

INTRODUCCION A GNU/LINUX MULTIMEDIA

Jesus Nuño – elTxus@gmx.net

1.- El montaje

Aunque hablemos de captura como ripeco, también vamos a hablar de captura como creación.

Una película es una sucesión de imágenes. El cine moderno sí se "inventó" a finales del siglo XIX, pero las inquietudes creativas de la Humanidad en éste campo hace mucho que comenzaron a despuntar. De todas formas, ésto tampoco es una clase de Historia, así que avanzaremos rápidamente hasta el siglo XX.

Sólo un buen montaje puede dar vida a una película. Los diferentes planos son sólo piezas perdidas de película hasta que son ensambladas hábilmente para explicar una historia coherente. Al cortar, quitamos todas las partes flojas de una película, suprimiendo el metraje superfluo: los falsos comienzos, la entradas y salidas innecesarias, etc. A menudo un buen montador puede mejorar la película original concebida por el director.

Desde los años veinte, cuando los teóricos del cine empezaron a darse cuenta de lo que se puede conseguir con el montaje, ésta ha sido la técnica cinematográfica más debatida. Algunos autores han encontrado erróneamente, en el montaje, la clave del buen cine, cuando muchas películas anteriores a 1904 constaban solamente de un único plano y por tanto no dependían del montaje.

El montaje no es necesariamente la técnica cinematográfica más importante, pero contribuye en gran medida a la organización de la película y al efecto que causará en los espectadores. El montaje es la fase de creación de una película en la que se ensamblan y ajustan las imágenes y los sonidos del film. También abarca tareas distintas, desde la organización de los planos en una secuencia, hasta la estructuración narrativa de las secuencias.

Uno de los principios básicos de un buen montaje es mantener el RACCORD, que es el perfecto ajuste de movimientos y detalles que afectan a la continuidad entre distintos planos. Raccordar significa unir dos planos de modo que el paso de uno a otro no dé lugar a una falta de coordinación entre ambos.

Una película está hecha de muchos planos. Cada plano requiere situar la cámara en la mejor posición para observar a los actores, el escenario y la acción en ese momento particular de la narración. Varios factores influyen en la colocación de la cámara y su ángulo. Éste último determina tanto el punto de vista del público como el área cubierta en el plano. Los ángulos de cámara pueden ser de tres tipos:

Ángulo de cámara objetivo: El público ve el suceso a través de los ojos de un observador invisible. Son ángulos impersonales. La gente filmada no parece consciente de la cámara y nunca mira directamente al objetivo. Éste es el ángulo desde el que más suelen ser filmadas las películas.

Ángulo de cámara subjetivo: Filma desde un punto de vista personal. El espectador es colocado dentro de la película, y ve el suceso filmado a través de los ojos del protagonista de la acción. También se implica al espectador cuando alguien mira directamente a la cámara, estableciendo de ésta manera una relación entre espectador y actor.

Ángulos de cámara de punto de vista: Toman la escena desde el punto de vista de un personaje determinado. El espectador no ve la acción a través de los ojos del protagonista, sino a través de su punto de vista, como si estuviera de pie a su lado.

Una película debe presentar un flujo continuo, suave y lógico de imágenes visuales, complementadas con un sonido que describa el proceso filmado de una forma coherente. Es el aspecto continuo de una película, es decir, su continuidad, lo que decide el éxito o el fracaso de una producción. Los elementos visuales y sonoros de una película deben estar integrados para que se complementen mutuamente a la hora de afectar al público.

Una película puede crear su propio tiempo y espacio, para adaptarse a cualquier tipo de narración. El tiempo se puede comprimir o incluso se puede mantener constante tanto tiempo como se desee. El espacio se puede empequeñecer o aumentar o simplemente reconstruir en un escenario que puede existir sólo en la película. Tiempo y espacio pueden ser eliminados, recreados y presentados de una manera que ayude a la comprensión del público.

El montaje es algo más que una técnica para ordenar los trozos de película con una cierta continuidad espacio-temporal. Lumiere era partidario del plano único, en el que los personajes se mueven calculadamente en la misma escena.

Meliés descubrió lo que más tarde se denominaría el truco de parada casi por casualidad, cuando el aparato se bloqueó. Éste truco sería la base del posterior montaje del cine de animación, y se comenzaron a crear escenas de transformación. También descubrirá el fundido y la sobreimpresión realizada de forma múltiple, simple o sobre fondo negro.

Sería Porter quien descubriera la importancia del montaje en el cine, debiendo basarse éste en la sucesión continua de empalmes y no en encuadres aislados (pieza incompleta de la acción, y unidad sobre la que debe construirse el film), sobrepasando la forma teatral de Meliés y dotándolo del principio del montaje. Descubrió así la posibilidad de representar dos o más acciones paralelas, comenzando a utilizar su técnica hacia 1906.

Y sería Griffith quien posteriormente estudiara el montaje de las secuencias y de los planos dentro de las secuencias. Aportó la utilización de un montaje elaborado de múltiples posibilidades, utilizando por primera vez el montaje por atracciones, el montaje de las retroacciones, la vuelta atrás y el flashback.

Poco a poco irían surgiendo teorías a favor y en contra del montaje (o de la manipulación de la realidad). Así, nos encontramos con el grupo del Cine-Ojo, que tenían como meta fundamental no falsear la realidad con artificios procedentes del teatro, y por ello prescinden del guión, los actores, los decorados, la iluminación, el maquillaje, y todo aquello que pueda manipular la realidad tal cual es.

Eisenstein es uno de los cineastas rusos más importantes en la cinematografía mundial, tanto en el ámbito práctico como en el teórico. Para Eisenstein, el montaje ayuda a solucionar la tarea de captación del film. Su eficacia reside en que incluye en el proceso creador las emociones y la inteligencia del espectador, quien es obligado a marchar por el mismo camino creador recorrido por el creador de la imagen. El espectador ve no sólo los elementos representativos de la obra ya terminada, sino que vive también el proceso dinámico de la aparición y composición de la imagen justamente como fue vivido por el autor. Éste es el más alto grado posible para transmitir visualmente las percepciones e intención del autor en toda su plenitud. En sus escritos, distingue cinco tipos de montaje: métrico, rítmico, tonal, sobretonal e intelectual. No profundizaremos más porque ésto tampoco es clase de Teoría de Comunicación

2.- Métodos de montaje visual:

Puede parecer que el montaje presente un dilema al autor. Por un lado, la interrupción física de un plano a otro tal vez parezca tener un efecto perturbador, interrumpiendo el flujo de atención del espectador. Pero no, el montaje es un medio fundamental para construir una película. Éste problema se solucionó planeando la fotografía cinematográfica y la puesta en escena, con el propósito de montar los planos de acuerdo con un sistema concreto.

