



3. UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO

INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA

“INTRODUCCIÓN A LA TECNOLOGÍA DE VIDEO MODERNA”

M O N O G R A F Í A

**PRESENTADA QUE PARA OBTENER EL TÍTULO DE:
INGENIERO EN ELECTRÓNICA Y TELECOMUNICACIONES**

**PRESENTA:
JOSUE NEFTALI ANGELES CRUZ**

**ASESORÓ:
ING. MARIANO ARUMIR RIVAS**

PACHUCA HIDALGO, FEBRERO DE 2006.

AGRADECIMIENTOS.

A todas aquellas personas que forman parte importante de mi vida, les agradezco, su confianza, enseñanzas, consejos y apoyo. Por que con ello son parte importante de mi superación personal.

Gracias a mis padres, mi hermana, mi novia, mis familiares, mis amigos y maestros.

Pero sobre todo gracias a Dios por permitirme vivir esta vida, este tiempo y este momento.

CONTENIDO

INTRODUCCIÓN	9
CAPITULO 1. FUNDAMENTOS DEL VIDEO ANÁLOGO	11
1.1 VIDEO ANÁLOGO	11
1.2 PROCESO DE EXPLORACIÓN DE LA IMAGEN.	12
1.2.1 Barrido entrelazado.	16
1.3 FRECUENCIAS DE EXPLORACIÓN HORIZONTAL Y VERTICAL.....	18
1.3.1 Tiempo de línea horizontal.	18
1.4 LAS SEÑALES DE COLOR.	18
1.5 SISTEMAS DE TELEVISIÓN A COLOR.....	22
1.5.1. Sistemas de exploración y sistemas de color.....	23
1.5.2 Compatibilidad y retrocompatibilidad.....	23
1.5.3 La cámara de color	24
1.5.4 Corrección de gamma.....	25
1.5.5 Inserción de la crominancia en la señal de video	25
1.6 GRABANDO Y CODIFICANDO EL VIDEO ANÁLOGO.....	26
1.6.1 Cómo interpretan el color las cámaras de video.....	26
1.6.2 La resolución de imagen.....	27
1.6.3 Visualizando una señal de video.	27
1.6.4 SMPTE timecode.....	28
1.7 FORMA DEL PIXEL.....	28
1.7.1 Frecuencia de Datos.	29
1.7.2 Área Visible.....	30
1.7.3 Relación de Aspecto.	31
1.7.4 Tamaño de la Ventana.	32
1.7.5 Editores.....	33
1.7.6 Círculos y Cuadrados.	33
1.7.7 Rotación de Pantalla.	34

1.7.8 Formato Panorámico (Wide Screen)	34
CAPITULO 2. VIDEO DIGITAL	37
2.1 DIGITALIZACIÓN DE LA SEÑAL DE TV	37
2.1.1 Formato de la señal.....	39
2.1.1.1Formato compuesto	39
2.1.1.2Formato en componentes.....	40
2.2 ESTANDAR DE CODIFICACIÓN DE VIDEO EN COMPONENTES	41
2.2.1Muestreo.....	41
2.2.2Margen dinámico.....	42
2.2.3Niveles de cuantificación.....	43
2.2.4Tasa binaria y ancho de banda.....	44
CAPITULO 3 ESTANDARES DE COMPRESIÓN DE VIDEO.....	45
3.1 ¿QUE ES MPEG?.....	45
3.2 COMPRESIÓN DE VIDEO EN EL ESTÁNDAR MPEG-1 (APLICACIONES MULTIMEDIA).....	49
3.2.1 Formato del video de entrada	50
3.2.2 Tipos de imagen MPEG	50
3.2.3 Descomposición en capas de una secuencia de video MPEG	54
3.2.4 Compensación en movimiento	58
3.2.5 Codificador y decodificador MPEG-1.....	59
3.2.6 Trama de vídeo MPEG-1.....	63
3.3 COMPRESIÓN DE VIDEO EN EL ESTANDAR MPEG-2.....	66
3.3.1 ESTÁNDAR MPEG-2 (APLICACIONES BROADCAST).....	72
3.3.1.1 PERFILES Y NIVELES MPEG-2	72
3.3.3 LA ESTRUCTURA "FRAME"	78
3.3.4 LA ESTRUCTURA "FIELD"	78

3.3.5 SCANNING (Exploración)	80
3.3.6 DESCRIPCIÓN DEL PROCESO DE CODIFICACIÓN MPEG-2.....	81
3.3.7 DESCRIPCIÓN DEL PROCESO DE DECODIFICACIÓN MPEG-2	84
3.3.8 CONSTITUCIÓN DEL PAQUETE DE TRANSPORTE MPEG-2	85
3.3.9 ORGANIZACIÓN DEL MÚLTIPLEX DE TRANSPORTE	89
3.3.9.1 Program association table (PAT).....	89
3.3.9.2 Program map table (PMT).....	89
3.3.9.3 Conditional access table (CAT).....	90
3.3.10 TABLAS PRINCIPALES DE DVB-SI.....	90
3.3.10.1 Network information table (NIT).....	90
3.3.10.2 Service description table (SDT).....	90
3.3.10.3 Event information table (EIT)..	90
3.3.10.4 Time and date table (TDT).....	90
3.3.11 TABLAS OPCIONALES DE DVB-SI	91
3.3.11.1 Bouquet association table (BAT).....	91
3.3.11.2 Running status table (RST).....	91
3.3.11.3 Stuffing tables (ST).....	91
3.3.12 INSERCIÓN DE LAS SECCIONES EN LOS PAQUETES DE TRANSPORTE	96
3.3.13 CÓMO DEMULTIPLEXAR MPEG-2	97
3.4 COMPRESIÓN DE VIDEO EN EL ESTANDAR MPEG-4.....	98
3.4.1 Lo nuevo del MPEG-4.....	99
3.4.2 Resoluciones del MPEG-4	100
3.4.3 Interactividad.....	101
3.4.4 Contenido sintético y más posibilidades	102
3.5 ESTÁNDAR MPEG-7 Y SUS APLICACIONES	102
3.5.1 Objetivos de MPEG-7	104
3.5.2 Principales Elementos del estándar MPEG-7.....	106
3.5.3 Áreas de aplicación del MPEG-7	109
3.5.4 El PostScript del multimedia	111
3.5.5 Las primeras soluciones con el lenguaje	113

3.6 MPEG-21	114
3.6.1 Definición MPEG-21	116
3.6.2 EJEMPLO DE USO DE MPEG-21	117
3.6.3 Objetos Digitales	118
3.7 MPEG AUDIO	122
3.7.1 El estándar MPEG audio.	122
3.7.2 Introducción al sistema MPEG-1.	123
3.7.3 MPEG-1 en detalle.	126
3.7.4 TRAMA DE AUDIO	135
3.7.5 MÁS SOBRE MPEG-1	137
3.7.6 MÁS SOBRE MPEG-2	140
3.7.7 COMPRESIÓN DE AUDIO	143
3.7.8 CODIFICACIÓN DE LA SUB-BANDA	144
3.7.9 CODIFICACIÓN POR TRANSFORMACIÓN.....	147
3.7.10 CODIFICACIÓN MPEG DE AUDIO	149
3.7.11 FORMATO DE LA TRAMA MPEG DE AUDIO	152

CAPTITULO 4 TEMAS RELACIONADOS CON EL VIDEO DIGITAL

.....	155
4.1 HDTV (TELEVISIÓN DE ALTA DEFINICIÓN ó TELEVISIÓN DIGITAL).....	155
4.1.1 USO DE LA ALTA DEFINICION	161
4.1.2 TELEVISION DE ALTA DEFINICION	161
4.1.3 CINEMATOGRAFIA	163
4.1.4 ALTA DEFINICION. LOS ESTANDARES.	165
4.1.5 AUDIO PARA ALTA DEFINICIÓN	171
4.1.5.1 FORMATOS DE AUDIO PARA ALTA DEFINICION.....	171
4.1.5.2 Sonido envolvente. Producción y postproducción.	173
4.1.5.3 Metadatos	174
4.1.5.4 Gestionando la mezcla.....	175

4.1.5.5 Reutilización de la mezcla.....	175
4.1.5.6 Emisión.....	176
4.1.6 Normas para HDTV.....	176
4.1.7 Formatos e intervalos.....	177
4.1.8 El HDV (Vídeo en alta definición a nivel doméstico)	181
4.1.9 HD. Recepción y visualización.	184
4.2 BALANCEO INTELIGENTE DE VIDEO.....	186
4.2.1 Conversión de videos para televisión a videos de computadora	187
4.2.2 Resolución	187
4.2.3 Señales Entrelazadas versus No entrelazadas.....	188
4.2.4 Formato de la señal.....	188
4.2.5 Duplicadores y Cuadruplicadores de Líneas	190
4.2.6 Un comentario importante sobre la compensación del movimiento	191
4.2.7 Limitaciones de los Duplicadores y Cuadruplicadores de líneas	192
4.2.8 Balanceo Inteligente de Video.....	194
4.3 VIDEO SOBRE REDES.....	196
4.3.1 Transmisión de vídeo: vídeo digital.	197
4.3.2 Digitalización.....	199
4.3.3 Soluciones de vídeo	204
4.3.3.1 APLICACIONES	207
4.3.4 CONECTIVIDAD	207
4.3.5 CONECTIVIDAD EN INFOGRAFÍA: SERVIDORES DE VIDEO Y AUDIO	208
4.3.5.1 Aplicaciones	211
4.3.6 CONECTIVIDAD EN MULTIMEDIA.....	213
4.3.7 LA RED GLOBAL: INTERNET.....	213
4.3.8 APLICACIONES	214
4.4 DVD (Disco Digital o Disco Versátil Digital).	218
4.4.1 ¿Cuales son las características del DVD-Video?	218
4.4.2 ¿Como es la calidad del DVD-Video?	220
4.4.3 ¿Cuales son las desventajas del DVD?.....	222

4.4.4 ¿Que títulos están disponibles en DVD?	222
4.4.5 ¿Que son los "códigos regionales" "códigos de país"?.....	223
4.4.6 ¿ Que hay acerca del DVD-Audio o Music DVD?	226
4.4.7 ¿Que estudios están apoyando el DVD?	227
4.4.8 ¿Puedo grabar en DVD desde VCR, TV, etc?.....	228
4.4.9 ¿Es diferente la carpeta de la del CD?.....	228
4.4.10 ¿Reemplazara el DVD al lector de cintas de video?	229
4.4.11 ¿Reemplazara el DVD al CD-ROM?	229
4.4.12 ¿Pueden los grabadores de CD-R crear DVDs?.....	229
4.4.13 ¿Cuanto duran los discos DVD?.....	230
4.4.14 ¿Que es un disco de doble capa? ¿Funcionara en todos los lectores?	230
4.4.15 ¿Es un estándar mundial el DVD-Video? ¿Trabaja con NTSC, PAL y SECAM?.....	231
4.4.16 ¿Que es el firmware y porque debo actualizarlo?.....	233
4.4.17 ¿Que pasa con las animaciones en el DVD? ¿Se comprimen mal?.....	233
4.4.18 ¿Por que algunos discos requieren cambio de cara? ¿No pueden los DVDs almacenar cuatro horas por cara?	234
4.4.19 ¿Porque esta la imagen apretada haciendo que las cosas parezcan muy flacas?	234
4.4.20 ¿Usan todos los lectores de video el Dolby Digital (AC3)? ¿Todos tienen 5.1 canales?.....	234
4.4.21 ¿Cual es la diferencia entre el DVD de primera, segunda y tercera generación?	235
4.4.22 ¿Que es lo que pasa entre el DTS y el DVD?.....	236
4.4.23 ¿Es compatible el CD-ROM con el DVD-ROM?	236
4.4.24 ¿Es compatible el CD-R con el DVD-ROM?.....	236
4.4.25 ¿Es compatible el video CD con el DVD?	237
4.4.26 ¿Es compatible el Super Video CD con el DVD?.....	237
4.4.27 ¿Es el Enhanced CD compatible con el DVD?	238
4.4.28 ¿Es compatible el MP3 con el DVD?.....	238
4.4.29 ¿Es compatible el laserdisc con el DVD?.....	238

4.4.30 ¿Que efecto tendra el FMD en el DVD?	239
4.4.31 ¿Como afecta el MPEG4 al DVD?.....	239
4.4.32 ¿Que pasa con la grabación en DVD: DVD-R, DVD-RAM, DVD-RW y DVD+RW?	240
4.4.33 ¿Que es el DeCSS y el DivX;-)?	240
4.5 CONVERTIDOR ANALOGICO-DIGITAL (ADC,Analogic to Digital Conversion)	242
4.5.1 Muestreo	245
4.5.2 Cuantización:	246
4.5.3 Codificación	247
CONCLUSIONES	249
GLOSARIO	251
BIBLIOGRAFIA	277
ANEXO	1

INTRODUCCIÓN

Etimológicamente, la palabra video proviene del latín *video, vidi, visum*; y significa “ver”. Esto en principio parece sencillo, pero adquiere cierto grado de complejidad al observar que el término video se utiliza con muy poca precisión, por lo que se presta a equívocos y generalizaciones diversas.

Originalmente, a esta tecnología de grabación electromagnética de la imagen visual se le llamó *videotape*, para diferenciar la filmación en celuloide tradicional del cine y la transmisión en directo de la televisión. El *videotape* o cinta de video registra tanto imágenes visuales como sonoras; sin embargo, el término hace referencia sólo al ámbito visual. En forma rigurosa se le debería llamar audiovideo, aunque desde el principio se le denominó “video”, porque es ahí donde radica su novedad primordial.

Así se nombra con la misma palabra video: al medio de comunicación, la cámara de registro y grabación, el aparato reproductor casero, al soporte en forma de casete y a los programas mismos.

El objetivo de este trabajo es hacer una recopilación bibliográfica sobre el tema de “VIDEO”, con la finalidad de que sirva como material de apoyo y consulta para alumnos y catedráticos de ingeniería, en especial para los de la carrera de Ingeniería en Electrónica y Telecomunicaciones. Ya que para los alumnos de esta carrera es importante que tengan conocimientos sobre el tema de video, y desafortunadamente este tema no es tratado en ninguna de las materias que integran el plan de estudios de la carrera. Si observamos los planes de estudio de la carrera de ing. Electrónica y Telecomunicaciones de otras universidades dicho tema es tratado como una materia mas del plan de estudios.

Por tal motivo me avoque a realizar un acervo bibliográfico, con los conceptos más importantes sobre dicho tema, esperando que sean útiles para los alumnos y catedráticos en específico de la carrera de Electrónica y Telecomunicaciones y también para los que estén interesados en saber un poco sobre este tema.

Dicho trabajo consta de cuatro capítulos, en los cuales se tratan los conceptos más importantes sobre el tema de video, como son el video analógico y el video digital.

En el primer capítulo comenzamos con los conceptos más importantes del video analógico, como son el barrido de la imagen, el proceso de exploración, y como esta formada la señal de video.

El segundo capítulo es una introducción al video digital, básicamente se explica como es digitalizada una señal de video y como se hace el muestreo y cuantificación de la señal digital.

El tercer capítulo trata sobre una parte importante del video digital, como lo son los estándares de compresión de video. En este capítulo se explica de manera detallada cada uno de los estándares MPEG, más utilizados en la actualidad.

El capítulo cuatro trata sobre algunos temas que se relaciona con el video.

