imágenes [en NTSC]. Es necesario que los aparatos reproductores de DVD tengan capacidad para leer la información de una manera rápida [414.720 píxeles en cada una de las imágenes]. Del mismo modo un sistema de transmisión digital como la TDT debe poder transmitir esa enorme cantidad de información.

La única forma de hacer esto viable es que la información esté comprimida mediante algún algoritmo que reduzca de manera efectiva su tamaño de una forma algo parecida a cómo lo hace el sistema de compresión JPEG con las imágenes bidimensionales.

Formatos y compresión

La televisión digital (incluyendo la HDTV de calidad superior) comenzó a extenderse en la mayoría de los países desarrollados a principios de la década de 2000. El video digital también se usa en teléfonos móviles y sistemas de videoconferencia. El video digital también se usa para la distribución en Internet de vídeo, incluyendo el vídeo streaming y el intercambio de películas en P2P entre particulares.

La compresión de vídeo surge de la necesidad de transmitir imágenes a través de un canal que contenga un ancho de banda aceptable. La televisión digital tiene muchas ventajas en comparación al antiguo sistema analógico. Generalmente, el equipo es más barato y la corrección de errores puede compensar la degradación de la señal introducida a lo largo de la trayectoria de la señal. Como la televisión digital trata únicamente con unos y ceros, los circuitos no requieren de la sofisticación que se necesitaba antes para mantener la calidad de la continua variación de voltaje de la señal analógica. Pero el almacenaje y recorrido de la señal digital requiere manejar grandes paquetes de información y puede plantear ciertos problemas. Para un sistema de televisión PAL de 625 líneas y un ratio de aspecto de pantalla 4:3, la imagen activa es: 720 píxeles (Y) + 360 píxeles (Cr) + 360 píxeles (Gb) = 1440 píxeles por línea. Las 576 líneas activas por imagen significan $1440 \text{ píxeles/línea} \times 576 = 829.440 \text{ bytes}$ o 830 bytes de almacenaje. Un segundo de imágenes requiere $830 \times 25 = 20.750 \text{ Kbytes}$ o 21 Mbytes. Ambos sistemas de 625 y 525 líneas requiere aproximadamente de la misma cantidad de almacenaje para un tiempo determinado. Un minuto requiere 1.26 Gbytes y una hora requiere 76 Gbytes. La capacidad de una cinta de vídeo Standard puede grabar únicamente un par de minutos de material de esta densidad.

Compresión

Para reducir la gran cantidad de información digital que genera la televisión se introdujo una técnica que bloqueaba cada fotograma de vídeo y solo transmitía la diferencia entre los fotogramas sucesivos. Con este método de codificación es posible descartar un largo porcentaje de información y seguir produciendo imágenes de TV aceptables. La proporción de información digital seleccionada puede ser comprimida para encajar la capacidad de almacenaje de grabación o para reducir el ancho de banda necesario para transmitir. Si se elimina parte de la información digital, la señal se puede transmitir a través de un canal de menor frecuencia. La relación entre la fuente y la proporción de bits del canal es denominada como factor de compresión. Al final del canal receptor, el expansor o decodificador tratará de restaurar la señal comprimida a sus valores originales. El compresor está diseñado para reconocer y pasar a la parte útil de la señal de entrada conocido como entropía. La parte remanente de la señal entrante es conocida como redundancia. Es redundante porque la información filtrada de salida puede ser predecida por lo que ha sido recibido anteriormente por el decodificador. Si el decodificador no puede reconstruir la información digital retenida, entonces la señal es incompleta y la compresión ha degradado la señal original. Este hecho puede ser o no aceptable cuando se visualiza la imagen recibida. Algunas porciones de una imagen pueden contener elementos irremplazables de un fotograma a otro. Se puede conseguir guardar una considerable cantidad de información digital transmitida si en un cambio de plano se transmite toda la imagen y posteriormente en cada fotograma sucesivo, únicamente se recogen los cambios o diferencias de fotograma a fotograma. La imagen puede ser reconstruida por el decodificador al añadir los elementos cambiantes de la imagen a las partes estáticas o invariables de la misma. El grado de compresión no puede ser muy elevado dado que puede originar pérdida de información.