Montaje Continuo. La finalidad básica del montaje continuo es hacer que la transición de un plano a otro sea suave. En primer lugar, las cualidades gráficas se mantienen continuas de un plano a otro, las figuras están equilibradas y simétricamente colocadas en el cuadro; la tonalidad de la iluminación se mantiene constante; la acción ocupa la parte central del fotograma. En segundo lugar, el ritmo del montaje se establece normalmente dependiendo de la distancia de la cámara del plano. Los planos generales se mantienen en pantalla más tiempo que los planos medios, y los planos medios más tiempo que los primeros planos. Se supone que el espectador necesita más tiempo para abarcar los planos que contienen más detalles. Puesto que el estilo continuo pretende presentar una acción narrativa, es principalmente mediante el manejo del espacio y del tiempo que el montaje fomenta la continuidad narrativa.

Montaje Discontinuo o Sintético. Las películas no narrativas evitan explícitamente el estilo continuo, ya que éste se basa en la presentación lógica de una historia. Las primeras películas de Eisenstein constituyen intentos de construir un film a partir de determinados recursos del montaje. En vez de subordinar los patrones de montaje a la ordenación de la historia, Eisenstein concibe éstas películas como construcciones de montaje. La continuidad espacial clásica se interrumpe con la intercalación de diferentes lugares. Al violar el espacio de ésta manera, la película nos invita a establecer conexiones emocionales y conceptuales.

Montaje Sonoro. Al contrario que el ojo, el oído es sensible a estímulos procedentes de cualquier dirección, siempre que tales estímulos tengan suficiente fuerza. Cuando un nuevo sonido entra en nuestro campo auditivo, no desplaza a los otros, sino que pasa a formar parte del conjunto que estamos percibiendo. No se puede cortar repentinamente la banda de efectos y pasar a la banda de diálogos: el corte llamaría mucho la atención. Lo que se hace en lugar de esto es bajar gradualmente el volumen del sonido, hasta dejarlo en un punto mínimo (estas transiciones deben graduarse bien para que pasen inadvertidas al público).

Una película es concebida en la cámara y ensamblada en la sala de montaje. Una película bien descrita en el guión de rodaje, planificada al minuto, generalmente se encontrará con problemas de montaje menores y fáciles de resolver. El montaje cinematográfico crea siempre un conjunto de relaciones nuevas que no existían previamente como tales en la realidad. Si pretendemos que la realidad cinematográfica sea igual a la realidad verdadera, deberíamos renunciar al montaje y recurrir a la profundidad de campo y al plano secuencia.

En la técnica de la profundidad de campo, en vez de ser la cámara la que pasa de un lugar a otro en planos de diferente duración que analizan la realidad, son los actores los que se desplazan con relación a la cámara, en el interior de un campo captado en un plano fijo.

El cine ha creado, como nunca antes había sucedido en ningún arte, un tiempo y un espacio propios. En el transcurso de una película de dos horas, pueden transcurrir años o sólo unos minutos. El tiempo en el cine se puede comprimir o alargar y no sólo entre la elipsis de las secuencias, sino también en el empalme entre planos dentro de una misma acción.

3.- Glosario y apendices

Antes de continuar, una apreciación:

No conviene pensar únicamente en términos de P2P. Me explico: a éste documento se le pueden dar dos usos, por una parte para crear (hacer un corto con los amigos, grabar las vacaciones en familia, etc., para posteriormente editar y guardar para esas maravillosas reuniones familiares ;)) y por otra parte para hacer "Copias de Seguridad" de nuestros originales en VHS o de nuestro programa favorito. Conviene tener la mente abierta y no comparar siempre lo aquí expuesto con el último "estreno" que acabamos de bajarnos o que acabamos de perpetuar en CD o DVD. Éste texto pretende ayudar a tu creatividad, primeramente dando a conocer las tecnologías existentes y su desarrollo hasta llegar a donde están, y también explicando cómo funcionan y por qué se encuentran en la posición en la que se encuentran (principalmente me refiero, en éste caso, al MPEG-4 y toda la familia de codecs que abarca: DivX, XviD, etc. Más adelante profundizaremos en ello).

Y ahora, vamos a aclarar algunos conceptos que se repetirán a lo largo de toda la charla y que nos vendrá bien tener claro para no perdernos:

ITU -- Unión Internacional de Telecomunicaciones (International Telecommunications Union). Organismo regulador de las Naciones Unidas que cubre todas las formas de comunicación. La ITU establece normas obligatorias y regula el espectro de radiofrecuencia.

ITU-R 601 -- Esta norma define los parámetros de codificación de la televisión digital para estudios. Es el estándar internacional para la digitalización de vídeo en componentes tanto para el sistema de 525 líneas como para el de 625. ITU-R 601 se aplica tanto a las señales diferencia de color (Y, R-Y, B-Y) como al vídeo RGB, y define sistemas de muestreo, valores de la matriz RGB/Y, R-Y, B-Y y características de filtrado. No define sin embargo el interfaz electro-mecánico.

ITU-R 601 normalmente se refiere al vídeo digital por componentes diferencia de color (en lugar de al RGB), para el cual define un muestreo 4:2:2 a 13,5 Mhz con 720 muestras de luminancia por línea activa y digitalización con 8 ó 10 bits. Se acepta una pequeña reserva por debajo del negro en el nivel 16 y por encima del blanco en el nivel 235 - para minimizar distorsiones de ruido y sobremodulaciones. Utilizando una digitalización con 8 bits son posibles aproximadamente 16 millones de colores diferentes: 28 cada uno de los bits para Y (luminancia), Cr y Cb (señales diferencias de color digitalizadas) = $2^8 = 256 = 16.777.216$ combinaciones posibles.

La frecuencia de muestreo de 13,5 Mhz se eligió con objeto de ofrecer una norma de muestreo común políticamente aceptable para los sistemas de 525/60 y 625/50, siendo múltiplo de 2,25 Mhz, la frecuencia común más baja que proporciona un patrón de muestreo estático para ambos.

ITU-R 656 -- Interfaces para las señales de vídeo digital en componentes en los sistemas de televisión de 525 y 625 líneas. Establece la norma internacional para interconectar equipos digitales de televisión que funcionan de acuerdo con la norma 4:2:2 definida en ITU-R 601. Define la señal de borrado, las palabras de sincronismo embebidas, los formatos de multiplexación de vídeo usados por los interfaces serie y paralelo, las características eléctricas del interfaz y los detalles mecánicos de los conectores.

DVE (Digital Video Effects) -- Sistema de Efectos de Vídeo Digital. Los DVE han venido suministrándose como máquinas separadas pero cada vez más se incluyen como parte integrante de los sistemas. La lista de efectos varía pero siempre incluye manipulaciones de la imagen como zoom y posición y puede llegar a rotaciones, perspectiva 3D, paso de página, imagen curvada, difuminados, etc. Además de la lista de efectos, la calidad de la imagen y sus control varían ampliamente.

DVTR -- Magnetoscopio Digital (Digital Video Tape Recorder). El primer DVTR para uso comercial apareció en 1986, funcionando según la norma digital por componentes ITU-R 601 y la norma asociada D1 para DVTRs. Utiliza cassettes de 19 mm. que graban 34, 78 o (usando cinta más delgada) 94 minutos. En la actualidad hay muchos formatos DVTR disponibles. Hay más información acerca de otros VTR's (D1, D2, D3 y D5) en el Glosario de los Apuntes.