- a) HDTV (televisión de alta definición), su uso sus estándares, los formatos de audio utilizados en la HDTV.
- b) Balanceo inteligente, como se duplican y cuadruplican las líneas, para que las señales analógicas puedan ser vistas en una TV digital.
- c) Video sobre redes, en el cual se tratan temas como el video Internet, video conferencias, streaming video, y la aplicación de cada uno de estos. El uso de servidores de video y audio.
- d) El DVD (disco versátil digital) es otro tema incluido en este capítulo, su definición, su uso, la calidad, etc.
- e) El convertidor analógico- digital, como es el muestreo, cuantización y codificación.

Finalmente como un anexo tenemos la comparación de formatos de video.

También es incluido un glosario con los términos más importantes y relacionados con el tema de video.

CAPITULO 1. FUNDAMENTOS DEL VIDEO ANÁLOGO

1.1 VIDEO ANÁLOGO

En un sistema análogo, la información se transmite mediante alguna variación infinita de un parámetro continuo como puede ser la tensión en un hilo o la intensidad de flujo de una cinta (véase la Figura 1). En un equipo de grabación, la distancia a lo largo del soporte físico es un elemento análogo continuo más del tiempo. No importa en que punto se examine una grabación a lo largo de toda su extensión: se encontrará un valor para la señal grabada. Dicho valor puede variar con una resolución infinita dentro de los límites físicos del sistema.

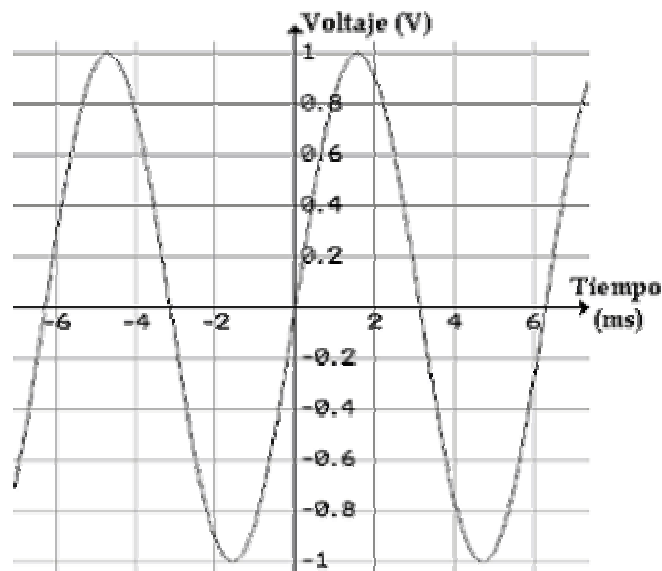


Figura 1. Variación infinita de un parámetro continuo en función del tiempo

Dichas características suponen la debilidad principal de las señales análogas. Dentro del ancho de banda permisible, cualquier forma de onda es válida. Si la velocidad del soporte no es constante, una forma de onda que sea válida pasará a ser otra forma de onda también válida; no es posible detectar un error de base de tiempos en un sistema análogo. Además, un error de tensión tan sólo hace variar un valor de tensión válido en otro; el ruido no puede detectarse en un sistema análogo. Se puede tener la sospecha de que existe ruido, pero no se sabe qué proporción de la señal recibida corresponde al ruido y cual es la señal original. Si

la función de transferencia de un sistema no es lineal, se producirán distorsiones, pero las formas de onda distorsionadas aún serán válidas; un sistema análogo es incapaz de detectar distorsiones.

Es característico de los sistemas análogos el hecho de que las degradaciones no puedan ser separadas de la señal original, por lo que nada pueda hacerse al respecto. Al final de un sistema determinado la señal estará formada por la suma de todas las degradaciones introducidas en cada etapa por las que haya pasado. Esto limita el número de etapas por las que una señal puede pasar sin que quede inutilizable. [1]

1.2 PROCESO DE EXPLORACIÓN DE LA IMAGEN.

Se debe recordar que todas las normas vigentes de televisión en la actualidad, NTSC (National Television Systems Comitee), PAL (Phase Alternation Line) y SECAM (Systeme Electronique Color Avec Memoire) se derivan, directa o indirectamente, de los estándares en blanco y negro definidos en los años 40 y 50.

Estas primeras emisiones utilizaban un barrido progresivo (todas las líneas de la imagen se barren consecutivamente, como se puede ver en la Figura 2).

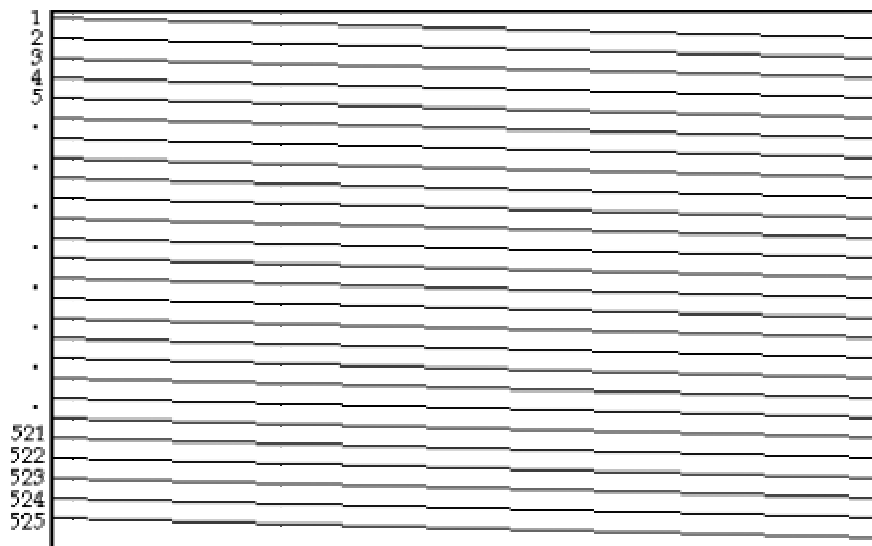


Figura 2. Representación simplificada del barrido progresivo

Por razones de orden práctico (radiaciones debidas a fugas magnéticas de los transformadores de alimentación, filtrados imperfectos), fue indispensable utilizar una frecuencia de imagen que estuviera relacionada con la frecuencia de la red (60 Hz en EE.UU., 50 Hz en Europa) para minimizar el efecto visual de estas imperfecciones; la frecuencia de exploración fue, por tanto, de 30 imágenes/s en EE.UU. y de 25 imágenes/s en Europa. Estas primeras imágenes presentaban un parpadeo bastante molesto (también llamado flicker de campo).

Tiempo después la captación de la imagen se hizo electrónica, haciendo que las definiciones alcanzaran un mayor número de líneas, esto gracias al barrido entrelazado. [1]

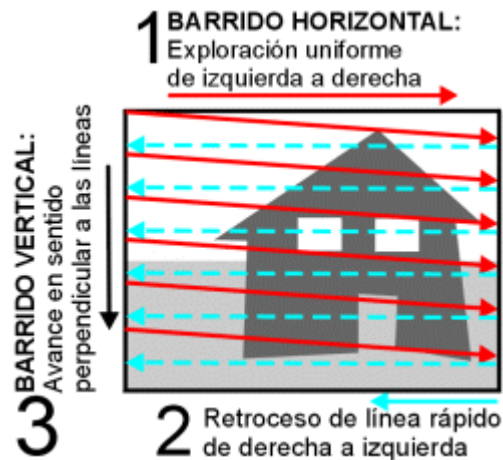


Fig. 3. representación electronica de una imagen.

Vamos a ver la forma en que representamos electrónicamente las imágenes, la señal de video y sus componentes. Tomemos una imagen, por ejemplo, la casa de la foto. Necesitamos convertir una superficie plana, con tonos de gris, zonas oscuras y claras en una sucesión de valores analógicos, una señal, para que la imagen pueda transmitirse. Bien, pues para obtener esta señal se escanea la imagen del siguiente modo: Empezando por la esquina superior izquierda se rastrea una línea de la que tomamos los valores de luminosidad, en adelante, Luminancia (Y). Este proceso se denomina Barrido horizontal.

Acabada la primera línea se rastrea la siguiente en sentido descendente hasta que, línea tras línea, tengamos rastreada toda la imagen. Este proceso se denomina Barrido vertical. Cada barrido vertical se corresponde con un cuadro de imagen o frame. De la sucesión rápida de cuadros en el tiempo se obtiene la sensación de movimiento. ¿Y el color? Se obtienen los tres valores de luminancia de cada color primario (en luz son el rojo, el verde y el azul). Es decir, se hacen en realidad, tres barridos simultáneos, uno para cada color básico y se obtienen, en consecuencia, tres señales de luminancia: RGB del inglés red, green & blue.

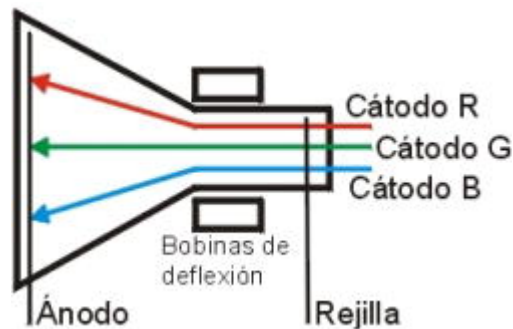


Fig. 4. tubo de rayos catódicos.

Las pantallas de TV y de ordenador son unos tubos de rayos catódicos adaptados a la proyección de imágenes. Tienen tres cátodos, uno para cada señal de luminancia (RGB) y el ánodo es, en realidad, la pantalla donde impacta el haz de electrones proveniente del cátodo. Para reconstruir la imagen se sigue el proceso que hemos explicado, pero a la inversa. Es decir, se dirigen los haces de electrones provenientes del cátodo para que efectúen los barridos horizontal y vertical sobre la pantalla. Esto se consigue mediante el uso de las bobinas deflectoras que se encargan de dirigir el haz de izquierda a derecha y de arriba abajo. A tal efecto, las bobinas reciben una señal sincrónica con el barrido, un diente de sierra que se construye a partir de los sincronismos horizontal y vertical.

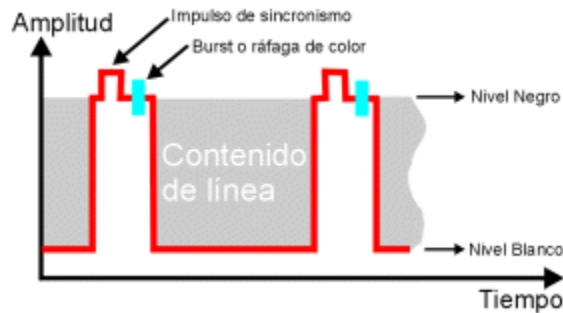


Fig. 5. contenido de línea.

Ahora tenemos tres señales de luminancia, una para cada color primario. ¿Cómo se vuelven a transformar en imágenes en una pantalla? Pues mediante la proyección de la señal en un tubo de rayos catódicos. Bueno, ahora también hay las pantallas TFT y de plasma, pero trabajan con la misma señal de video por lo que el ejemplo del tubo nos sirve perfectamente. Un tubo de rayos catódicos no es otra cosa que eso, un tubo, donde se ha hecho el vacío y se han colocado dos terminales eléctricos, el ánodo (-) y el cátodo (+). Si sometemos estos terminales a una tensión eléctrica suficiente, se producirá una corriente eléctrica entre ellos.

La señal de Vídeo Compuesto

Bien, pues ya tenemos todos los componentes de la señal de vídeo: las tres señales Y (RGB) más la señal de sincronismos (Horizontal y vertical). O lo que es lo mismo, ya sabemos que la señal de vídeo se compone en realidad de 5 señales, tres de información y dos de sincronismo. Para poder transmitir todo eso por un único canal, los precursores de la TV se las apañaron para componer todas las señales en una sola... El vídeo compuesto o línea. Este ha sido el estándar para la transmisión de vídeo durante muchos años y todos nuestros aparatos tienen entradas y salidas para este formato

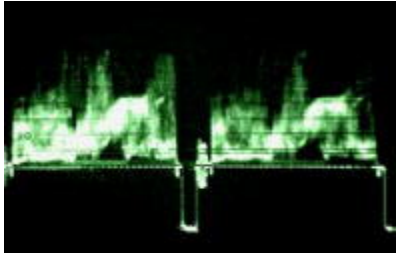
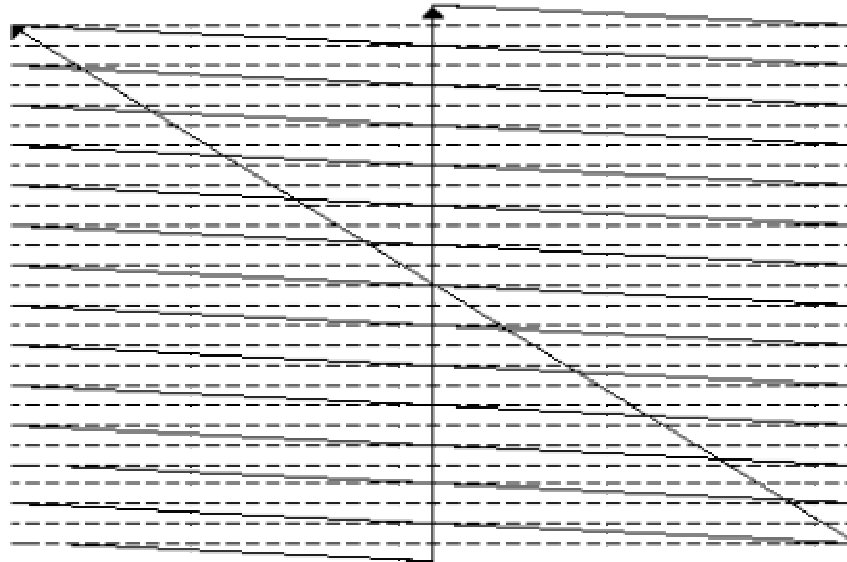


Fig. 6. video compuesto o linea.

En el gráfico se puede ver cómo lo hicieron. Se transmite en una sola señal la suma de las tres Y (RGB) los impulsos de sincronismo H y V y una tercera señal denominada ráfaga o burst de color o croma (Cr) que sirve para descomponer la suma de luminancias y volver a obtener RGB en el receptor. De este modo obtenemos tres señales en una: Sincronismo, Y (suma) y Cr (color). Este es el formato más simple y más práctico porque nos permite conectar dos dispositivos por el mismo cable, ya que el vídeo compuesto contiene todos los componentes de la señal. Ahora bien, es el de peor calidad ya que el receptor tiene que descomponerlo todo y la señal se degrada a medida que vamos acumulando etapas y procesos. Para conectar dos aparatos (TV y PC o lo que sea) por vídeo compuesto se suelen usar conectores RCA y cable coaxial. Sólo en aparatos antiguos o de gama profesional se usan conectores BNC. [2]

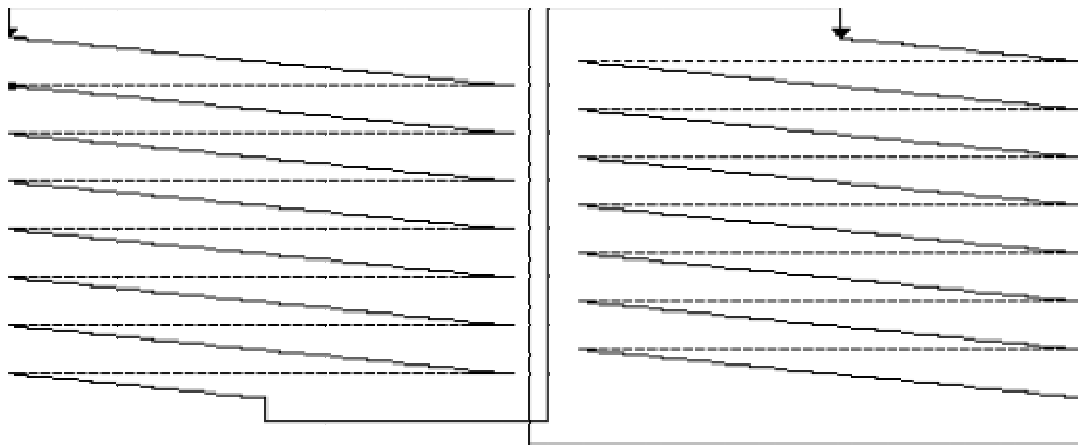
1.2.1 Barrido entrelazado.