Compresión interframe. Tan sólo se guardan de forma completa los llamados "keyframes", frames o imágenes de la película que contengan gran información. Del resto de imágenes sólo se guarda aquello que cambia en relación a los keyframes.

La imagen de una persona subiendo una escalera es tomada con una cámara travelling a la misma velocidad que la del viandante. El único movimiento en el cuadro se produciría con otra persona que ande escaleras abajo. La compresión de movimiento es compensada, al producir que la información de movimiento resulte redundante, midiendo sucesivas áreas de imágenes que contienen movimiento y produciendo vectores de movimiento. Estos se aplican al objeto y pronostican la nueva posición reconstruida.

Cualquier error es eliminado al compararse el movimiento reconstruido con el movimiento actual de la imagen original. El codificador envía los vectores de movimiento y las discrepancias de la imagen a través del canal al decodificador, que mueve la imagen previa mediante los vectores y añade las discrepancias para reproducir la siguiente imagen. Así, se ahorra una cantidad de información digital que necesita ser transmitida a través del canal, incluso con movimiento.

Proveyendo únicamente la diferencia entre una imagen y la siguiente significa que en cualquier instante del tiempo una imagen solo puede ser reconstruida por referencia a una imagen completa previa. Si existe un movimiento significativo en el cuadro, habrá poca redundancia y por tanto, poca posible compresión.

Estos métodos de compresión, recurren a los procedimientos generales de compresión de datos, aprovechando además la redundancia espacial de una imagen [áreas uniformes], la correlación entre puntos cercanos y la menor sensibilidad del ojo a los detalles finos de las imágenes fijas [JPEG] y, para imágenes animadas (MPEG), se saca provecho también de la redundancia temporal entre imágenes sucesivas.

Compresión intraframe. Cada frame, del mismo modo que en la compresión JPEG optimiza el almacenamiento de cada imagen al mínimo.

Cuando las imágenes individuales son comprimidas sin referencia a las demás, el eje del tiempo no entra en el proceso de compresión, esto por lo tanto se denomina codificación intraframe o codificación espacial. A medida que la codificación

espacial trata cada imagen independientemente, esta puede emplear ciertas técnicas de compresión desarrolladas para las imágenes fijas. El estándar de compresión ISO JPEG [Joint Photographic Experts Group], está en esta categoría. Donde una sucesión de imágenes codificadas en JPEG también se usan para la televisión, esto es llamado MPEG, "JPEG en movimiento".

Existen muchos tipos de compresión de vídeo para servir vídeo digital sobre Internet, y en DVD. Aunque el vídeo DV no está comprimido más allá de su propio códec mientras se edita, los tamaños de archivo que resultan no son prácticos para la entrega en discos ópticos o sobre Internet. Con Windows Media, MPEG2, MPEG4, Real Media, el más reciente H.264, o el códec de vídeo Sorenson se hace posible su almacenamiento y distribución. Probablemente los formatos más ampliamente usados para entregar vídeo sobre Internet son MPEG4 y Windows Media, mientras MPEG2 se usa casi exclusivamente para DVD, proporcionando una imagen excepcional en el mínimo tamaño pero resultando en un alto nivel de consumo de CPU para descomprimir. Los procedimientos de compresión utilizados en el vídeo digital son los siguientes:

Estándares de vídeo

MPEG 1 es el nombre de un grupo de estándares de codificación de audio y vídeo normalizados por el grupo MPEG (Moving Pictures Experts Group). *MPEG-1 video* se utiliza en el formato Video CD. La calidad de salida con la tasa de compresión usual usada en VCD es similar a la de un cassette vídeo VHS doméstico.

MPEG 2 es la designación para un grupo de estándares de codificación de audio y vídeo publicados como estándar ISO 13818. MPEG-2 es por lo general usado para codificar audio y vídeo para señales de transmisión, que incluyen televisión digital terrestre, por satélite o cable. MPEG-2. Con algunas modificaciones, es también el formato de codificación usado por los discos SVCD's y DVD's comerciales de películas. MPEG 2 es similar a MPEG 1, pero también proporciona soporte para vídeo entrelazado, el formato utilizado por las televisiones.

MPEG 3, diseñado originalmente para la HDTV, la televisión de alta definición, pero fue abandonado posteriormente a favor de MPEG 2.