FRECUENCIA -- Número de oscilaciones de una señal en un periodo de tiempo determinado (normalmente un segundo). Por ejemplo define las frecuencias de subportadora en los sistemas analógicos, o las frecuencias de reloj en los sistemas digitales.

COMPONENTE (vídeo) -- La interpretación normal de una señal de vídeo por componentes es aquella en la cual la luminancia y la crominancia permanecen como componentes separados, por ejemplo, los componentes analógicos en los VTRs Betacam y MII, o los componentes digitales Y/Cr/Cb en ITU-R 601. RGB es también una señal por componentes. Las señales de vídeo en componentes conservan el máximo ancho de banda de luminancia y crominancia.

LUMINANCIA -- Un componente, el elemento de brillo o de blanco y negro, de una imagen. Se designa como Y, así que la Y de YUV, YIQ, Y/Cr/Cb es la información de luminancia de la señal. En un sistema de televisión en color la señal de luminancia se suele obtener a partir de las señales RGB, que proceden de una cámara o de un telecine, mediante una matriz o una suma que consiste aproximadamente en: $Y = 0,3R + 0,6G + 0,1B$

CROMINANCIA -- La información de color de una señal, relacionada con el tono y la saturación pero no con el brillo o luminancia de la señal. Así, el negro, el gris y el blanco no tienen crominancia, pero cualquier señal coloreada tiene tanto crominancia como luminancia. U y V, Cr y Cb, I y Q, (R-Y) y (B-Y) representan la información de crominancia de una señal. Son los componentes de color Rojo y Azul. El Verde se ha eliminado por ocupar demasiado espacio, además de que puede conseguirse por adición.

LLAVE DE CROMA -- Proceso consistente en superponer una señal de vídeo sobre otra, estando definidas las áreas de la superposición por una gama específica de color, o crominancia, de una de las señales. Para que esto funcione de manera fiable, la crominancia debe tener suficiente resolución, o ancho de banda. Los sistemas de vídeo codificados (compuestos) tienen un ancho de banda de crominancia restringido y por lo tanto no resultan adecuados para una llave de croma de alta calidad.

RELACIÓN DE ASPECTO -- El formato típico del cine es el 16:9. Los televisores convencionales tienen un formato de 4:3. Como se puede observar, las reglas aritméticas dan siempre como resultado 1,33333. Es lo que se llama Relación de Aspecto (de pixels). Se define como la relación de aspecto del área de una imagen asignada a un pixel. La norma de codificación digital ITU-R 601 define pixels de luminancia que no son cuadrados. En el formato 525/60 hay 486 líneas activas cada una con 720 muestras de las cuales sólo 711 son visibles debido al periodo de supresión.

Por lo tanto la relación de aspecto del pixel en una pantalla 4:3 es:

$$486/711 \times 4/3 = 0,911 \text{ (es decir, los pixels son más altos que anchos)}$$

Para el formato 625/50 hay 576 líneas activas cada una con 720 muestras de las cuales 702 son visibles de modo que la relación de aspecto del pixel es:

$$576/702 \times 4/3 = 1,094 \text{ (es decir, los pixels son más anchos que altos)}$$

Hay que tener en cuenta las relaciones de aspecto de los pixels - por ejemplo al realizar un movimiento en un DVE - al rotar un círculo, ya que el círculo debe permanecer circular y no convertirse en elíptico. Otro área en que es importante es la transferencia de imágenes entre plataformas de ordenadores (estándar) y sistemas de televisión. Los ordenadores casi siempre usan pixels cuadrados así que su relación de aspecto debe ajustarse para adaptarse a la de la televisión. Este proceso lleva tiempo y no será perfecto, siendo la calidad del resultado función de la calidad del proceso utilizado.

VÍDEO COMPUESTO -- Señal eléctrica que se compone de la información de luminancia -luminancia- más los impulsos de sincronismo vertical y horizontal. La luminancia y la crominancia se combinan utilizando uno de los sistemas de codificación (NTSC, PAL o SECAM) - para generar el vídeo compuesto. El proceso, que es análogo al de compresión de vídeo, limita el ancho de banda (detalle de la imagen) de los componentes. La crominancia se añade a la luminancia mediante una técnica visualmente aceptable pero resulta difícil, si no imposible, realizar fielmente el proceso contrario (decodificar) para recuperar la luminancia y crominancia originales. Esto puede causar problemas, sobre todo en postproducción.

A pesar de todo, las señales compuestas constituyen un medio sumamente eficiente y económico para transmitir y grabar programas. Las técnicas de procesamiento digital ofrecen la posibilidad de transmitir señales codificadas de pantalla ancha (no de alta definición) a los hogares. Muchos de estos sistemas se están experimentando en todo el mundo, como por ejemplo el PALplus en Europa y el EDTV-II en Japón.

COMPRESION DE VÍDEO -- Proceso que consiste en reducir el ancho de banda o la velocidad de la señal de vídeo. Las normas que se utilizan en la actualidad, PAL, NTSC y SECAM son sistemas de compresión de vídeo analógico. En los sistemas digitales se analizan las imágenes para detectar la redundancia y la repetición y eliminar así los datos innecesarios. Las técnicas se desarrollaron inicialmente para transmisión digital pero han sido adoptadas como medio para manipular el vídeo digital en ordenadores y reducir las necesidades de almacenamiento en los VTRs digitales.

HI-COLOR -- Hace referencia a los píxeles que emplean 16 bits para determinar la información.

Actualmente existen dos principales sistemas de codificación de la señal. Como siempre, uno es el más extendido en los USA y el otro lo es en el resto del mundo:

PAL -- Fase Alternada en cada Línea (Phase Alternating Line). Sistema de codificación para televisión en color ampliamente utilizado en Europa y en todo el mundo, casi siempre con el sistema de 625/50 líneas/campo. Procede del sistema NTSC pero, al invertir la fase de la señal de referencia de color (burst) en líneas alternas (de ahí el nombre) es capaz de corregir las variaciones de tono generadas por errores de fase durante el proceso de transmisión.

El ancho de banda para el sistema PAL-I es de 5,5 Mhz para la luminancia, y 1,3 Mhz para cada señal diferencia de color, U y V.

Hay otros tipos de PAL, como el PAL-M, versión del PAL estándar, pero utilizando una estructura de 525 líneas y 60 campos, y que sólo se utiliza en algunas zonas de Sudamérica (por ejemplo, en Brasil), o el PALplus, sistema de codificación de pantalla ancha compatible con los receptores PAL 4:3 existentes así como con los nuevos receptores PALplus 16:9.

PALplus utiliza la estructura normal de 625/50 líneas/campo así como codificación de color PAL de forma que los receptores 4:3 "normales" pueden presentar la imagen en color con relación de aspecto 16:9 en formato buzón sobre 432 líneas de TV (en lugar de las 576 usuales). Los receptores PALplus 16:9 presentan las 576 líneas completas con los detalles de alta frecuencia adicionales añadidos a las imágenes mediante la señal de realce vertical ("vertical helper") que se transmite por encima y por debajo de la banda central de 432 líneas. Dicha señal aparece como negro en los receptores normales. Este sistema se basa en el procesamiento digital en el codificador y en cada receptor PALplus de pantalla ancha.