Consiste en la transmisión de un primer campo compuesto por las líneas impares de la imagen y a continuación un segundo campo formado por las líneas pares, como se ve en la Figura 7. Esta forma de barrer la imagen, permite duplicar la frecuencia de refresco de la pantalla (50 o 60 Hz, en lugar de los 25 o 30 Hz) sin aumentar el ancho de banda para un número de líneas dado. Como se ve en la Figura 8, el barrido entrelazado se obtiene utilizando un número impar de líneas, por ejemplo 525 o 625 líneas que constituyen un cuadro, de manera que el primer campo comience en una línea completa, terminando en la mitad de otra línea, y el segundo campo comience en la mitad de una línea y finalice con una línea completa. En los países donde la frecuencia de la red es de 60 Hz, la velocidad de cuadro es de 30 por segundo y, por consiguiente, la frecuencia de campo es de 60 Hz.



Cuadro completo

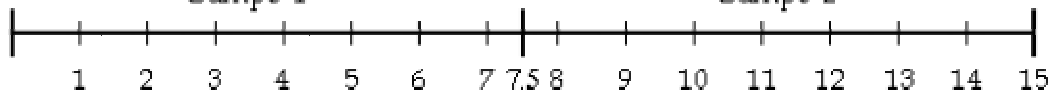
Figura 7. Barrido entrelazado 2:1



El primer campo comienza con una línea completa y finaliza con media línea El segundo campo comienza con media línea y finaliza con una línea completa

Campo 1

Campo 2



Debe haber un número impar de líneas en cada cuadro

Figura 8. Los campos de un entrelazado 2:1

1.3 FRECUENCIAS DE EXPLORACIÓN HORIZONTAL Y VERTICAL.

La velocidad de campo de 60 Hz es la frecuencia de exploración vertical. Este es el ritmo con que el haz electrónico completa su ciclo de movimiento vertical, desde la parte superior hasta la parte inferior de la pantalla para volver nuevamente a la parte superior.

El número de líneas de exploración horizontal de un campo es la mitad del total de las 525 líneas de un cuadro completo (en el sistema NTSC), ya que un campo contiene la mitad de las líneas. Esto da por resultado 262.5 líneas horizontales para cada campo.

Como el tiempo que corresponde a un campo es 1/60s y cada campo contiene 262.5 líneas, el número de líneas por segundo es:

$$262.5 \times 60 = 15750 \text{ líneas/s}$$

Esta frecuencia de 15750 Hz es la velocidad con que el haz electrónico completa su ciclo de movimiento horizontal de izquierda a derecha y regresa nuevamente a la izquierda.

1.3.1 Tiempo de línea horizontal.

El tiempo durante el cual se realiza la exploración de una línea horizontal es: $1/15750 \times 63.5 \mu \text{ s}$

1.4 LAS SEÑALES DE COLOR.

El sistema para la televisión en color es el mismo que para la televisión monocromática excepto que también se utiliza la información de color. Esto se realiza considerando la información de imágenes en términos de rojo, verde y azul. Cuando es explorada la imagen en la cámara, se producen señales de video separadas para la información de rojo, verde y azul de la imagen. Filtros de color separan los colores para la cámara. Sin embargo, para el canal estándar de 6 MHz de televisión, las señales de video de rojo, verde y azul son combinadas de modo que se forman dos señales equivalentes, una correspondiente al brillo y otra para el color. Específicamente las dos señales transmitidas son las siguientes:

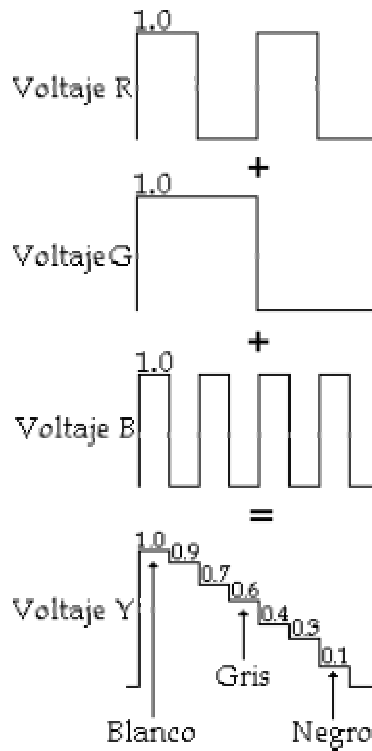
Señal de luminancia: Contiene solo variaciones de brillo de la información de la imagen, incluyendo los detalles finos, lo mismo que en una señal monocromática. La señal de luminancia se utiliza para reproducir la imagen en blanco y negro, o monocroma. La señal de luminancia o Y se forma combinando 30% de la señal de video roja (R), 59% de la señal de video verde (G) y 11% de la señal de video azul (B), y su expresión es:

$$Y = 0.30R + 0.59G + 0.11B \dots \text{ec (1)}$$

Los porcentajes que se muestran en la ecuación corresponden a la brillantez relativa de los tres colores primarios. En consecuencia, una escena reproducida en blanco y negro por la señal Y tiene exactamente la misma brillantez que la imagen original. La Figura 9 muestra como el voltaje de la señal Y se compone de varios valores de R, G y B. La señal Y tiene una máxima amplitud relativa de unidad, la cual es 100% blanca. Para los máximos valores de R, G y B (1V cada uno), el valor de brillantez se determina de la siguiente manera:

$$Y = 0.30 (1) + 0.59 (1) + 0.11 (1) = 1 \text{ lumen} \dots \text{ec (2)}$$

Los valores de voltaje para Y que se ilustran en la Figura 10 son los valores de luminancia relativos que corresponden a cada color.



$$Y = 0.30R + 0.59G + 0.11B$$

Figura 9. Obtención de la señal Y

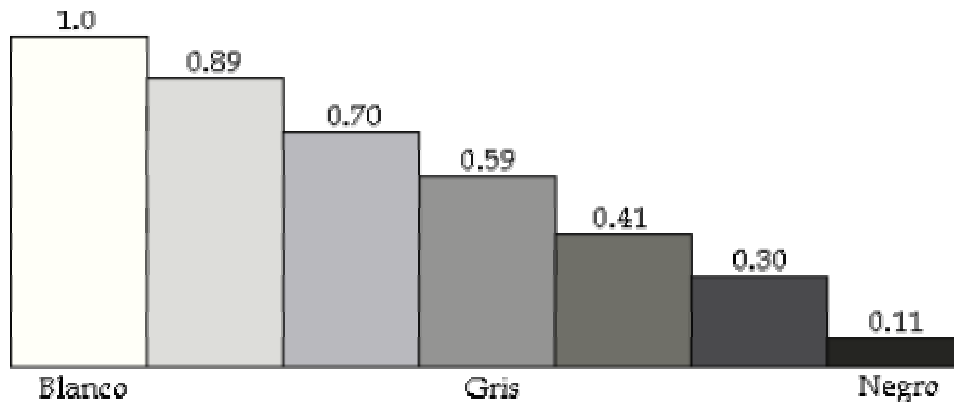


Figura 10. Valores de luminancia relativa

Señal de crominancia: La señal de crominancia es una combinación de las señales de color I y Q. La señal I o señal de color en fase se genera combinando el 60% de la señal de video en rojo (R), 28% de la señal de video en verde (G) invertida y 32% de la señal de video en azul (B) invertida, y se expresa como:

$$I = 0.60\tilde{R} + 0.28\tilde{G} + 0.32\tilde{B} \dots \text{ec (3)}$$

La señal Q o señal de color en cuadratura se genera combinando el 21% de la señal de video en rojo (R), 52% de la señal de video en verde (G) invertido y 31% de la señal de video en azul (B), y su expresión es:

$$Q = 0.21\tilde{R} - 0.52\tilde{G} + 0.31\tilde{B} \dots \text{ec (4)}$$

Las señales I y Q se combinan para producir la señal C y debido a que las señales I y Q están en cuadratura, la señal C o crominancia es la suma vectorial de estas, y su expresión es:

$$C = \sqrt{I^2 + Q^2}$$

$$\tan^{-1} \frac{Q}{I} \dots \text{ec (5)}$$

Las amplitudes de las señales I y Q son, en cambio, proporcionales a las señales de video R, G y B. La Figura 11 muestra la rueda de colores para la radiodifusión de televisión. Las señales \tilde{R}_Y y \tilde{B}_Y se utilizan en la mayor parte de los receptores de televisión a color para demodular las señales de video R, G y B. En el receptor, la señal C reproduce colores en proporción a las amplitudes de las señales I y Q. El matiz (o tono del color) se determina por la fase de la señal C y la profundidad o saturación es proporcional a la magnitud de la señal C. La parte exterior del círculo corresponde al valor relativo de 1.

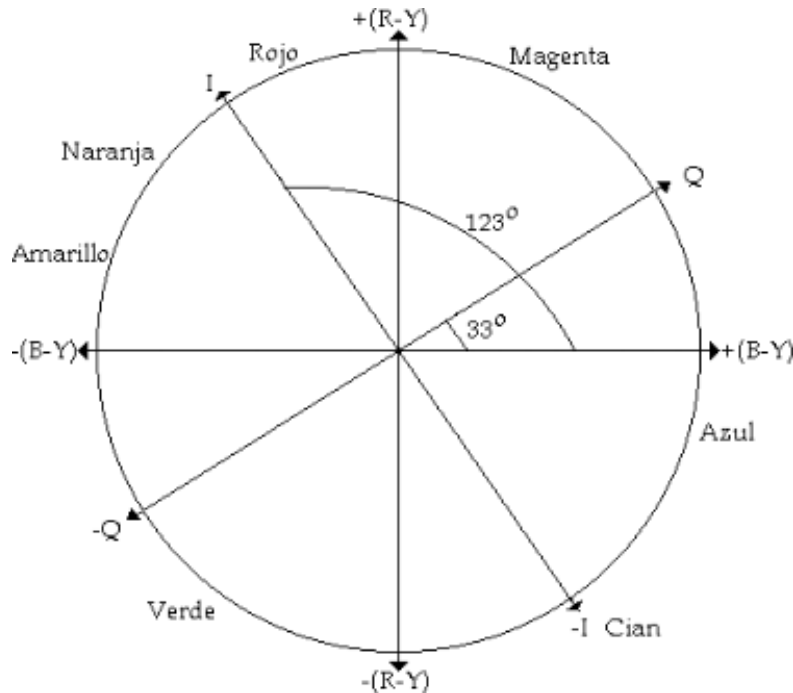


Figura 11. Representación de los colores en NTSC

Así se consigue que los sistemas de color y monocromáticos sean completamente compatibles. [1]

1.5 SISTEMAS DE TELEVISIÓN A COLOR

En la década del 40 solo existía televisión monocromática o blanco y negro, es decir, que la señal de video transmitida por las estaciones existentes, solo incluían la información de brillo de la imagen, la cual era representada en la pantalla del receptor como una sucesión de puntos con mayor o menor intensidad (tonos de grises).

Si bien hoy en día, ver imágenes en blanco y negro pueden no resultar atractivas, esta modalidad de transmisión logra cumplir con un objetivo muy necesario: dotar a la imagen reproducida de definición suficiente para que el espectador pueda discriminar dentro de la imagen, las formas, y tamaños relativos de los componentes de la escena. Cuando la tecnología pudo agregarle color a la imagen, hubo que analizar la forma de incluir dentro del canal de televisión, la información de color (crominancia), sin detrimento de la información de brillo (luminancia), ya existente.

1.5.1. Sistemas de exploración y sistemas de color

Conviene hacer una aclaración importante: hay que poder discernir entre 2 conceptos distintos, que en la práctica se suelen tratar indistintamente.

Los sistemas de exploración de imágenes de televisión, se refieren a la manera en que la imagen es barrida por el haz, la cantidad de líneas de definición, las frecuencias vertical y horizontal, y otras características. Así, existen normas como la M en EE.UU., Brasil, Japón, la norma N de Argentina, La norma B en Europa, etc.

Los sistemas de codificación de color de imágenes de televisión, se refieren a la manera en que se agrega la información de color a la imagen. Así, existen 3 sistemas clásicos que se implementaron en el mundo: NTSC, PAL y SECAM.

1.5.2 Compatibilidad y retrocompatibilidad

Se hace necesario agregarle a la señal monocromática de luminancia, la señal de crominancia. Para conseguir esto se debe mantener 2 condiciones importantes:

Compatibilidad

Es la propiedad de un sistema de televisión color que permite la reproducción de las emisiones a color, en los receptores monocromáticos existentes (por supuesto se verán las imágenes en blanco y negro, aunque se hayan generado en el transmisor a color).

Retrocompatibilidad o compatibilidad inversa

Es la propiedad de un sistema de televisión en colores que permite a los receptores de televisión en colores, reproducir en blanco y negro, las emisiones de un sistema existente en blanco y negro.

En ambos casos, las imágenes deben ser de buena calidad, por lo que la emisión en colores debe mantenerse dentro del canal de frecuencias previsto para blanco y negro, sin invadir canales adyacentes.

1.5.3 La cámara de color

Básicamente será igual a la monocromática, pero deberá tener algún agregado que le permita discriminar entre los 3 colores primarios que componen la imagen de toma, separarlos y obtener sendas señales de R, G y B. Esto se consigue con la inclusión dentro de la cámara de espejos muy especiales que en lugar de reflejar toda la radiación incidente, solo lo hacen con una pequeña banda de la misma, permitiendo que el resto de la radiación sea atravesada. Estos espejos se llaman dicróicos.

Entonces, con un juego de 2 espejos dicróicos y otro espejo normal se consigue separar la onda incidente en la cámara en sus 3 componentes primarias.

Con estos 3 colores se podrán reproducir la mayoría de los colores existentes en la naturaleza, por lo tanto, si se transmitieran estas 3 señales se podrían reproducir en un receptor destinado para este fin; sin embargo un televisor monocromático pre-existente no está preparado para recibir estas 3 señales, sino solo la Y.

Se deduce, de la colorimetría, que la relación entre Y los 3 colores primarios está establecida por la llamada ecuación fundamental de la Luminancia (ec. 1)

Conocido el hecho que se necesitan 3 señales para reproducir una imagen coloreada y una de las señales a transmitir es Y, resta todavía obtener 2 señales más, que conformaran la señal vectorial de crominancia. Estas 2 señales deberán tener la particularidad de anularse en caso de tratarse de una imagen monocromática (solo brillo). Este hecho, al igual que el anterior son necesarios en relación a las 2 premisas anteriormente nombradas, compatibilidad y retrocompatibilidad. Observando que el blanco se obtiene con iguales cantidades de los 3 primarios, por ejemplo el blanco de máximo brillo se obtiene con señales normalizadas con $R=G=B=1v$, se comprueba que $Y=1v$ también. Por lo tanto la señal de crominancia estará formada por 2 de las 3 señales diferencia de color R-Y, G-Y, B-Y. Solo será necesario enviar 2 señales, además de Y, dado que la tercera es combinación lineal de las otras. En el receptor, de igual manera se podrán recuperar las componentes R, G y B a partir de Y y C, donde C es la señal vectorial de crominancia formada por 2 señales de diferencia de color. Al deducir la expresión analítica de las diferencias de color, se

comprueba que la diferencia al verde G-Y es la que tiene coeficientes menores y por ende menor potencia por lo que será más susceptible al ruido. Por lo tanto la señal de crominancia C estará compuesta por la diferencia al rojo y diferencia al azul, también simbolizadas Cr y Cb.