MPEG 4, introducido a finales de 1998. Los usos principales del estándar MPEG-4 son los flujos de medios audiovisuales, la distribución en CD, la transmisión bidireccional por videófono y emisión de televisión. soporte para la gestión de derechos y variados tipos de interactividad.

Edición

El vídeo digital puede procesarse y editarse en una estación de edición, un dispositivo construido exclusivamente para editar vídeo y audio que puede importar tanto fuentes analógicas como digitales, pero no está pensada para hacer algo diferente a editar vídeos. El vídeo digital también puede editarse en un ordenador personal que tenga hardware y software adecuado. El software y hardware de Avid es utilizado en estaciones de edición; Adobe Premiere es el más usado en ordenadores personales.

Desktop Video, vídeo para el ordenador

El video digital para el mercado de consumo apareció por primera vez con QuickTime, de Apple y con Video for Windows de Microsoft. Los archivos en formatos mov y avi podían verse en un ordenador con capacidad multimedia, en un primer momento de un CD y, más tarde, desde Internet en streaming.

Las primeras herramientas de edición podían digitalizar una fuente de vídeo analógica a un formato digital en baja calidad, 320 x 240 píxeles como máximo. En todo caso, el video digital para el ordenador mejoró rápidamente su calidad, primero con la introducción de estándares de reproducción como MPEG 1 y MPEG 2 (adoptados para el uso en las transmisiones de televisión y el soporte DVD), y después con la introducción del formato de cinta DV permitiendo grabar directamente a datos digitales y simplificando el proceso de edición, permitiendo utilizar completamente sistemas de edición no lineal en ordenadores de sobremesa.

Formas de emisión de la televisión digital

TDT, televisión digital terrestre, que admite varios canales de audio, subtulado y sistemas de pay per view [pago por visión] por programas separados.

Recepción directa vía satélite, con las mismas características que el anterior sistema pero a través de una antena parabólica.

Cable, distribuido en la actualidad con servicios de Internet y telefonía.

Internet. A través de la web, en baja resolución, en formato FLV [tecnología Flash o Silver Light] que permite la recepción en streaming pero no el almacenamiento en el equipo del usuario.

Video on demand, video y televisión bajo demanda

El objetivo de los sistemas de distribución de vídeo y televisión es establecer procedimientos de disfrute conforme a la demanda de los usuarios. De esta forma, los espectadores deciden qué quieren ver en cada momento sin depender de los horarios cerrados de los operadores de televisión. Esto también permite establecer distintos niveles de pago por contenidos; mientras una parte sería de acceso gratuito, otros contenidos se adquirirían en paquetes o como programas

suelos. Esta última opción es la que se usa en los llamados sistemas de “pay per view” para ver acontecimientos deportivos especiales o películas de calidad.

Codecs

Códec es la abreviatura de codificador-decodificador. Describe una especificación desarrollada en software, hardware o una combinación de ambos, capaz de transformar un archivo con un flujo de datos (stream) o una señal. Los códecs pueden codificar el flujo o la señal para la transmisión, el almacenaje o el cifrado y recuperarlo o descifrarlo del mismo modo para la reproducción o la manipulación en un formato más apropiado para estas operaciones. Los códecs son usados a menudo en videoconferencias y emisiones de medios de comunicación. Un endec es un concepto similar para el hardware.

La mayor parte de codecs provoca pérdidas de información para conseguir un tamaño lo más pequeño posible del archivo destino. Hay también códecs sin pérdidas (lossless), pero en la mayor parte de aplicaciones, para un aumento casi imperceptible de la calidad no merece la pena un aumento considerable del tamaño de los datos. Otra cosa es que se trate de información que deba editarse más adelante y conviene mantenerla inalterada.

Muchos archivos multimedia contienen tanto datos de audio como de vídeo, y a menudo alguna referencia que permite la sincronización del audio y el vídeo. Cada uno de estos tres flujos de datos puede ser manejado con programas, procesos, o hardware diferentes; pero para que estos streams sean útiles para almacenarlos o transmitirlos, deben ser encapsulados juntos. Esta función es realizada por un formato de archivo de vídeo (contenedor), como .mpg, .avi, .mov, .mp4, .rm, .ogg, .mkv o .tta. Algunos de estos formatos están limitados a contener streams que se reducen a un pequeño juego de códecs, mientras que otros son usados para objetivos más generales.