Uno de los objetivos del sistema es servir de transición hacia los nuevos formatos de TV y resoluciones. En los estudios se recomienda la señal ITU-R 601 usual con muestreo de 13,5 Mhz de forma que se puedan utilizar la mayoría de los equipos de producción existentes, incluidos VTRs digitales por componentes, sin ninguna modificación - la forma de la imagen ha cambiado pero la naturaleza de la señal no. Naturalmente los mezcladores, generadores de caracteres, DVEs y sistemas gráficos necesitan tener en cuenta el cambio geométrico, pero la electrónica principal - incluidas las memorias de cuadro - no resulta afectada. Las mayores alteraciones se producen en los monitores y en las cámaras - donde las unidades existentes se pueden utilizar con lentes anamórficas o chips CCD de formato ancho (16x9).

NTSC -- Comité de Sistemas de Televisión Nacional. Sistema de televisión en color utilizado en USA, Canadá, México y Japón donde NTSC M es el estándar de transmisión (M define el formato de campo y línea de 525/60, y con frecuencia el sistema se suele denominar simplemente NTSC). Fue definido por el NTSC. El ancho de banda en el sistema NTSC es de 4,2 Mhz para la señal de luminancia y de 1,3 y 0,4 Mhz para los canales de color I y Q.

4:1:1 -- Conjunto de frecuencias de muestreo que guardan la relación 4:1:1, utilizadas para digitalizar la luminancia y las componentes diferencia de color (Y, R-Y, B-Y) de una señal de vídeo. El 4 representa 13.5 Mhz, la frecuencia de muestreo de Y, y cada 1, 3.75 Mhz para R-Y y B-Y. Se suele utilizar como un método económico para formatos de imagen de 525 líneas, ya que muestrea la información de color a la mitad de la frecuencia que en el sistema 4:2:2. Tanto la luminancia como las señales diferencia de color se siguen muestreando en cada línea pero estas últimas tienen la mitad de resolución horizontal que con 4:2:2, aunque se mantiene la resolución vertical de la información de color. Para las imágenes de 525 líneas, esto quiere decir que la resolución horizontal y vertical del color es prácticamente la misma.

4:2:0 -- Sistema de muestreo que se utiliza para digitalizar la luminancia y las componentes diferencia de color (Y, R-Y, B-Y) de una señal de vídeo. El 4 representa la frecuencia de muestreo de Y de 13.5 Mhz mientras que las señales R-Y y B-Y se muestrean a 6.75 Mhz- en una línea sí y otra no (es decir, una línea se muestrea a 4:0:0, sólo la luminancia, y la siguiente a 4:2:2). Esto se suele utilizar como un sistema más económico que el 4:2:2 para los formatos de 625 líneas, de forma que las señales de color tienen una resolución horizontal y vertical prácticamente uniformes.

4:2:2 -- Relación de frecuencias de muestreo que se utiliza para digitalizar la luminancia y las componentes diferencia de color (Y, R-Y, B-Y) de una señal de vídeo. Se suele utilizar como referencia a ITU-R 601. El término 4:2:2 describe que por cada cuatro muestras de Y, hay dos muestras de R-Y y dos de B-Y, lo que proporciona más ancho de banda de crominancia en relación a la luminancia que con el sistema 4:1:1. ITU-R 601, 4:2:2 es la norma para los equipos de estudio digitales y los términos '4:2:2' y '601' suelen considerarse (aunque de manera técnicamente incorrecta) sinónimos. La frecuencia de muestreo de Y es de 13.5 Mhz y la de R-Y y B-Y de 6.75 Mhz cada una, proporcionando un ancho de banda de color máximo de 3.37 Mhz - suficiente para realizar croma-keys de alta calidad.

4:2:2:4 -- Es lo mismo que el 4:2:2 pero incluyendo la señal de llave como cuarto componente, también muestreada a 13.5 Mhz.

4:4:4 -- Una de las relaciones de frecuencias de muestreo que se utiliza para digitalizar la luminancia y las componentes diferencia de color (Y, R-Y, B-Y) o las componentes RGB de una señal de vídeo. En esta relación existe siempre un número igual de muestras de todos los componentes. RGB 4:4:4 se suele utilizar en equipos basados en ordenadores con plataformas estándar. Aunque ofrece la ventaja de proporcionar mejores imágenes formadas a partir de más datos, los sistemas de transmisión y grabación de televisión suelen utilizar muestreos 4:2:2 o 4:2:0 ya que los beneficios son mínimos e incluso pueden desaparecer con la doble conversión entre los sistemas de muestreo.

Y, Cr, Cb -- Señales digitales de luminancia y diferencia de color en una codificación ITU-R 601. La señal de luminancia Y se muestrea a 13,5 Mhz y las dos señales diferencia de color se muestrean a 6,75 Mhz simultáneamente con una de las muestras de luminancia. Cr es la versión digitalizada del componente analógico (R-Y), al igual que Cb es la versión digitalizada de (B-Y).

YUV -- Abreviatura que se suele utilizar, incorrectamente, para describir las señales de luminancia y diferencia de color analógicas en los sistemas de vídeo en componentes. Y es la denominación correcta para luminancia pero U y V son, de hecho, los dos ejes de modulación de la subportadora que se utilizan en el sistema de codificación de color PAL. Se utilizan versiones de las señales diferencia de color B-Y y R-Y multiplicadas por un factor y filtradas, para modular la subportadora PAL en los ejes U y V respectivamente. La confusión surge porque U y V se asocian con las señales diferencia de color aunque claramente no sean lo mismo.

YIQ -- Abreviatura práctica que se suele utilizar, incorrectamente, para describir las señales de luminancia y diferencia de color analógicas en los sistemas de vídeo en componentes. Y es la designación correcta para luminancia pero I y Q son los dos ejes de modulación de la subportadora (I - Inphase y Q - Quadrature) que se utilizan en el sistema de codificación de color NTSC. Se utilizan versiones de las señales diferencia de color R-Y y B-Y multiplicadas por un factor y filtradas, para modular la subportadora NTSC en los ejes I y Q respectivamente. La confusión surge porque I y Q se asocian con las señales diferencia de color aunque claramente no sean lo mismo.

Y, (R-Y), (B-Y) -- Son las señales analógicas de luminancia, Y, y diferencia de color, (R-Y) y (B-Y) del vídeo en componentes. Y contiene información de luminancia únicamente mientras que las dos señales diferencia de color juntas proporcionan la información de color. Estas últimas son la diferencia entre un color y la luminancia: rojo - luminancia y azul - luminancia. Las señales se obtienen de la fuente RGB original (por ejemplo, una cámara o un telecine). Las señales Y, (R-Y) y (B-Y) son fundamentales en televisión. Por ejemplo, en ITU-R 601 son estas señales las que se digitalizan para conseguir el vídeo digital en componentes 4:2:2 y en los sistemas de TV PAL y NTSC se utilizan para generar la señal codificada compuesta final.