1.5.4 Corrección de gamma

El brillo de una pantalla de un Tubo de Rayos catódicos TRC no guarda relación lineal con respecto a la intensidad del haz, sino más bien es una relación de cuadrática a cúbica. Debido a esto se hace necesario una corrección de Gamma en estudio, siendo Gamma (γ) el exponente de la intensidad I cuando se cumple la ecuación $L=k.I^\gamma$. La condición ideal sería un Gamma igual a 1.

1.5.5 Inserción de la crominancia en la señal de video

Se comprueba experimentalmente que la crominancia requiere un ancho de banda menor que el de luminancia, pero aun así, necesitan ubicarse ambas señales dentro del mismo canal de frecuencia, para cumplir la compatibilidad exigida. Para lograr esto hay que notar que el espectro de luminancia como el que se obtiene de crominancia, tienen la particularidad de ser discretos, es decir, están compuestos por rayas espectrales y no por una banda continua de frecuencias. La razón de esto hay que buscarla en el mismo proceso de generación de la imagen: a partir de la exploración punto a punto, lo que lo hace discretos.

Entonces, la solución se basa en intercalar ambos espectros, el de Y y el de C de manera de conseguir en el receptor una fácil separación de ambos. Esto se logra premodulando la croma con una frecuencia de subportadora color, que cambia según el sistema o la norma, pero que en todos los casos persigue un mismo objetivo, ubicar la porción de mayor potencia del espectro de croma en una zona donde el espectro de luma sea notoriamente inferior, admitiendo un posible caso que la separación de ambos espectros no se pueda conseguir con absoluta eficacia. Aunque este sea el caso, de todas formas se contribuye a la compatibilidad y las imágenes recibidas seguirán siendo de buena calidad, aunque se vea afectada una pequeña porción de espectro.

1.6 GRABANDO Y CODIFICANDO EL VIDEO ANÁLOGO.

Esta sección describe cómo las cámaras de video graban y codifican las señales de la imagen, y también cómo las cámaras de video interpretan el color y miden la resolución de las imágenes.

Las cámaras de video convencionales contienen dispositivos foto sensibles llamados Charge-coupled devices (CCDs), que digitalizan o capturan las imágenes individuales como las imágenes ópticas y las convierten o ponen en códigos o señales electromagnéticas. Una vez la señal de video analógico ha sido codificada por la cámara, ésta puede ser transmitida, grabada hacia video analógico, o se puede grabar digitalmente hacia un dispositivo de almacenamiento de disco. Las señales eléctricas capturadas por una cámara de video representan el color e información de brillo de la imagen. Las cámaras se clasifican, entre otras cosas, por su respuesta al color y resolución de imagen.

1.6.1 Cómo interpretan el color las cámaras de video.

Las cámaras de video interpretan el color como una combinación de los tres colores primarios aditivos: rojo, verde y azul. Este modelo se basa en la teoría del color a partir de la luz y normalmente se conoce como color RGB. Todas Las cámaras difieren en la manera como codifican la señal de color en un video, algunas procesan la señal separando en canales los componentes del RGB (Señal S-VHS), otras procesan la señal por croma (color) y luminancia (brillo), información que produce el campo de video, este proceso es el más común, poner la información del código RGB y la luminancia en una señal, conocida como señal compuesta (señal VHS).

En Estados Unidos y Japón la señal compuesta normal adoptada por la televisión y el video industrial es la NATIONAL TELEVISION STANDARS COMMITTEE (NTSC) Una señal NTSC tiene una proporción de 30 fps. (O más precisamente 29,97 fps.). En Europa, la señal compuesta más común es la señal PHASE ALTERNATING LINE (PAL) que tiene una proporción de 25 fps.

1.6.2 La resolución de imagen.

Otro concepto importante de una señal de video es la resolución de la imagen, cuya calidad depende del número de elementos del fotograma; Cada elemento se llama pixel.

Una imagen de video proyectada es una conglomeración de pixeles los cuales generan el color y brillo de la imagen. La calidad de la imagen incrementa si la cantidad de pixeles en el área de la imagen es mayor. Una cámara de video codifica la información de la imagen como una reja de pixeles, algo así como una colección de azulejos en un mosaico. En una señal NTSC un cuadro de video contiene 486 líneas horizontales de pixeles visibles, y cada una de estas líneas contiene 720 pixeles. Así, un cuadro de video NTSC esta conformado por 349.920 pixeles (486 x 720).

1.6.3 Visualizando una señal de video.

Para que la señal de video análogo o digital sea reconocible como una imagen debe correrse a través de un decodificador y ser convertida a una imagen. El decodificador lee la señal compuesta en RGB y permite que la imagen pueda desplegarse en pantalla. Las pantallas de televisor están hechas de diminutas partículas de fósforo que emiten intensidades variables de rojo ligero, verde, y azul luminoso y que son dirigidos en forma de haz de luz. Para que una señal normal pueda ser reconocida y proyectada en una pantalla debe reconocer 525 líneas 30 veces cada segundo.

En la realidad, una pantalla de televisor lee esta información de modo *entrelazado*, es decir, examina los electrones por líneas pares e impares en un fotograma, estas son inferidas como campos separados (2 por fotograma). Para mantener la tasa de 30 fps. La pantalla debe examinar los electrones a una velocidad de 60 campos por segundo. Cuando se congela un cuadro de video, realmente se ven dos cuadros y se presentan alternadamente pues previamente el monitor NTSC los examinó.

Una pantalla de computadora opera de modo "*noninterleaved*". Es decir, examina el haz de electrones (todas las filas de fósforos) consecutivamente para crear en pantalla la imagen y repite el proceso aproximadamente 60 a 75 veces por segundo refrescando la pantalla.

1.6.4 SMPTE timecode.

A la duración de un clip de video, desde su fotograma inicial hasta el último cuadro usado, se le asigna una unidad de tiempo o dirección llamada el *TIMECODE* (punto de corte). Timecode identifica cada cuadro dentro de un video en su edición y en su transmisión. Cuando se revisa un video, el timecode permite localizar los cuadros con precisión y sincronizar las imágenes con los elementos de audio (también es llamado sincronización exacta de cuadros).

El timecode usado por la SOCIETY OF MOTION PICTURE AN TELEVISION INGENEERS (SMPTE) identifica cada fotograma con una única dirección en forma de horas: minutos: segundos: cuadros. Un clip con una duración de 00:02:31:15, dura entonces 0 horas, 2 minutos, 31 segundos, y 15 fotogramas. A razón de 30 fps. Por segundo y es visualizado durante 2 minutos y 31.5 segundos.

Los estándares SMPTE Timecode son normas estrictas y claras para las diferentes ratas usadas en películas, videos, e industrias de televisión. Por razones técnicas involucradas en la transmisión, NTSC adoptó una norma de 29,97 fps En lugar de 30 fps, la cual se usó originalmente en el proceso de la programación de televisión en blanco y negro. El timecode de SMPTE para video norma NTSC asume una proporción de 30 fps que produce un 0,1 por ciento de diferencia en la proyección entre el tiempo real y la duración timecode.

[3]

1.7 FORMA DEL PIXEL

Todos los monitores de ordenador y formatos gráficos usan pixels cuadrados. ¿Por qué? Simplemente porque cada pixel representa la misma distancia horizontal que vertical. El número de pixels por pulgada (ppi) puede tener el valor que queramos. Así, si una imagen tiene 100 ppi, diremos que tiene 100 pixels por pulgada en cada una de sus direcciones.

Todo esto resulta obvio, pero si disponemos de una imagen con un valor de ppi horizontal y otro valor distinto en vertical, está claro que existe una manipulación de la imagen. Esto es lo que pasa con algunos formatos de video, especialmente en dos que son los más

populares, DV y MPEG. En estos formatos, los pixels no son cuadrados sino rectangulares, o sea que los valores de ppi horizontal y vertical, son distintos.

¿Por qué se usan los pixels rectangulares? La razón principal es porque existen dos sistemas de televisión a nivel mundial, cada uno con unas características distintas.

-El sistema Europeo está basado en 25 cuadros/segundo, 576 líneas/cuadro

- El sistema Americano está basado en 30 cuadros/segundo, 480 líneas/cuadro

Los términos NTSC, PAL, y SECAM se refieren a la técnica de codificación del color, no a la frecuencia de cuadro, aunque en la práctica todos los sistemas de 30 cuadros/segundo (fps) usan NTSC y los sistemas de 25 fps usan PAL o SECAM

La razón de las diferentes frecuencias de cuadro es histórica y está basada en la frecuencia local de la red eléctrica (50 ó 60 Hz). Para evitar parpadeos debido a la persistencia de las pantallas de televisión, se optó por dividir los cuadros de la imagen en dos campos, par e impar, explorando la imagen en campos alternos. El problema, técnicamente, desapareció hace más de 25 años, pero la compatibilidad hacia atrás de los equipos domésticos hace que todavía perdure hasta la fecha.

Un cálculo rápido de los dos sistemas, revela una interesante conexión: ambos tienen más de 14.400 líneas/segundo de video. Esto fue por diseño, no por coincidencia.

1.7.1 Frecuencia de Datos.

Una imagen de vídeo analógico es eso, una línea horizontal analógica continua. Verticalmente está dividida por un número fijo de líneas. La resolución que una TV es capaz de mostrarnos, depende de cómo de rápida pueda explorar esa señal toda la pantalla. Para digitalizar una imagen de TV, la señal ha de ser muestreada mientras explora la pantalla. La frecuencia de muestreo ideal reflejará la resolución inherente de la señal analógica.

En ambos sistemas de TV, el número de líneas/segundo es el mismo, y cada línea contiene la misma cantidad de información. Es lógico muestrear las líneas de TV a la misma frecuencia para ambos sistemas. Esto conlleva que el proceso de muestreo obtenga el mismo número de bits/segundo en el resultado final. Para los procesadores digitales de señal, esto minimiza la diferencia entre sistemas.

Así es exactamente como se escribió la norma internacional para la digitalización de la imagen de video y televisión, hace años. Se decidió que una línea de la imagen de televisión digital tuviera 720 pixels horizontales, tanto en los sistemas de 25 y 30 fps. (Técnicamente ambos sistemas muestrean exactamente a 13.5 millones de pixel por segundo)

Sin embargo, el número de pixels verticales de cada cuadro es igual al número de líneas de TV, siendo diferentes para ambos sistemas, 576 o 480 líneas.

Todos los formatos estándar de video digital tendrán 720 pixels de anchura: 720 x 576 para 25 fps y 720 x 480 para 30 fps. Esto incluye las emisoras de TV digital, reproductores de DVD, así como todos los formatos de cámaras de DV, tanto domésticas como profesionales.

1.7.2 Área Visible.

La señal de televisión analógica contiene más información que no se muestra en pantalla. Estas son las líneas extras (los sistemas tienen 625 y 525 líneas en total, respectivamente) no todas llevan información y están situadas al principio y al final del área visible de la pantalla.

El estándar DV no es diferente, los 720 pixels que forman la anchura de la imagen incluyen ocho en cada uno de los extremos del área visible especificada. Esto hace que en el centro de la pantalla sólo se representen 704 pixels reales de la imagen de televisión.

1.7.3 Relación de Aspecto.

Veamos el efecto de todo esto sobre un formato de pantalla. La imagen de TV se muestra en la pantalla con una relación de aspecto de 4:3 (ignoremos, de momento, el formato panorámico) Si los pixels son cuadrados, una imagen de 576 líneas tendrá 768 pixels de anchura, y una imagen de 480 líneas tendrá 640 pixels de anchura.

Sin embargo, en el formato DV siempre tenemos 720 pixels de anchura y sólo se muestran correctamente cuando los 704 pixels están centrados ocupando una anchura física de $\frac{4}{3}$ de su altura. El resultado que nos da son los pixels rectangulares - bajos y anchos a 25 fps (576 líneas) y altos y estrechos a 30 fps (480 líneas).

La relación de aspecto actual de los pixels es de 12:11 a 25 fps y de 10:11 a 30 fps. Tal como se ve en el dibujo:

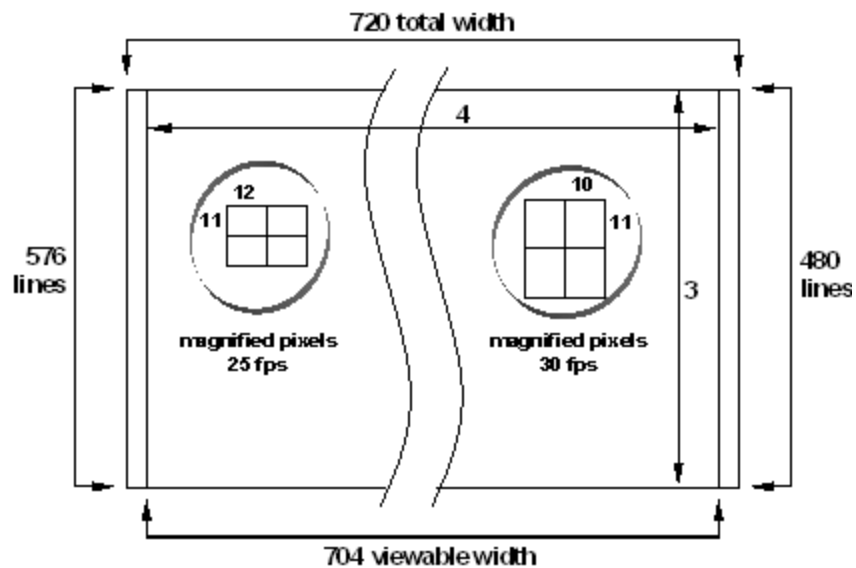


Fig. 12. La relación de aspecto actual de los pixels.

1.7.4 Tamaño de la Ventana.

Cuando pre-visualizamos un vídeo mientras editamos en la pantalla del ordenador, vemos las imágenes en la misma proporción que se verán en una televisión. Muchos programas de edición permiten elegir el tamaño de esa ventana, permitiéndonos mantener la relación de aspecto de la TV. Pero si, cuando trabajamos en DV, escalamos la imagen acorde al tamaño del pixel de vídeo (como lo hacen por defecto todos los programas de edición), veremos una imagen deformada.

Para evitar esto, el programa, haciendo uso intensivo del procesador, reescala la imagen de vídeo en una sola dimensión y luego nos la muestra. De esta forma minimiza la degradación de la imagen, manteniendo un número de líneas familiar. Es más sencillo mantener la anchura en 720 pixels y ajustar la altura.

Recordando que la imagen de TV es siempre de 704, esto nos da una altura total de 528 ($704:528 = 4:3$). El tamaño de la ventana para ver el formato completo de DV será de 720 x 528, como se ve en la figura:

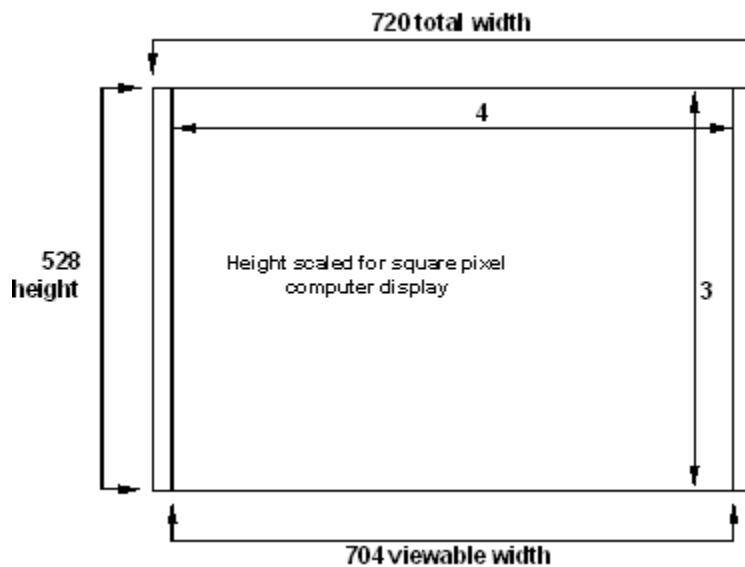


Fig. 13. El tamaño de la ventana para ver el formato completo de DV.

Para ventanas pequeñas, podemos usar valores fraccionales que minimizan el esfuerzo de reescalado y retienen la calidad de imagen. Con $\frac{1}{4}$ de pantalla tendremos una ventana de 360 x 264, etc. Todos estos tamaños son válidos para sistemas de 25 y 30 fps.

No debemos preocuparnos cuando editemos con el ordenador, si las imágenes están un poco distorsionadas (podemos ver la imagen alta y estrecha en América y baja y ancha en Europa) Todo esto se corregirá cuando pasemos la imagen a la televisión.

Sin embargo, hay situaciones en las cuales se pueden inducir errores de distorsión no visibles. Estas son:

- Importación de imágenes fijas
- Efectos que incluyan círculos o cuadrados
- Efectos que incluyan rotaciones

Nota: es conveniente disponer de un monitor de televisión externo cuando realicemos la edición con el ordenador. De este modo podemos controlar la inclusión de efectos geométricos en la edición.

1.7.5 Editores.

La mayoría de los paquetes de edición de video actuales, contemplan el concepto de relación de aspecto así como el tamaño y número de pixels. Adobe Premiere, Ulead MediaStudio, Sonic Vegas Video, EditDV, Xpress DV, etc. Realizan todos los cálculos referentes a la relación de aspecto y compensan automáticamente el tamaño del pixel antes de mostrarlo en pantalla. Por tanto, no debemos de preocuparnos más que de usar los valores correctos de la imagen que queramos capturar (720 x 576, 25 fps. En el caso de Europa)

1.7.6 Círculos y Cuadrados.

Trabajando en DV, si aplicamos un efecto o transición que contenga, por ejemplo, un círculo, el editor calculará el círculo asumiendo que los pixels son cuadrados. Sobre una televisión, el círculo lo veremos un poco ovalado. Si esto nos causa algún problema,

podremos ajustar esa transición (si lo permite) hasta encontrar el resultado deseado. En caso contrario, siempre podremos elegir otro tipo de efecto.

El mismo argumento lo aplicaremos a los cuadrados, aunque el ojo humano puede detectar las pequeñas desviaciones de un círculo perfecto mejor que las de un cuadrado perfecto.

1.7.7 Rotación de Pantalla.

Si usamos un efecto que implique una rotación de pantalla, tal como una imagen sobre imagen en movimiento, tendremos un problema similar. Si empezamos moviendo una imagen posicionada horizontalmente, con la altura y anchura ajustada para compensar la relación de aspecto, el programa mantendrá el tamaño del pixel constante, pero cuando giremos 90° la imagen, ésta se verá descompensada al hacer los cálculos. Será en este punto donde tengamos que realizar los ajustes.

Si en el efecto elegido podemos añadir puntos intermedios (keyframes), podremos añadir uno de esos puntos donde la imagen esté alineada horizontal o verticalmente. Entonces ajustaremos la relación de aspecto de la imagen en cada uno de esos puntos, teniendo en cuenta su orientación.

Como alternativa, también podremos no aplicar ninguna compensación. La relación de aspecto se mantendrá a medida que vaya girando la imagen y, a lo mejor, obtenemos un resultado de nuestro agrado.

1.7.8 Formato Panorámico (Wide Screen)

El formato DV también soporta el formato panorámico con relación de aspecto de 16:9. Sin embargo, esa imagen seguirá teniendo 720 pixels de anchura aunque, en el flujo de datos, se añadirá una información que permita a las televisiones de formato panorámico que muestren esas imágenes en toda su anchura.

Desde el punto de vista de la edición, el único cambio es que la diferencia entre la relación de aspecto del pixel y la relación de aspecto visible, es mucho más pronunciada

que cuando usamos el modo 4:3. Si nuestro programa de edición de video soporta la edición en modo 16:9, simplemente lo seleccionaremos. Él hará el resto.

Si usamos otro tipo de editores de vídeo, podremos aplicar los principios descritos anteriormente, pero usando diferentes factores de corrección. La mayor diferencia es que donde antes podíamos ignorar la distorsión debido a la relación de aspecto 4:3, ahora tendremos que trabajar mucho, ya que aquella será mucho mayor. Un ejemplo es el cuadro girando descrito anteriormente, que lo veremos muy descompensado en 16:9 a menos que no lo compensemos para las diferentes orientaciones.

Las ventanas de previsualización pueden tener 720 x 396, 360 x 198, etc. El factor de corrección será de 16/11 a 25 fps y de 40/33 (aprox. 6/5) a 30 fps. [4]

Las siguientes figuras muestran las dimensiones para el formato panorámico.

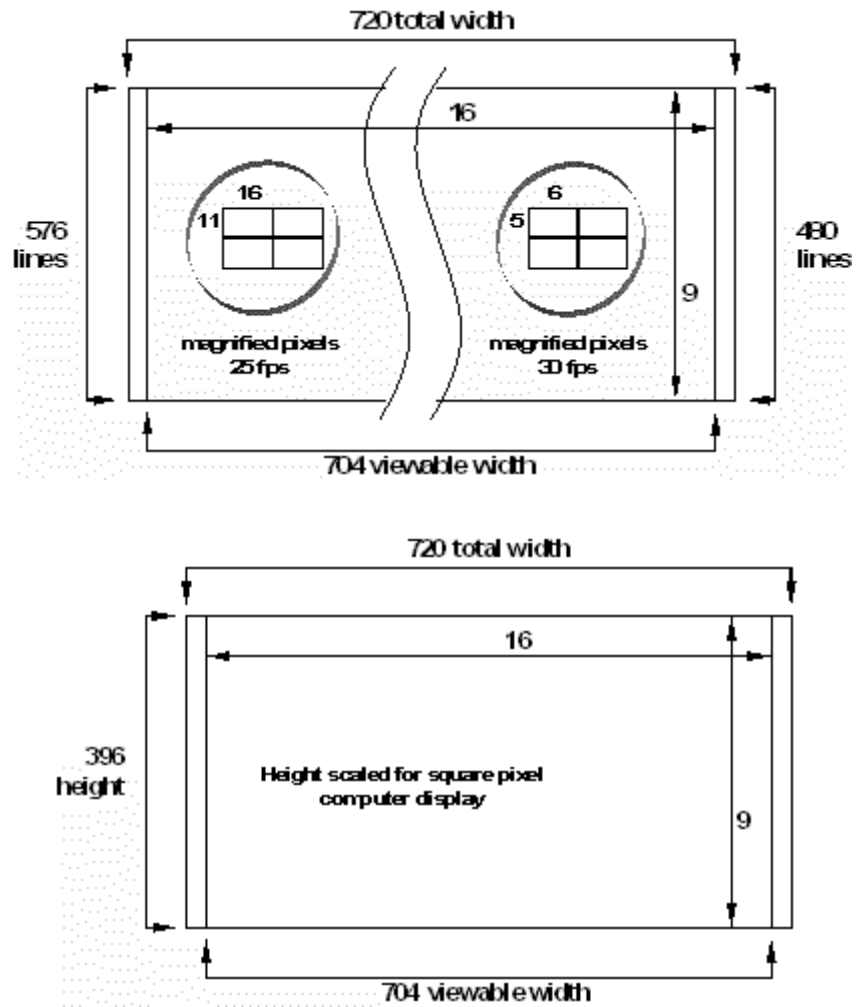


Fig. 13. Formatos panorámicos.

CAPITULO 2. VIDEO DIGITAL

La viabilidad de la puesta en practica del *tratamiento digital de la señal* de televisión y el desarrollo de *estándares de codificación y transmisión* permiten el establecimiento de una nueva tecnología con muchas ventajas y potenciales frente a la TV analógica actual.

Las características ventajosas de un sistema digital de TV pueden resumirse en:

- Ventajas para el usuario:

1. Posibilidad de incremento en el número de programas
2. Recepción de la señal desde receptores móviles sin pérdida de calidad
3. Integración de todos los aparatos audiovisuales domésticos
4. Canal de retorno: espectador “interactivo”
5. Globalización del concepto de pago por visión

- Ventajas técnicas:

1. Mayor eficiencia espectral: mayor número de canales en el mismo ancho de banda
2. Estándar a nivel mundial
3. Mayor robustez en la señal frente a ruido e interferencias multitrayecto

Todos estos desarrollos se unifican en el proyecto europeo DVB (*Digital Video Broadcasting*) que ha estandarizado la transmisión de señales digitales de TV por todos los medios conocidos: satélite, cable y terrestre. En dicho proyecto se adoptó el estándar MPEG-2 para la codificación de audio y video, cuyo resultado es la señal fuente a transmitir en TV digital.

2.1 DIGITALIZACIÓN DE LA SEÑAL DE TV

A pesar de la creciente introducción de la tecnología digital en la captación y representación de imágenes, todavía las cámaras de mayor calidad siguen siendo dispositivos esencialmente analógicos, así como lo son los instrumentos de visualización (pantallas de los televisores).

Por ello, para tener una imagen digital que forme la señal de TV digital deberemos digitalizar las señales proporcionadas por las cámaras e, inversamente, la señal recibida en un televisor deberá ser devuelta a su formato analógico para que sirva de entrada al TRC. Con ello el esquema básico a seguir en TV digital es el mostrado en la Figura 14.

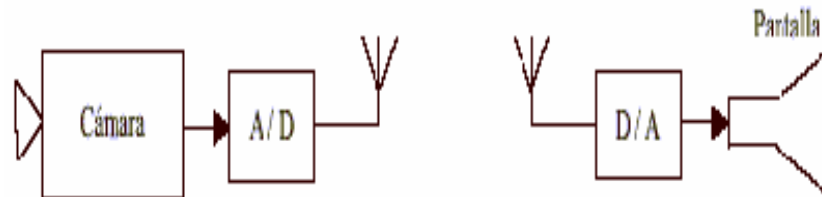


Figura 14. Esquema básico a seguir en TV digital

Es sabido, por el criterio de Nyquist, que la frecuencia de muestreo (número de muestras por segundo) debe ser mayor o igual que el doble de la frecuencia máxima de la señal continua (o de su ancho de banda si está en banda base). Cumpliendo esta premisa será posible reconstruir la señal original a partir de la muestreada. Tras el muestreo, el valor de la señal obtenido en cada muestra (aun en un rango continuo) se codifica digitalmente. La codificación se inicia con una cuantificación, que asigna el valor muestreado a un valor de entre un grupo finito de posibilidades. La elección de la cuantificación depende del tipo de señal y de la calidad deseada, ya que este proceso produce un error en cada muestra que se conoce globalmente como ruido de cuantificación.

Por ejemplo, para una señal analógica de amplitud variable en el rango $[-A, A]$ el paso o escalón de cuantificación Δ será: $\Delta = \frac{2A}{N} = \frac{2A}{2^n}$ (1) siendo N el número de niveles o escalones y n el número de bits elegidos para codificar cada muestra. Así por ejemplo, si se usan 8 bits/muestra se tiene un escalón del $1/256$ o $0,4\%$ del rango total de la señal. Si se emplearan 10 bits sería un escalón menor del $0,1\%$. Por dar una idea, la relación señal-a-ruido (de cuantificación) puede obtenerse con la siguiente fórmula:

$$SN = 6n + 10,8 \text{ (dB)} \quad (2) \dots \text{ec (6)}$$

Una vez codificadas las muestras se genera una trama de bits en serie cuya frecuencia o *tasa binaria* R_b , definida como número de bits por segundo, depende de la frecuencia de muestreo f_s y del número de bits por muestra n :

$R_b = f_s \cdot n$ (3) y aplicando el criterio de Nyquist para una señal de ancho de banda B :

$R_b > 2B \cdot n$ (4)

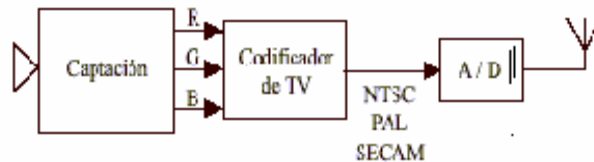


Fig 15. Esquema básico de digitalización en formato compuesto.

2.1.1 Formato de la señal

Una vez repasados los conceptos básicos de muestreo y cuantificación de cualquier señal, debemos pasar a ocuparnos expresamente de la señal de TV. Existen dos alternativas en la elección de las señales a digitalizar para representar la señal de TV color: digitalizar la señal de video compuesta o digitalizar cada componente.

2.1.1.1 Formato compuesto

Esta primera opción consiste en digitalizar la señal de video compuesta, que ha sido generada y codificada según algún estándar conocido: PAL, NTSC, SECAM, etc.

El diagrama de bloques genérico de este modo de funcionamiento se esquematiza en la Figura 15.

Las ventajas de este método son:

- Se necesita un único digitalizador
- Presenta una sencilla compatibilidad con los equipos analógicos actuales y pasados

Y entre los inconvenientes figuran:

- Las diferentes normas de TV color siguen siendo incompatibles entre si

- Los problemas inherentes a cada norma se mantienen: errores de tono en NTSC y saturación en PAL, cross-color, etc.

2.1.1.2 Formato en componentes.

En esta segunda opción, tal como se esquematiza en la Figura 16, se digitalizan tres de las señales que nos permitan formar la señal total de TV color. La elección de las tres señales o componentes, por simplicidad, nos lleva a la luminancia Y a las dos componentes de la crominancia $U = k_1(R - Y)$ y $V = k_2(B - Y)$. Si la señal nos la entrega directamente la fuente de imagen (cámara) el esquema es el presentado en la Figura 16a. En caso de tener como origen una señal ya en video compuesto (por ejemplo de un magnetoscopio) deberíamos insertar un descodificador del sistema usado (por ejemplo PAL), que nos entregue las tres componentes, que a continuación serían codificadas. Este caso se contempla en la Figura 16b.

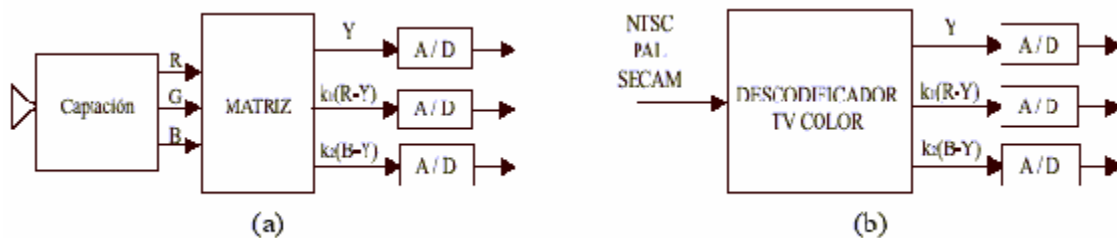


Figura 16. Esquema básico de digitalización en formato en componentes. (a) A partir de RGB. (b) A partir de un sistema analógico de TV color

Las ventajas de este formato se derivan, en primer término, de la eliminación de los inconvenientes del formato compuesto:

- Desaparecen las incompatibilidades entre sistemas, al partir de una señal previa a la aplicación de cualquier norma
- Desaparecen los problemas intrínsecos de los sistemas analógicos (imbricación de espectros, separación luminancia–crominancia, etc.)