DCT -- Algoritmo de compresión. Transformada discreta del coseno (Discrete Cosine Transform). Método muy extendido de compresión de datos de imágenes de vídeo digital que consiste básicamente en analizar bloques de la imagen (normalmente de 8 x 8 pixels) según frecuencias, amplitudes y colores. JPEG se basa en DCT.

CUADRO -- Composición de todas las líneas que forman la imagen. En el sistema PAL, 625 líneas. En el NTSC, 525.

CUADRO CLAVE O KEYFRAME -- Conjunto de parámetros que definen un punto en la transición de un efecto en un sistema DVE. Por ejemplo un keyframe puede definir el tamaño, posición y rotación de una imagen. Cualquier efecto digital debe tener un mínimo de dos keyframes, el principio y el final, aunque movimientos más complejos utilizarán más, incluso hasta 100.

CUADROS B -- Denominados también B-frames. Cuadros predictivos bidireccionales utilizados en la señal MPEG. Se componen evaluando la diferencia entre el cuadro anterior y el posterior en una secuencia de imágenes de televisión. Como contienen únicamente información predictiva no constituyen una imagen completa y por eso tienen la ventaja de ocupar mucho menos que los cuadros I. Sin embargo, para ver la imagen original es necesario decodificar la secuencia completa de cuadros MPEG.

CUADROS I -- Uno de los tres tipos de cuadros que se utilizan en las señales codificadas MPEG. Contienen datos para construir una imagen completa ya que se componen de la información contenida en el propio cuadro (compresión intra-cuadro). La información original se comprime utilizando la técnica DCT.

CUADROS P -- Uno de los tres tipos de cuadros que se utilizan en una señal codificada MPEG-2. Incluyen únicamente información predictiva (no una imagen completa) generada al examinar la diferencia entre el cuadro actual y el anterior. Contienen muchos menos datos que los cuadros I y eso hace posible las bajas velocidades de transmisión que se pueden conseguir con la señal MPEG. Para ver la imagen original que corresponde a un cuadro P es necesario decodificar un GOP (Grupo de imágenes - Group of Pictures) MPEG-2 completo.

CUANTIFICACION O MUESTREO -- Proceso que consiste en muestrear una señal analógica para obtener paquetes de información digital que representan la señal analógica original. El típico cielo nocturno "gris" o la sensación de colores oscuros "apagados" son señal de ruido de cuantificación.

DETECCION DE ERRORES, ENCUBRIMIENTO Y DETENCION -- Ningún medio de grabación digital es perfecto. Tanto los discos como las cintas magnéticas contienen unas pocas áreas de grabación marginales donde la grabación y la reproducción son difíciles o incluso imposibles. Sin embargo, los errores se pueden detectar y remediar por medio de encubrimientos o correcciones. Lo primero intenta ocultar el problema haciendo que no sea tan apreciable mientras que lo último corrige realmente el error de forma que la salida de datos sea perfecta.

Cuando los datos grabados son una imagen, un error se puede disimular sencillamente utilizando datos de líneas de TV o de campos previos o posteriores. No se garantiza que el resultado sea idéntico al original pero el proceso es relativamente simple y rápido. Si la información almacenada procede de una base de datos, un programa de ordenador o es para un procesamiento de imagen más especial, entonces es esencial una exactitud de datos del 100%. Esto puede asegurarse grabando los datos de manera que se pueda detectar cualquier error y entonces calcular el dato correcto exacto a partir de otra información grabada. Entonces se trata de corrección de errores.

Una diferencia entre los sistemas de ordenadores y la TV es que esta última es continua y no puede esperar una corrección posterior. O bien el resultado correcto tiene que estar listo a tiempo o se tiene que realizar alguna otra acción; la función debe continuar, dejando un margen muy estrecho de tiempo para cualquier corrección de errores en TV. En contraste con esto, un ordenador por lo general puede funcionar más lento durante un momento o esperar unos pocos milisegundos.

DITHER (Oscilación) -- En televisión digital, las imágenes analógicas originales se convierten en dígitos: un intervalo continuo de valores de luminancia y crominancia se traducen en un conjunto de dígitos. Mientras que algunos valores analógicos se corresponderán exactamente con dichos dígitos, otros caerán, inevitablemente, en medio. Dado que siempre existirá un cierto grado de ruido en la señal analógica original, puede existir "dither" en los dígitos en el bit menos significativo (Least Significant Bit, LSB) entre los dos valores más cercanos. Esto tiene la ventaja de permitir al sistema digital que refleje los valores analógicos entre LSBs para proporcionar una representación digital muy exacta del mundo analógico. Si la imagen es generada por un ordenador, o como resultado de un procesamiento digital, el "dither" puede no existir, dando lugar a efectos de "contorneado". Utilizando Redondeo Dinámico se puede añadir "dither" a las imágenes para ofrecer un resultado más exacto.

4.- Formatos de video profesionales y domesticos

4.1.- Formatos analogicos

VHS

Si algún formato de vídeo se ha convertido en el estándar de la industria, éste sería el VHS (Video Home System). Gracias a una política correcta de ventas, JVC (Japan Victor Company) logró conseguir la implantación mundial del sistema que inventó en 1976. Las características básicas del VHS son similares a las del formato BETA. Posee cinta de media pulgada, dos cabezales de grabación, eliminación de la banda de guarda, y giro de azimut para los cabezales. Las principales diferencias estriban en el modo de realizar el enhebrado y en la forma de proteger la cromina.

La primera fue la Larga Duración (LP, Long Play), método por el cual la cinta, al desplazarse a mitad de velocidad que la normal, consigue que los registros sobre la cinta duren el doble de tiempo. Por contra, la calidad del sistema es menor. Es importante, para las cabezas, leer sólo la señal sobre sus propias pistas y no sobre las adyacentes. Por lo tanto, en modo SP [al grabar], se incorporan bandas de guarda entre pistas, sin la señal grabada. Sin embargo, las bandas de guarda reducen significativamente el tiempo de reproducción de la cinta. Entonces, se recurre a un truco para evitar la invasión de las pistas adyacentes por parte de los cabezales: éstos son orientados con ángulos diferentes respecto de la cinta (no confundir éste ángulo con el ángulo formado por la pista respecto a la cinta de vídeo). En modo LP, las pistas son empaquetadas sin banda de guarda, y con un solapamiento de 9 micras.

Cuando se presiona el botón Pausa, la cinta se detiene, pero el tambor portacabezales sigue girando, de manera que la misma pista es leída y releída. Bajo condiciones normales, el ángulo de la pista es la suma vectorial de la velocidad de la cinta y el ángulo de inclinación de la cinta sobre el cabezal. Al pararse la cinta, el ángulo se vuelve un poco más acusado, porque la velocidad de la cinta es removida. A pesar de la ayuda provista por el ángulo azimutal, en "pausa", la imagen es degradada.

El audio grabado en las pistas longitudinales de los magnetoscopios es de muy baja calidad debido a que la velocidad lineal de desplazamiento de la cinta es muy baja y la anchura de la pista de sonido muy pequeña. Por ello, se pensó en dotar a los magnetoscopios VHS de un sistema de mejora de audio. Se logra en 1983 de una forma bastante curiosa. A los aparatos que incorporan éste sistema se les denomina HI-FI. El método consiste en poder grabar el sonido en la misma zona en donde se graba el vídeo. Para ello primeramente se modulan en FM los dos canales de sonido estéreo (L y R), con portadoras en 1,4 y 1,8 MHz. Como en éstas frecuencias se cruza con la banda de vídeo, se debe recurrir a algún método para separar ambas señales: los cabezales de audio hi-fi van instalados en el tambor al igual que los de vídeo, pero desplazados 60 grados. Además, se colocan a diferente altura para que puedan leer la misma pista. Durante la grabación, el campo magnético aplicado a los cabezales de audio hi-fi es mayor que el aplicado a los cabezales de vídeo, por lo que se graba en la cinta a una gran profundidad (es por ello que, cuando reusamos una cinta muchas veces, se pueden percibir sonidos de lo que se grabó previamente). Cuando grabamos luego la imagen, ésta sólo consigue llegar a la mitad de profundidad que el sonido, por lo que no lo destruye, pudiéndose reproducir con posterioridad. Aprovechando ésta característica de modulación, las cintas que hay en el mercado bajo la denominación HI-FI, suelen incorporar dos capas magnéticas de menor grosor, pero acondicionadas una para la grabación del audio en profundidad y otra en superficie para el vídeo.

Otra mejora introducida en 1985 fue el sistema HQ, el cual incrementaba el nivel de recorte de blancos en un 20%. Durante la grabación, a la señal se la somete a un proceso de preénfasis, lo que se traduce en una pequeña mejora para las transiciones abruptas de negro a blanco y luego a negro.

Por último, la denominada S-VHS, introducida en 1987, mejora la calidad del sistema ampliamente, por lo que puede ser considerado como un sistema semiprofesional. Entre sus principales diferencias están el cambio de las frecuencias de modulación para permitir un mayor ancho de banda, alcanzando casi 400 líneas de resolución frente a las 270 del VHS estándar. Otra de sus características es que las componentes de luma y croma se introducen en el magnetoscopio por canales diferentes. Lo habitual es disponer de un conector especial denominado S que dispone de cinco patillas. Por enumerar algunas más de sus innovaciones, la cinta contiene partículas magnéticas más finas y con una mayor coercitividad, introduce mejoras en los circuitos de preénfasis y dispone de dos pistas de audio longitudinales además de las dos pistas de audio hi-fi.

8MM

Nace a raíz de una propuesta hecha por parte de Sony, Hitachi, Matsushita, JVC y Philips a principios de 1982, para la estandarización del formato doméstico de vídeo. Su nombre viene del ancho de las cintas de éste formato, que ha sido el último analógico que apareció en el mercado. Entre las características del sistema se encuentra la eliminación de la pista de sincronismos (CTL), y su alta densidad de grabación, ya que usa cintas de metal, con lo que se pueden obtener grabaciones de alta frecuencia usando velocidades menores. Otra novedad que incorpora es que, para hacer el seguimiento de las pistas, los cabezales van montados sobre un material piezoeléctrico, el cual se contrae y expande al aplicarle una tensión eléctrica. Gracias a éste sistema, la cabeza se puede colocar sobre la pista grabada de una manera más precisa que en otros formatos domésticos. En éste sistema se ha establecido que el audio se grabe en la pista de vídeo, modulado en FM, para lo cual se le ha dejado un espacio entre la luminancia y la croma centrado en 1,5 MHz. En éste sistema se ha previsto la posibilidad de grabar sonido en formato digital PCM, para lo cual se utilizaría parte de la pista del vídeo.

Al igual que el VHS evolucionó a un formato semiprofesional, el S-VHS, el formato 8mm condujo al formato Hi8, en el cual las mejoras y la calidad obtenida son similares al S-VHS. Sus ventajas, entre otras, son un menor tamaño de cinta y la incorporación del código de tiempo en una zona de la pista de vídeo.

4.1.2.- Formatos digitales

En el año 1986, SONY presenta el primer magnetoscopio que es capaz de grabar en formato digital. Éste formato recibió el nombre de D1 y la primera máquina fabricada con ésta tecnología, DVR-1000 (antes mencionado en la definición de VTR, y con más info disponible en el Glosario Completo). Éste formato de vídeo es capaz de grabar la señal de vídeo digital en componentes según la norma 601 del CCIR sobre televisión digital en formato 4:2:2. Posee cuatro canales de audio de alta calidad grabados en digital. Gracias a su alta calidad, se pueden conseguir hasta 20 generaciones en multicopia.

Con posterioridad a la salida del formato D1, aparecieron en el mercado otros formatos de grabación digital, entre ellos el D2 y D3, que realizan la grabación en digital pero de la señal en compuesto, muestreando ésta a cuatro veces la frecuencia de la subportadora de color.

Por último, ha aparecido en el mercado un formato que graba en componentes digitales, el BetaCam digital, que usa compresión 2:1 y que muy probablemente se convierta en el estándar de la industria al ser poco voluminoso, poder configurarse como camacopio (un camacopio es un equipo que está formado por una cámara y por un magnetoscopio de vídeo, ambos unidos como una única pieza. Por ejemplo, los equipos de los cámaras de calle), trabajar en componentes y poseer cierta compatibilidad con el formato BetaCam SP.

5.- Presente de Video y Comunicacion: LA REVOLUCIÓN DEL MPEG-4

Miembros de la ITU, de la "International Organization of Standardization" (ISO) y la "International Electrotechnical Commission" (IEC) se pusieron de acuerdo en la base técnica de este sistema, el cual fue llamado por la ITU H.264/AVC y por la ISO/IEC denominado 14496-10/MPEG-4 AVC. El primer objetivo de la norma MPEG-4 fue suceder a las normas MPEG-1 para la compresión y la transferencia de audio/vídeo, y MPEG-2 para la televisión digital.

Algunos de los miembros de la comunidad MPEG-4 están interesados en una comunicación a un rango de velocidad muy bajo, y otros lo están en la TV interactiva. Ambos comparten, a su vez, la creencia de que la codificación basada en objetos es necesaria para romper las limitaciones en las transmisiones de hoy en día. El cambio está siendo empujado por el crecimiento vertiginoso de las tecnologías y sus nuevas aplicaciones: la tecnología de los semiconductores en los procesadores y en las memorias está abriendo una oportunidad para algoritmos que no se podían aplicar anteriormente, por las limitaciones existentes.

¿Qué significa ver? La respuesta de un hombre corriente (y la de Aristóteles) sería "conocer qué hay a través de la visión". En otras palabras, la visión es el proceso de descubrir a partir de imágenes qué hay en el mundo y dónde está. Los estándares de codificación de imágenes disponibles actualmente, y los modelos de datos sobre imágenes subyacentes, principalmente gestionan este proceso proporcionando una representación de las imágenes en términos de secuencias de marcos rectangulares 2D que dan al usuario "una ventana al mundo real". Esta idea se conoce como el paradigma de la televisión. Sin embargo, el proceso de la visión es solamente una parte de la tarea que las personas llevan a cabo, puesto que normalmente el ser humano, además de ver, necesita interactuar con los objetos que componen el mundo que está viendo. Un razonamiento similar puede hacerse respecto al proceso de escuchar y a los correspondientes modelos de representación de audio.