Evidentemente, este procedimiento acarrea algunos inconvenientes:

- Se necesitan tres digitalizadores
- En principio necesitaríamos tres canales de transmisión (uno por componente) o un canal con un gran ancho de banda para enviar por el toda la información

Las necesidades de ancho de banda del canal se analizaran con más detalle posteriormente. Se describe a continuación una norma mundial que se ha convertido en estándar para codificación de video y que, como resultado, constituye también la base para la televisión digital que tratamos en este tema. La norma se conoce como *Recomendación 601* del CCIR o de la UIT, y consta de una familia de normas con diferentes niveles dependiendo de las aplicaciones de la codificación digital propuesta. En esta norma se trata siempre de una digitalización en componentes, siendo las señales a digitalizar: Y, k1 (R–Y) y k2 (B–Y). A continuación se tratan los diferentes aspectos contemplados en la norma.

1Las constantes k1 y k2 se fijaran mas adelante.

2.2 ESTANDAR DE CODIFICACIÓN DE VIDEO EN COMPONENTES

2.2.1 Muestreo.

Como ya se ha expuesto al inicio del tema, es necesario fijar una frecuencia de muestreo para digitalizar las tres componentes analógicas de la señal. La frecuencia de muestreo deberá cumplir dos requisitos fundamentales:

1. Criterio de Nyquist: $f_s > 2f_{max}$. Para que esta frecuencia de muestreo tenga una aplicación universal, deberemos buscar la mayor f_{max} de todos los estándares conocidos.

Concretamente, hay algunas variantes del SECAM que emplean $f_{max} = 6$ MHz, luego se deberá escoger $f_s > 12$ MHz.

2. Para favorecer la uniformidad y el funcionamiento de las memorias digitales, el muestreo deberá ser *ortogonal*, es decir, que todas las líneas contengan el mismo número de muestras. Esto se traduce en que $f_s = n f_H$. Esta segunda condición debe cumplirse también para todos los sistemas existentes. Las frecuencias de línea conocidas son dos: $f_H = 15625$ Hz (sistemas de 625 líneas) y $f_H = 15734,265732$ Hz (sistemas de 525 líneas). El mínimo común múltiplo de ambas es 2,25 MHz (con un error mínimo).

Atendiendo conjuntamente a ambas condiciones, la frecuencia mínima que las cumple es

$$f_s = 6 \times 2,25 = 13,5 \text{ MHz.}$$

Cuando se muestrean las tres componentes con $f_s = 13,5 \text{ MHz}$ se ésta llevando a cabo el llamado muestreo 4:4:4. Estos dígitos nos permitirán definir todo el conjunto de posibles muestreos de la norma, según el tipo de servicio.

El muestreo 4:4:4 no tiene en cuenta que el ojo humano exhibe una sensibilidad menor al color que a la luminancia. La recomendación 601 prevé otros muestreos, como 4:2:2, 4:1:1, 4:2:0, 2:1:1, 8:4:4, etc. que se corresponderán con diferentes niveles de calidad.

Uno de los muestreos asignados a la norma básica de televisión es el 4:2:2, que significa que por cada 4 muestras de luminancia se toman 2 de (R - Y) y 2 de (B - Y). Las frecuencias de muestreo serán $f_s = 13,5 \text{ MHz}$ para la luminancia y $f_s = 6,75 \text{ MHz}$ para las dos componentes de la crominancia. Con estas frecuencias los sistemas de 625 líneas poseen en cada línea 864 muestras de luminancia, 432 de (R - Y) y 432 de (B - Y), mientras que los sistemas de 525 líneas toman 858 muestras de luminancia, 429 de (R - Y) y 429 de (B - Y) en cada línea. Con el fin de homogeneizar el proceso posterior, se fijó el mismo número de muestras activas por línea para los dos sistemas: 720 de luminancia y 360 de crominancia. La posición de las muestras activas en cada línea se representa en la Figura 17.

Además, antes del muestreo las señales son filtradas con filtro pasa bajo cuyas frecuencias de corte son 5,75 MHz para la luminancia y 2,75 MHz para la crominancia. De este modo se evitan problemas de aliasing.

2.2.2 Margen dinámico

El margen dinámico de las tres señales debe ser 1 V, yendo de 0 a 1 V la luminancia, y de -0,5 a 0,5 V las componentes en diferencia de color (B - Y y R - Y). Inicialmente las señales Y, R, G y B están normalizadas entre 0 y 1 V, luego los márgenes de R - Y y B - Y serán $\pm 0,701$ y $\pm 0,886$ respectivamente. Para reducirlos al margen $\pm 0,5$ se emplearan los coeficientes antes introducidos: $k_1 = 0,5/0,701 = 0,713$ y $k_2 = 0,5/0,886 = 0,564$.

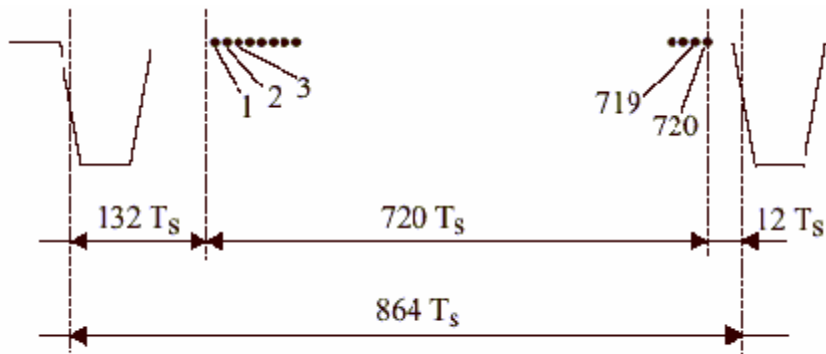


Fig. 17. posiciones de las muestras sobre cada línea.

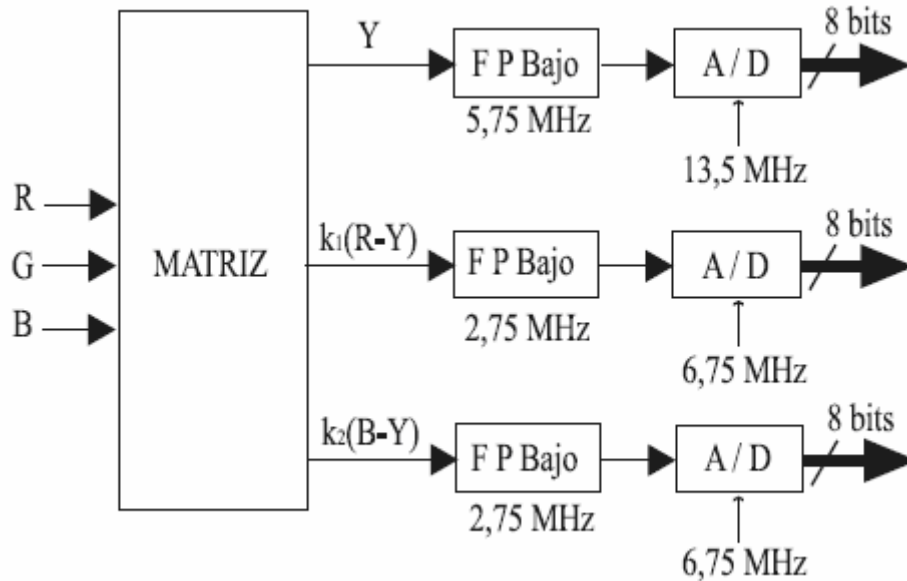


Fig. 18. esquema de muestreo 4:2:2 en la recomendación 601.

En definitiva, las señales que se digitalizan son: Y , $CR = 0.713(R-Y)$ y $CB=0.564(B-Y)$. El proceso de muestreo queda reflejado en la Figura 18.

2.2.3 Niveles de cuantificación

La recomendación propone una cuantificación con 8 bits por muestra para las tres componentes. Con ello se obtienen 256 niveles equidistantes para cada señal (de 0 a 255). Por norma se decidió que los niveles 0 y 255 se ocuparan para datos, y el resto para video.

Además, finalmente la luminancia sólo se cuantificara con 220 niveles y la crominancia con 225 niveles.

2.2.4 Tasa binaria y ancho de banda

Atendiendo al muestreo y a la cuantificación elegida podemos calcular la tasa binaria que se produce para esta norma de codificación.

Si se elige el muestreo 4:2:2 tendremos una tasa binaria de valor $R_b = 13,5 \text{ MHz} \times 8 \text{ bits} + 6,75 \text{ MHz} \times 8 \text{ bits} + 6,75 \text{ MHz} \times 8 \text{ bits} = 216 \text{ Mbits/s}$.

El muestreo escogido en el estándar convencional de TV digital es el 4:2:0, que aprovecha aún mas la poca sensibilidad al color del sistema visual humano y por cada linea sólo manda información de una de las dos componentes de color. Con este muestreo el régimen binario es: $R_b = 13,5 \text{ MHz} \times 8 \text{ bits} + 6,75 \text{ MHz} \times 8 \text{ bits} = 162 \text{ Mbits/s}$.

Si estudiamos la necesidad de ancho de banda para enviar estas tasas binarias, veremos que serian necesarios (como mínimo teórico) canales de 108 MHz y 81 MHz respectivamente.

Por lo tanto, se hace imprescindible reducir la tasa binaria necesaria para enviar la señal de TV, dando lugar a la introducción del concepto de compresión de la información que se explica en la siguiente sección. [5]

CAPITULO 3 ESTANDARES DE COMPRESIÓN DE VIDEO.

3.1 ¿QUE ES MPEG?

PEG es una familia de sistemas de compresión de imágenes estáticas, multinivel o multicapa pensado para fotografía.

JPEG incluye dos modos básicos de compresión:

compresión sin pérdidas basado en codificación predictiva seguida de codificación aritmética, donde la señal decodificada es idéntica a la original, pero el factor de compresión es muy pequeño de 1,6:1 a 2:1.

Compresión con pérdidas con factor de compresión variable y seleccionable por el usuario, puede estar entre 8:1 y 100:1.

Las normas MPEG

Estas normas se agrupan en varias especificaciones:

MPEG-1, esta norma se realiza para intentar ubicar imágenes en movimiento en soportes de tipo CD-ROM. Recordemos que la frecuencia de transferencia en este soporte (x1) es de 1,2 Mbits/sg.

MPEG-2 es una familia con capacidad de proporcionar imágenes y sonido para la recepción hasta HDTV (alta definición). La resolución de las imágenes MPEG-2 equivalen a la norma 601 (720 x 576 pixel/cuadro en la norma 625/50 y 720 x 480 pixel/cuadro en la norma 525/60), la velocidad de transmisión puede ir de 2 a 40 Mbits/sg y puede soportar relaciones de aspecto de 4/3 y 16/9, así como múltiples canales de audio y video en el mismo tren binario.

Se emplea codificación intercuadro con predicción y compensación de movimiento. En el proceso de predicción los macrobloques de 16 x 16 de un campo son comparados con todos los macrobloques del cuadro anterior, sobre un área de búsqueda dada.

En MPEG se encuentran imágenes codificadas intracuadro tipo "I" e imágenes predecidas intercuadro tipo "P", la primera imagen de un grupo debe ser una imagen "I" para poder arrancar la decodificación, el número de imágenes situadas entre dos imágenes "I" se denomina GOP.

MPEG define también otro tipo de imágenes denominadas bidireccionales "B" que pueden predecirse, desde imágenes posteriores o anteriores. La predicción bidireccional proporciona una mejor aproximación al valor real.

Las imágenes "B" pueden obtenerse a partir tanto de imágenes "I" como "P" pero no intervienen como fuente de predicción.

Un GOP de 8 indica que debe realizarse ediciones cada 8 campos a no ser que se decodifique.

Como se consigue una frecuencia de 1.5 Mbits en el MPEG-1. No sólo se elimina la redundancia, si no también buena parte de la entropía. Si la señal de entrada es del tipo 4:2:2, se comienza por descartar uno de cada dos campos a continuación la señal 4:2:2 se convierte en otra de tipo 2:1:0, lo que reduce los datos originales a 3/16. Esto genera lo que se conoce como formato de entrada SIF (Source Input Format 352 x 288).

MPEG-2 tiene diferentes **niveles y perfiles** que se presentan en el siguiente cuadro (tab. 1).

NIVEL (625/50)	PERFIL						
	SIMPLE		PRINCIPAL		SRN	ESPACIAL	
	Imag. 4:2:0	BNO	Imag. 4:2:0	B	Imag.B 4:2:0	Imag. 4:2:0	B 4:2:2
ALTO 1920 píxeles 1152 Líneas			80 Mb/s (MAX)				100 Mb/s (MAX)
ALTO 1440 píxeles 1152 Líneas			60 Mb/s (MAX)			60 Mb/s (MAX)	80 Mb/s (MAX)
PRINCIPAL 720 píxeles 576 Líneas	15 Mb/s (MAX)		15 Mb/s (MAX)		15 Mb/s (MAX)		20 Mb/s (MAX)
BAJO. 352 píxeles 288 Líneas			4 Mb/s (MAX)		4 Mb/s (MAX)		

Tabla 1. niveles y perfiles del MPEG-2

El límite superior de la frecuencia binaria del principal, es de 15 Mbit/sg, las evaluaciones del grupo MPEG, han demostrado que una imagen de calidad subjetiva similar a CCIR 601 se puede alcanzar con frecuencias de unos 9Mbits/sg, para aplicaciones de primeras generaciones, pero esto solo se logra con GOP de 12, lo que significa una precisión en la edición de 12 cuadros; válido para la radiodifusión pero no para su aplicación en estudio.

M-JPEG, en la actualidad **muchos sistemas de edición no lineal utilizan una variante de la norma JPEG diseñada para imágenes estáticas.**

Sin embargo, esta solución presenta las siguientes **desventajas**:

El M-JPEG no está normalizado, de forma que las aplicaciones de distintos fabricantes no son compatibles entre ellas.

La mayoría de las aplicaciones M-JPEG producen trenes binarios de frecuencia variable, lo que es apropiado para almacenamiento en disco duro, pero no para la grabación en cinta ni la distribución por canales de ancho de banda estables.

La transmisión digitalizada de imágenes se enfrenta al elevado ancho de banda necesario y a la elevada calidad exigida. Este es el caso de la televisión de alta definición (HDTV).

En 1988 se crea el comité MPEG, (Moving Pictures Expert Group, formalmente conocido como ISO/IEC JTC1/SC29/WG11) con el fin de desarrollar una técnica de codificación de vídeo y su audio asociado capaz de reducir la tasa binaria al entorno de los 1,5 Mbit/s. La calidad de la señal codificada debería ser lo suficientemente apropiada para su posible empleo en aplicaciones que integrasen vídeo, sonidos, imágenes, texto y gráficos.

Las tres partes fundamentales de este estándar describen un algoritmo de descompresión de vídeo, audio y una trama para el transporte de los datos obtenidos.

Los respectivos procesos de compresión no están definidos, lo que permite plena libertad en su realización, aunque su resultado debe garantizar la sintaxis establecida.