MPEG-4 puede ser utilizado tanto para construir aplicaciones totalmente nuevas como para mejorar las existentes. A diferencia de MPEG-2 (televisión digital), el objetivo de MPEG-4 no es una única aplicación, sino que abre muchas nuevas fronteras. Manipular las escenas audiovisuales, crear, reutilizar, acceder y consumir contenidos audiovisuales será mucho más sencillo. Ahora pueden desarrollarse aplicaciones nuevas y mucho más ricas, como por ejemplo vigilancia remota, comunicaciones personales, juegos, multimedia en soportes móviles, etc.; incluso es posible combinar distintos servicios tradicionalmente separados: "emisiones", "interacción directa" y "comunicaciones". En definitiva, MPEG-4 permite la convergencia de distintas áreas, entendida en términos de proliferación de elementos multimedia en todo tipo de servicios y (acceso a) redes. Cuando nos referimos al uso del MPEG-4 con medios de televisión digital, hablamos de transmisiones DVB (Digital Video Broadcast).

El MPEG-4 no es un patrón que dependa de un solo algoritmo. Ningún algoritmo existente puede considerar el rango de funcionalidad de las aplicaciones actuales, pero es conveniente señalar que no es indispensable utilizar un solo algoritmo para incrementar la funcionalidad de las aplicaciones ya que esto redundaría en una falta de efectividad y eficiencia del sistema por lo cual es conveniente conmutar varios algoritmos ó incluso utilizar algunos nuevos algoritmos. Lo anterior permitirá futuros avances en las técnicas de codificación que serán incluidas en los patrones, de tal forma que el MPEG-4 establece un juego de herramientas extensibles de codificación que puedan ser combinadas de varias formas y los mismos puedan ser vendidos para utilizarlos en aplicaciones específicas.

En este sentido el MPEG-4 será el API (Programa de Aplicación para Interfases) para las representaciones codificadas de datos audiovisuales. Quien unirá los objetos codificados en el MPEG-4 será el MSDL (M Syntax Description Lenguaje), y que abarca varios de los siguientes componentes:

- 1.- Definición de la interfase para la codificación de objetos.
- 2.- Un mecanismo de codificación de objetos para construir algoritmos codificados.
- 3.- un mecanismo que permita la descarga de nuevos códigos objetos.
- 4.- Conjunto de reglas para organizar esta sintaxis.

Ésta estructura implica una transmisión multifase de los datos del MPEG-4, y la misma comienza con un intercambio entre un usuario y una base de datos ó entre dos usuarios: en ésta fase de configuración el codificador y el decodificador determinan los objetos a codificar que serán usados, y su configuración, si no los poseen existirá una fase de aprendizaje durante la cual los objetos codificados serán entregados.

La norma MPEG-4 ofrece una solución radicalmente diferente para la codificación de la señal de vídeo. Las escenas audiovisuales están compuestas de varios objetos jerarquizados. De este modo, en todos los niveles de dicha jerarquía, encontramos: imágenes fijas (background), objetos de vídeo (objetos en movimiento sin background) y objetos de audio (voz asociada al objeto en movimiento). Una escena audiovisual, codificada en MPEG-4, se describe como un conjunto de elementos individualizados. Esta contiene componentes media simples y reagrupados por tipos. Esta construcción permite la fijación de escenas complejas, autorizando al usuario a manipular sólo una parte de los objetos. Para llegar a este objetivo, la norma ofrece soluciones y algoritmos, reagrupando la mayoría de sus funcionalidades, como la compresión de las imágenes y de los vídeos, la compresión de las texturas mapping para las imágenes en 2D y 3D, la compresión de las imágenes en 2D implícitas, la compresión de los campos de animación geométrica de éstas, el acceso aleatorio de todos los tipos de versiones originales (VO), la extensión de las funcionalidades de manipulación de las imágenes y secuencias de vídeo, la codificación de los vídeos e imágenes basados en el contenido, el redimensionamiento espacial, temporal y cualitativo y la robustez y resistencia a las tasas de error en cualquier ambiente.

Una escena audiovisual de MPEG-4 es descrita y compuesta en objetos individuales. La escena puede contener objetos compuestos de diversos medias, los cuales agrupan un número de objetos media primitivos. Por hacer una analogía, los objetos primitivos se corresponderían con las hojas que componen un árbol, mientras que los objetos media compuestos serían las ramas del mismo árbol. Por ejemplo, el objeto visual de una persona hablando y la correspondiente voz se unen para formar un nuevo objeto compuesto, conteniendo ambos componentes, sonido e imagen natural de la persona hablando. MPEG-4 ofrece una manera estandarizada de describir una escena, permitiendo, por ejemplo:

- Colocar los objetos media en cualquier lugar del sistema de coordenadas.
- Aplicar transformaciones para cambiar la geometría o la apariencia acústica del objeto.
- Agrupar objetos primitivos para formar objetos compuestos.
- Aplicar datos adicionales para modificar los atributos del objeto (sonido, textura, animación, etc.).
- Cambiar, interactivamente, el punto de vista y audición del usuario en cualquier lugar de la escena.

Ventajas que ésta conlleva:

- La codificación basada en contenidos de imágenes y vídeo permite la decodificación desordenada: la reconstrucción de objetos de vídeo se realiza de forma arbitraria.
- El acceso aleatorio al contenido de una secuencia de vídeo permite funciones tales como pausa, reproducción rápida hacia delante y rebobinado rápido de los objetos de vídeo almacenados.
- La manipulación del contenido extendida permite, en las secuencias de vídeo, funciones como la inserción de texto sintético o natural, imagen y/o vídeo. Un ejemplo es el mapeado de texto sobre un objeto, y que se mueve de la misma manera y acorde a dicho objeto.

Uno de los elementos más importantes es el modelado de mallas 2D. Una malla 2D es una división de una superficie plana en patrones poligonales. Los vértices de éstos polígonos son conocidos como Puntos Nodales, o nodos de la malla. MPEG-4 considera únicamente mallas triangulares, donde los patrones son, obviamente, triángulos. Una malla dinámica 2D es la información de la geometría y movimiento de los nodos: deformando la malla por el movimiento de los puntos nodales, podemos simular movimiento del objeto. La textura, dentro de cada uno de éstos patrones de referencia, se transforma a la del frame en cuestión mediante un mapeado paramétrico: la forma lineal de las mallas triangulares implica una baja complejidad computacional para éste mapeado. Las mallas triangulares han sido muy usadas en geometría de objetos 3D, modelado y renderizado de objetos y gráficos. El modelado de mallas 2D puede ser considerado como una proyección de la de 3D a una imagen plana.

Y éste modelado de mallas triangulares es tan crucial porque puede ser utilizado para la compresión si se elige transmitir los mapas de textura sólo en los keyframes o cuadros clave (que nosotros estableceremos previamente a la hora de capturar el vídeo) y los mapas de textura animadas (sin predicción de error) para los frames intermedios. Es lo que también se conoce como "auto-transfiguración de los keyframes" seleccionados usando la información de las mallas 2D.

Así pues, nos explicamos el cómo y el por qué códecs como DivX o XviD son los más usados: ofrecen una relativa rapidez de codificación en cuanto a consumo de CPU se refiere, la decodificación se optimiza en todo momento en función de la carga (aunque con posibilidades de indicar la prioridad de uso de procesador manualmente), y permite ratios de compresión excelentes para las capacidades tecnológicas actuales en gran parte del mundo.

6.- Futuro de Video y Comunicación: ¿MPEG-7?

Avalado por éxitos anteriores, MPEG (Moving Pictures Experts Group, y en el que trabajan fabricantes de la importancia de Canon, Sony, NTT, General Electric o France Télécom) está trabajando en un nuevo estándar de representación audiovisual destinado a solucionar el problema de describir información audiovisual para permitir búsqueda eficiente, procesado y filtrado de distintos tipos de materiales multimedia de interés para el usuario: MPEG-7, llamado oficialmente "Interfaz de Descripción de Contenidos Multimedia".

MPEG-7, como otros miembros de la familia MPEG, es una representación estándar de información audiovisual satisfaciendo requisitos particulares. El standard MPEG-7 se construye sobre otras representaciones estándar tales como PCM, MPEG-1, MPEG-2 y MPEG-4. Una de las funciones del estándar es proporcionar referencias para porciones ideales de ellos. Por ejemplo, tal vez una forma de descriptor utilizada en MPEG-4 es útil en un contexto MPEG-7 también, y lo mismo se puede aplicar a los campos de vectores móviles usados en MPEG-1 y 2.

Los descriptores de MPEG-7, sin embargo, no dependen de la manera en la que el contenido descrito es codificado o almacenado. Es posible pegar una descripción MPEG-7 a una película analógica o a una imagen impresa en papel. Aunque la descripción MPEG-7 no depende de la representación (codificada) del material, el estándar de alguna manera se construye sobre MPEG-4, el cual proporciona el medio para codificar material audiovisual como un objeto teniendo ciertas relaciones en el tiempo (sincronización) y espacio (en la pantalla para vídeo, o en una habitación de audio). Usando una codificación MPEG-4, será posible pegar descripciones a elementos (objetos) dentro de la escena, tales como objetos audio y visuales. MPEG-7 permitirá diferentes granulosos en sus descripciones, ofreciendo la posibilidad de tener diferentes niveles de discriminación.

Los datos MPEG-7 han de ser físicamente localizados con el material AV (Audio/Vídeo) asociado, en el mismo chorro de datos o sobre el mismo sistema de almacenamiento, pero las descripciones también podrían estar en cualquier otra parte del globo. Cuando el contenido y su descripción no están co-localizados, es útil un mecanismo que enlace sus descripciones MPEG-7; estos enlaces no deberían trabajar en ambas direcciones.

El énfasis de MPEG-7 será la aportación de soluciones nuevas para descripciones de contenido audiovisual. Por tanto, el direccionamiento de sólo texto no estará entre los objetivos de MPEG-7. Sin embargo, el contenido audio visual debe incluir o referirse a texto amén de su información audiovisual. MPEG-7, por lo tanto, considerará soluciones ya existentes y desarrolladas por otras organizaciones de estandarización para documentos de sólo texto y los mantendrá como deba. Aparte de los descriptores mismos, la estructura de la base de datos juega un papel crucial en la final ejecución de recuperación. Para permitir el deseado juicio rápido acerca de si el material es de interés, la información indexada tendrá que ser estructurada, por ejemplo, de una forma jerárquica o asociativa.

Hay muchas aplicaciones y dominios de aplicaciones los cuales se beneficiarán del estándar MPEG-7. Algunos pocos ejemplos son:

- Librerías digitales (catálogo de imagen, diccionario musical...)
- Servicios de guía telefónica multimedia (p.ej. páginas amarillas)
- Selección de medios de radiodifusión (canal de radio, canal de televisión...)
- Edición multimedia (servicio personalizado de noticias electrónicas)

La forma de datos MPEG-7 será utilizada para responder a las peticiones de usuario fuera del alcance del estándar. En principio, cualquier tipo de material AV puede ser almacenado por medio de cualquier tipo de material cuestionado. Esto significa, por ejemplo, que material de vídeo puede ser solicitado usando vídeo, música, voz, etc. Es misión de la máquina de búsqueda el combinar el dato en cuestión y la descripción AV MPEG-7. Algunos pocos ejemplos son:

Música: Tocar unas pocas notas en un teclado y obtener una lista de piezas musicales conteniendo (o casi) la melodía requerida.

Gráficos: Dibujar unas cuantas líneas en una pantalla y obtener como respuesta un grupo de imágenes que contengan gráficos similares, logos...

Imagen: Definir objetos, incluyendo manchas coloridas o texturas y obtener como respuesta ejemplos entre los cuales seleccionemos los objetos de interés para componer la imagen.

Movimiento: Dado un grupo de objetos, describir movimientos y relaciones entre objetos y obtener como respuesta una lista de animaciones completando las relaciones descritas temporales y espaciales.

Escenario: Dado un contenido, describir acciones y obtener una lista de escenarios donde hayan sucedido acciones similares.

Voz: Utilizar un extracto de la voz de Pavarotti, y obtener una lista de grabaciones de Pavarotti, videoclips donde aparezca éste cantando o videoclips donde esté presentado.

7.- Bibliografía y Enlaces Recomendados

<http://perso.wanadoo.fr/debian/dists/unstable/main/source/>

Multitud de paquetes de contenido multimedia. También puede añadirse la siguiente línea a /etc/apt/sources.list: **deb http://marillat.free.fr/ testing main**

<http://chicosmalos.virtualave.net/apuntes.htm>

Apuntes, enlaces, noticias, etc., relacionadas con el mundo de la Comunicación Audiovisual.

<http://www.escomposlinux.org/rvm/index.php?doc=tele.html>

La página de RVM, con multitud de textos y paquetes: ver TV, aplicaciones de teletexto, páginas del MAN de xawtv traducidas al español, etc.

<http://bytesex.org/snapshot/>

Página de donde obtener todos los programas necesarios para compilar, instalar y ver/grabar TV bajo nuestro S.O. favorito: bttv, btaudio, v4l2, etc.

<http://www.escomposlinux.org/jcantero/linux/debian/debian-TV.php>

Un documento un poco antiguo (basado en la versión 2.2 "potato" de Debian) pero muy detallado en cuanto a la instalación de los módulos compilándolos dentro del kernel.

<http://will.freehosting.net/bttv-HOWTO.html>

El mini-Como del bttv. Documento bastante interesante si quieres conocer un poco más a fondo como funciona.

– <http://www.linux-hogyan.hu/LDP/HOWTO/VCR-HOWTO-1.html>

Y el VCR-Como. Convierte tu PC en un Video Cassete Recorder (o sea, captura tus videos favoritos ;))