Posteriormente el comité MPEG prosiguió en la búsqueda de un algoritmo de codificación de alta calidad para cubrir los campos de distribución de señales de TV y HDTV con sonido multicanal (envolvente o surround), manteniendo una cierta compatibilidad con MPEG-1. Surge así el MPEG-2 o recomendación ISO/IEC IS 13818.

El MPEG-2 mantiene una estructura muy similar a la de MPEG-1 pero incluye mejoras y abarca más formatos de codificación y multiplexación.

TRAMA DE SISTEMA

Los sistemas MPEG definen las reglas sintácticas y semánticas que permiten la multiplexación en una sola trama (trama de sistema) de diferentes canales de audio, vídeo y datos junto con la información temporal necesaria para permitir la reproducción sincronizada y continuada de los mismos.

Esta trama es adecuada tanto para almacenamiento en sistemas digitales como para su transmisión por diferentes medios (radio, satélite, cable,...).

3.2 COMPRESIÓN DE VIDEO EN EL ESTÁNDAR MPEG-1 (APLICACIONES MULTIMEDIA).

Su principal objetivo es alcanzar un flujo de transmisión de datos constante de 1,5 Mbits/s (flujo de un CD-ROM de simple velocidad) del cual, 1.15 Mbits/s son para el video y los 350 Kbits/s restantes son para el sonido (estéreo) y para datos auxiliares.

La compresión de video utiliza los mismos principios que JPEG con pérdidas, a la que se le añaden nuevas técnicas que, juntas, forman el MPEG-1, que permiten reducir considerablemente la cantidad de información necesaria para la transmisión de imágenes sucesivas muy correlacionadas temporalmente.

Estas técnicas, llamadas de "predicción con compensación de movimiento", consisten en reducir, con un mínimo de información adicional, la mayoría de las imágenes precedentes (incluso las que le siguen).

Esto requiere un dispositivo de estimación de movimiento en el decodificador, que es la parte más compleja.

Tratándose de imágenes en movimiento o animadas, la descompresión deberá poder hacerse en "tiempo real" durante la reproducción. Por otro lado, la necesidad de un tiempo

de sincronización y de una respuesta de acceso aleatorio a una secuencia no demasiado largos (0.5 segundos máximo) limita el número de imágenes que pueden depender de la misma primera imagen a diez o doce para un sistema de 25 imágenes por segundo.

3.2.1 Formato del video de entrada

MPEG-1 se considera como un video solamente progresivo (no entrelazado), que alcanza un bitrate de 1.5 Mbps. La entrada de video es usualmente convertida primero al formato estándar de entrada MPEG SIF (Standard Input Format). El espacio de color adoptado es

Y- Cr- Cb según la recomendación CCIR 601. En el MPEG-1 SIF el canal de luminancia es de 352 pixeles x 240 líneas y 30 cuadros/segundo.

Los componentes de luminancia y crominancia son representados por 8 bit/pixel, y el componente de crominancia es submuestreado por 2 en ambas direcciones tanto vertical como horizontal. Mientras tanto los parámetros de video, los cuales son el tamaño de la imagen y la razón temporal, se pueden especificar, y por lo tanto son arbitrarios.

El siguiente conjunto de consideraciones contiene los parámetros específicos que ayudan a la implementación del hardware.

- Máximo número de pixeles/línea: 720
- Máximo número de líneas/imágenes: 576
- Máximo número de imágenes/seg: 30
- Máximo número de macrobloques/imagen: 396
- Máximo número de macrobloques/seg: 9900
- Máximo bitrate: 1.86 Mbits/seg
- Máximo tamaño del buffer del decodificador: 376832 bits

3.2.2 Tipos de imagen MPEG

MPEG define tres tipos de imágenes que se encadenan según el esquema de la Figura 19. Los cuales son el soporte de la codificación diferencial y bidireccional, minimizando la propagación de errores.

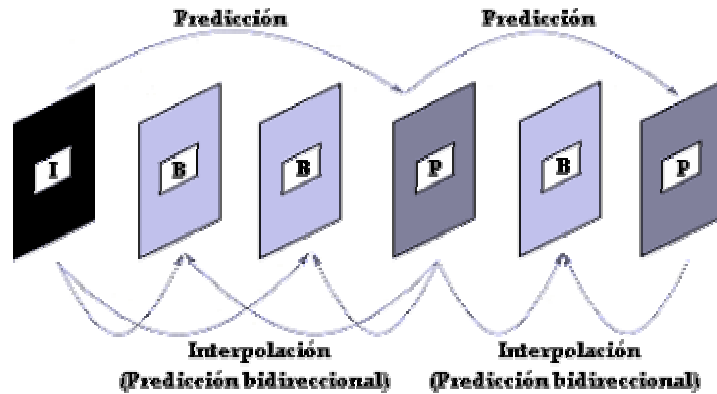


Figura 19. Encadenamiento de los 3 tipos de imágenes MPEG

1. Las imágenes I (intra)

Son imágenes que no requieren información adicional para su decodificación. Son codificadas sin ninguna referencia a otras imágenes, como en JPEG, es decir, que contiene todos los elementos necesarios para su reconstrucción por el decodificador y son, por ello, el punto de entrada obligatorio para el acceso a una secuencia.

La tasa de compresión de imágenes I es relativamente pequeña, comparable con la de JPEG con pérdidas. Ellas consisten ante todo de los coeficientes transformados y no contienen vectores de movimiento.

2. Las imágenes P (previstas)

Se codifican con respecto a las imágenes de tipo I o P anteriores, gracias a las técnicas de predicción con compensación de movimiento. Como la compensación de movimiento no es perfecta, no se podrá multiplicar indefinidamente el número de imágenes I, ya que, como se utilizan para decodificar otras imágenes P o B, se propagan amplificando cualquier error de codificación.

Su tasa de compresión es claramente mayor que la de las imágenes I. Las imágenes P requieren aproximadamente la mitad de los datos de las imágenes I.

3. Las imágenes B (Bidireccionales)

Se codifican por interpolación entre dos imágenes de tipo I o P precedentes y siguiente que las enmarcan. Como no se utilizan para describir otras imágenes, las imágenes B no propagan los posibles errores de codificación.

Este tipo de imágenes es el que ofrece el factor de compresión más alto, que generalmente es de una cuarta parte de los datos de las imágenes I.

Dependiendo de la complejidad del codificador utilizado, se podrán codificar solo las imágenes I, las imágenes I y P o las imágenes I, P y B; sin duda, con resultados absolutamente diferentes a nivel del factor de compresión y en cuanto a las posibilidades de acceso aleatorio, así como del tiempo de codificación y de la calidad percibida.

Los dos parámetros M y N definen la manera en que las imágenes I, P y B se encadenan:

- M es la distancia (en número de imágenes) entre dos imágenes P (previstas) sucesivas.
- N es la distancia entre dos imágenes I (intra) sucesivas.

Para alcanzar un flujo de video de 1.15 Mbits/s con una calidad satisfactoria, al tiempo que se mantiene una resolución de acceso aleatorio aceptable (< 0.5 segundos), los parámetros comúnmente utilizados son $M=3$ y $N=12$ como se muestra en la Figura 20.

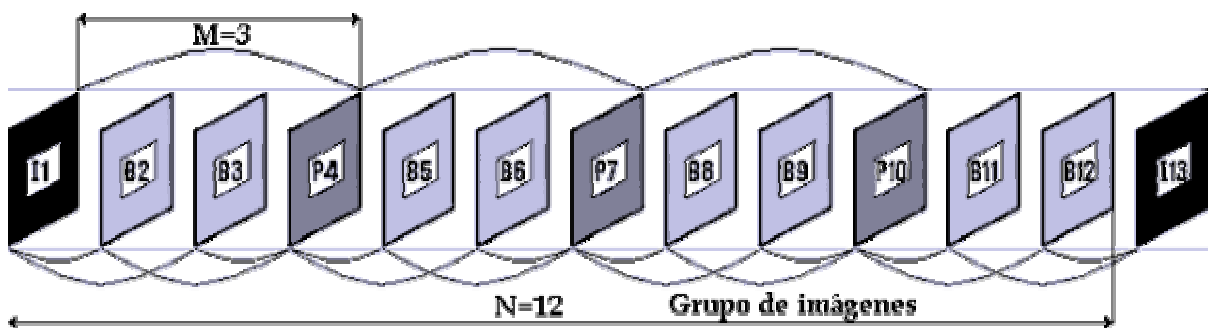


Figura 20. Ejemplo de grupo de imágenes, para $M=3$, $N=12$

En este caso, una secuencia de video se compone de $1/12$ (8.33%) de imágenes I, $1/4$ (25%) de imágenes P y de $2/3$ (66.66%) de imágenes B. El factor de compresión global se ve

favorecida por el hecho de que son las imágenes más frecuentes las que tienen un factor de compresión más alto.

En la visualización, tras la codificación y decodificación, es evidente que las imágenes de la secuencia de video deben ser reproducidas en el mismo orden en que se capturaron.

Con los parámetros definidos anteriormente ($M=3$, $N=12$), el modo de codificación de imágenes sucesivas se traduce por la correspondencia número □ tipo de imagen siguiente:

1(I) 2(B) 3(B) 4(P) 5(B) 6(B) 7(P) 8(B) 9(B) 10(P) 11(B) 12(B) 13(I) 14(B) 15(B) 16(P).....

Sin embargo, para codificar o decodificar una imagen B (Bidireccional), el codificador y el decodificador necesitarán la imagen I o P que la precede y la imagen P o I que la sigue. El orden de las imágenes será, por tanto, modificado antes de la codificación, de forma que el codificador y el decodificador dispongan, antes que las imágenes B, de las imágenes I y/o P necesarias para su tratamiento, o sea (Ver Figura 21):

1(I) 4(P) 2(B) 3(B) 7(P) 5(B) 6(B) 10(P) 8(B) 9(B) 13(I) 11(B) 12(B) 16(P) 14(B) 15(B).....

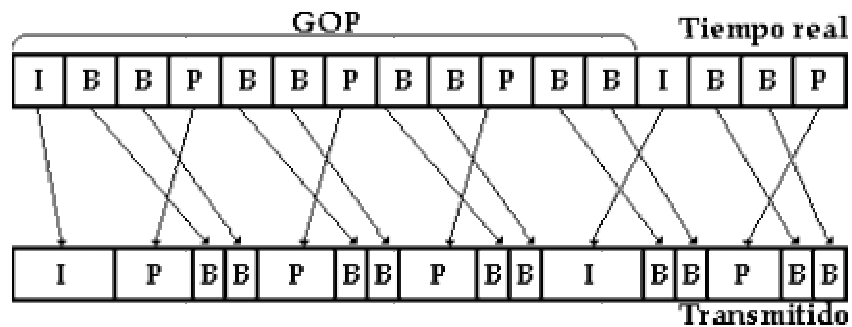


Figura 21. Comparación de las imágenes antes y después de la compresión, mostrando un cambio de secuencia

El aumento del factor de compresión facilitado por las imágenes B se paga, pues, con un tiempo de codificación/decodificación más largo (duración de dos imágenes) y un aumento en el tamaño de la memoria necesaria tanto en el codificador como en el decodificador (hay que almacenar una imagen suplementaria).

La Figura 22 muestra una curva de calidad constante donde la tasa de bits cambia con el tiempo de codificación. A la izquierda, solamente se utilizan imágenes I o codificación espacial, mientras que a la derecha solo se utilizan imágenes sucesivas IBBP. Esto significa que hay una codificación bidireccional de imágenes entre imágenes de codificación espacial (I) e imágenes previstas (P).

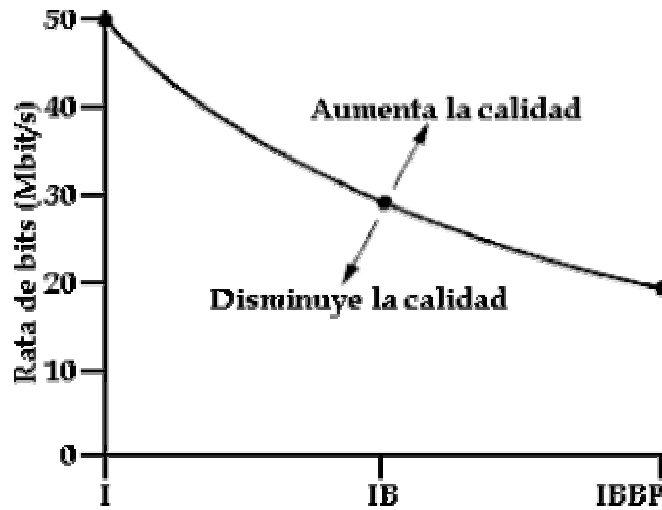


Figura 22. Curva de calidad constante

3.2.3 Descomposición en capas de una secuencia de video MPEG

Una secuencia de video MPEG es básicamente la salida del material en bruto de un codificador y contiene no más que lo necesario para que un decodificador restablezca la imagen original. La sintaxis de la señal comprimida es definida de manera rigurosa por MPEG, así se asegura que el decodificador cumpla con esta.

La Figura 23 muestra la construcción de una secuencia de video MPEG constituida por capas bien definidas.

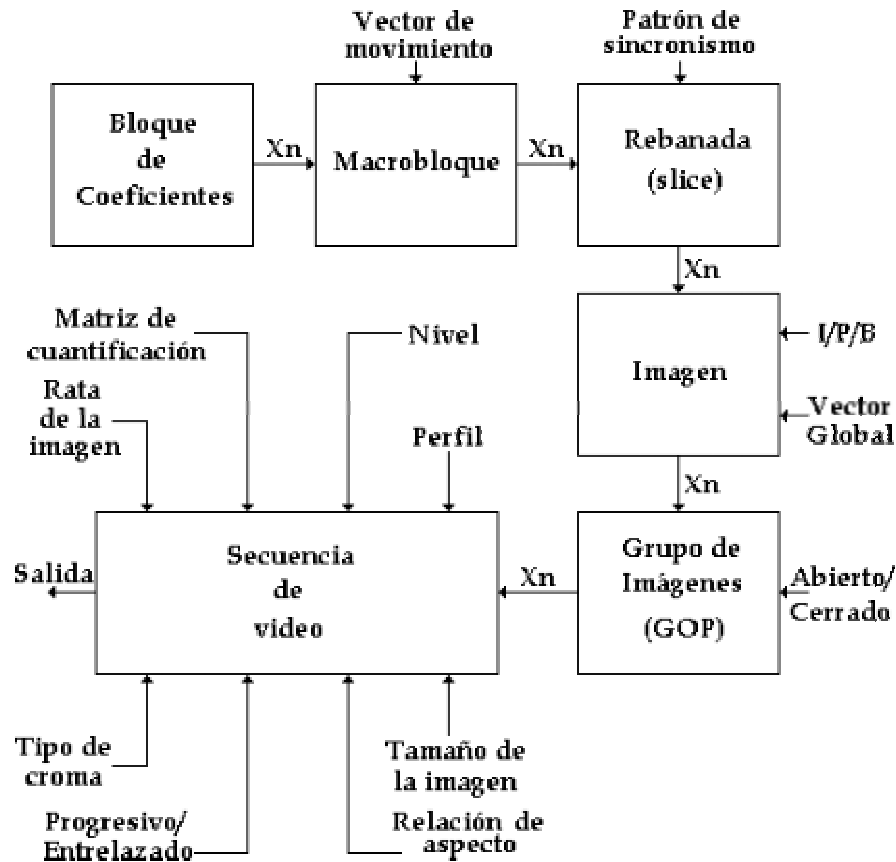


Figura 23. Estructura de una secuencia de video MPEG

1. Bloque (Block)

Es la unidad fundamental de la información de la imagen y esta representada por un bloque de coeficientes DCT, que tienen un tamaño de 8x8 pixeles, los cuales representan datos Y, Cr o Cb.

Aquí el coeficiente DC es enviado primero ya que este representa con mayor precisión la información de este bloque. Los demás coeficientes son enviados al final de este.

2. Macrobloque (Macroblock)

Es la unidad fundamental de la imagen que además está compensada en movimiento. Cada macrobloque es un vector de desplazamiento en dos dimensiones situado en la parte superior de la secuencia. En una imagen B, el vector puede ser hacia adelante o hacia atrás.

La compensación de movimiento puede ser en modo de cuadro o en modo de campo, el cual es indicado. La escala utilizada para la recuantificación de los coeficientes también es indicada. Usando los vectores, el decodificador obtiene información acerca de las imágenes anteriores y las posteriores, produciendo así una predicción de imágenes. Los bloques son transformados inversamente para producir una imagen de rectificación que es adicionada a la imagen prevista que ha sido producida a la salida del decodificador.

En un formato de codificación 4:2:0, cada macrobloque tendrá 4 bloques Y, y dos bloques de color diferente. Para hacer posible la identificación de cada bloque y sus componentes, estos se envían en un orden específico. Cada macrobloque tiene un tamaño de 16 x16 pixeles.

3. Rebanada (Slice)

Los macrobloques son reunidos en rebanadas, y aquellas siempre deben representar una fila horizontal que está ordenada de izquierda a derecha.

En MPEG, las rebanadas pueden comenzar en cualquier sentido y ser de tamaño arbitrario, pero las ATSC (Advance Television Systems Committee) establecen que ellas deben comenzar en el borde izquierdo de la imagen. Las rebanadas son la unidad fundamental de sincronización para la codificación de la longitud variable y diferencial, los vectores iniciales en una rebanada son enviados completamente, mientras que los demás vectores son transmitidos diferencialmente.

En imágenes I, los primeros coeficientes DC de las rebanadas son enviados completamente y los demás coeficientes DC son transmitidos en forma diferencial. En imágenes de diferencia, esta técnica no se utiliza.

4. Imagen (Picture) de tipo I, P o B

Cuando un número de rebanas se combinan, construyen una imagen, la cual es la parte activa de un campo o un cuadro.

La imagen de soporte inicial define qué imágenes I, P o B codifica e incluye una referencia temporal para que la imagen pueda ser representada en el momento adecuado. En el caso de tomas panorámicas e inclinaciones, los vectores en cada macrobloque serán los mismos. Un vector global puede ser enviado para toda la imagen, y luego se pueden enviar vectores individuales que lleguen a crear la diferencia en el vector global.

5. Grupo de imágenes (Group Of Pictures o GOP)

Las imágenes pueden ser combinadas para producir un GOP (grupo de imágenes) que comienza con una imagen I. El GOP es la unidad fundamental de codificación temporal. En el estándar MPEG, el uso de GOP es opcional, pero esta en la práctica es necesaria. Entre imágenes I, un número variable de imágenes P y/o B pueden ser colocadas como ya se ha descrito. Un GOP puede ser abierto o cerrado. En un GOP cerrado, las últimas imágenes B requieren de una imagen I para el siguiente GOP por decodificar y la secuencia de bits puede ser cortada al final de la GOP.

6. Secuencia

Cuando algunas GOP son combinadas se produce una secuencia de video con un código de inicio, seguido por un encabezamiento, y luego termina con un código final. Códigos de soporte adicional pueden ser situados al inicio de la secuencia. La secuencia de soporte especifica el tamaño horizontal y vertical de la imagen, norma de barrido, la rata de imágenes, si se usa un barrido progresivo o entrelazado, el perfil, nivel, velocidad de transferencia de bits, y cuales matrices de cuantificación se usan para codificar imágenes espaciales y temporales.

Sin la secuencia de soporte de datos, un decodificador no puede comprender el flujo de bits y por lo tanto no puede comenzar la operación de decodificación correcta. Esto ocurre generalmente cuando un televidente está cambiando canales de un lugar a otro en su televisor.

3.2.4 Compensación en movimiento

Es un proceso mediante el cual se mide eficazmente el movimiento de los objetos de una imagen a otra. De este modo se consigue medir qué tipos de movimientos redundan entre imágenes. La Figura 24 muestra que la imagen en movimiento puede ser expresada en un espacio de tres dimensiones que resulta de mover a lo largo del eje del tiempo dos imágenes consecutivas.

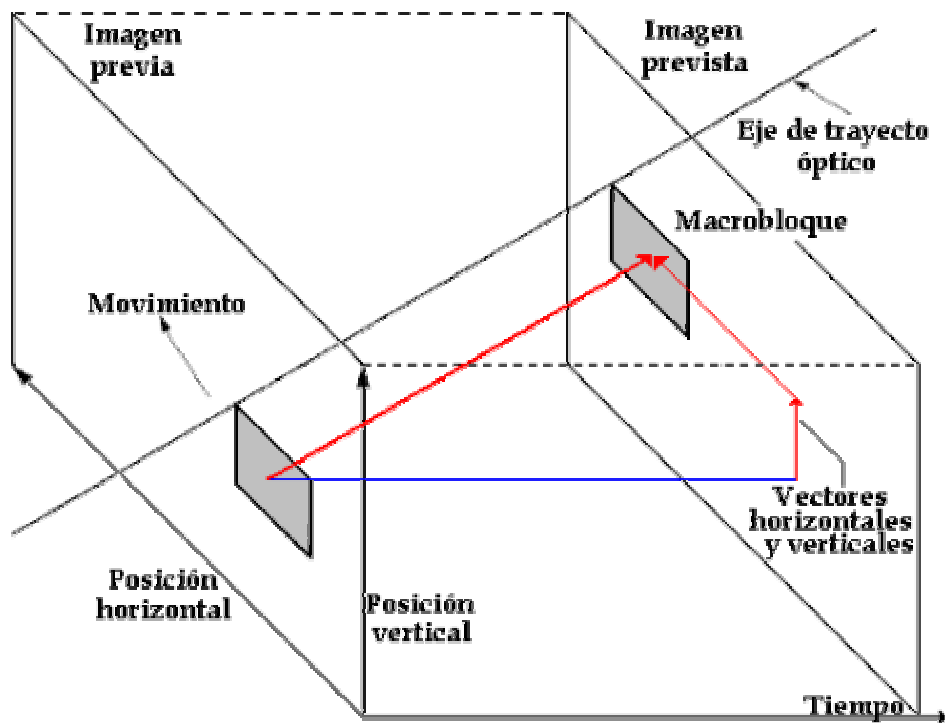


Figura 24. Objetos viajando en un espacio de tres dimensiones.

En el caso de un objeto que permanezca estático, el movimiento de este solo se ve en el eje del tiempo. Sin embargo, cuando un objeto está en movimiento, este se mueve en el eje de trayecto óptico (eje horizontal y vertical en el tiempo) que no es paralelo al eje del tiempo.

El eje de trayecto óptico une los puntos de movimiento de un objeto a medida que este se mueve a través de varias imágenes. Este trayecto puede ser nulo en valores que representen un objeto en movimiento que solo cambia con respecto al eje del tiempo. De igual manera, al mirar un objeto en movimiento que cambia su apariencia; uno de estos movimientos es el de rotación.

Para un simple movimiento de traslación los datos que representan un objeto son altamente redundantes con respecto al eje de trayecto óptico. Así, que de este modo, el eje de trayecto óptico puede ser hallado, generando un código de ganancia cada vez que se observa la presencia de un mismo objeto en movimiento.

Un codificador de compensación de movimiento trabaja de la siguiente forma. Una imagen I es enviada, pero esta es almacenada de tal modo que pueda ser comparada con la siguiente imagen de entrada para encontrar así varios vectores de movimiento, los cuales pueden ser utilizados en diferentes áreas de la imagen. Luego la imagen I es combinada de acuerdo a estos vectores o cancelada a una codificación espacial debido a su no conveniencia. La imagen prevista resultante es comparada con la imagen actual para producir una predicción de error también llamada residual. La predicción de error es transmitida con los vectores de movimiento. En el receptor la imagen I original es también retenida en la memoria, esta es cambiada de acuerdo con los vectores de movimiento transmitidos para crear la imagen prevista y luego la predicción de error es adicionada recreando la imagen original. Cuando una imagen es codificada de esta manera, es llamada imagen P en MPEG.

3.2.5 Codificador y decodificador MPEG-1

Un codificador de MPEG-1 incluye módulos de estimación de movimiento, selección de modos de compresión por macrobloques, conjunto de valores para la matriz de cuantificación, predicción de compensación de movimiento, cuantificador y decuantificador, transformada discreta del coseno (DCT), transformada inversa discreta del coseno (IDCT), códigos de longitud variable (VLC), un multiplexor, una memoria intermedia y un regulador de memoria. Ver Figura 25.

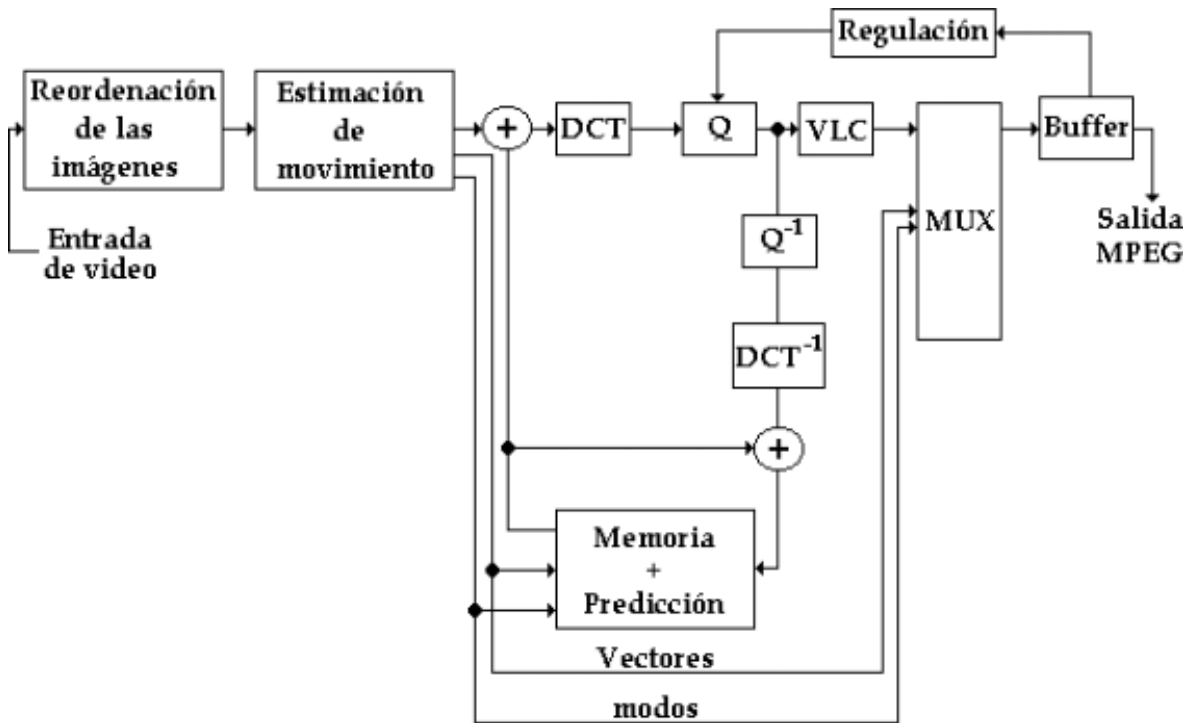


Figura 25. Esquema simplificado del codificador MPEG-1

La decuantificación y la transformada discreta del coseno inversa (IDCT) se necesitan en el codificador, porque la predicción de imágenes está basada en la reconstrucción de datos. El número relativo de imágenes I, P y B en la GOP depende específicamente de la aplicación a utilizar.

El estándar MPEG-1 especifica que se debe utilizar mínimo una imagen I, cada 132 imágenes para así evitar la propagación de errores en el módulo IDCT, creando de esta manera diferencias entre el codificador y el decodificador.

El uso de imágenes B es opcional. Ni el algoritmo de estimación de movimiento, ni el criterio de selección del modo de compresión, ni la matriz de cuantificación son parte del estándar. En general, la estimación de movimiento es calculada usando los datos de luminancia solamente. Un solo vector de movimiento es calculado para cada macrobloque. El rango para estimar el movimiento de un objeto en un macrobloque es de medio píxel. Un vector de movimiento que se refiera a unos píxeles que se hallan fuera de la imagen actual no es tomado en cuenta.

En resumen, un codificador típico de MPEG-1 realiza los siguientes pasos:

- Define los parámetros M y N que determinan de qué forma se encadenan las imágenes I, P y B en una GOP.
- Estima los vectores de movimiento para cada macrobloque en las imágenes P y B.
- Determina el modo de compresión para cada macrobloque de la imagen seleccionada (compresión espacial o temporal).
- Selecciona la matriz de cuantificación.

Un decodificador MPEG-1 revierte las operaciones del codificador. Donde la secuencia de video codificada de entrada es demultiplexada en coeficientes DCT y por otro lado es sacada la información de modos de compresión, vector de movimiento y matriz de cuantificación. Ver Figura 26.

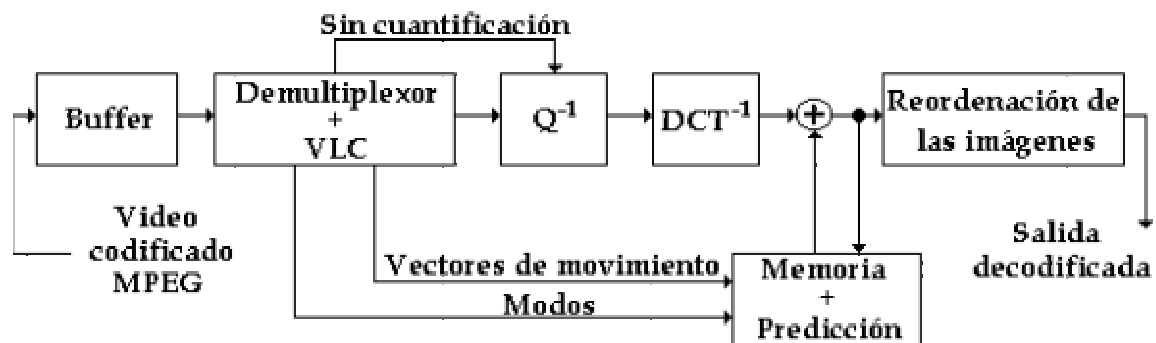


Figura 26. Esquema simplificado del decodificador MPEG-1

La ventaja del decodificador, es que no tiene que encargarse de realizar la estimación de movimiento, y su diseño resulta más simple, cumpliéndose uno de los objetivos principales planteados durante la definición de la norma.

Trama de sistema MPEG-1

Además de la multiplexación de los diferentes canales, la capa de sistema incluye información suficiente para llevar a cabo las siguientes funciones:

- 1- Presentación sincronizada de los diferentes canales.
- 2- Manejo de los buffers de los decodificadores de audio y video.

- 3- Arranque de los decodificadores.
- 4- Acceso aleatorio dentro de la trama.
- 5- Identificación de tiempo.

Sincronización y manejo de buffers:

Estas funciones se llevan a cabo mediante marcas temporales (time-stamps) incluidas dentro de la trama.

Estas marcas se denominan:

- PTS: Presentation Time Stamp.
- DTS: Decoding Time Stamp.
- SCR: Source Clock Reference.

Se miden en unidades del reloj del sistema (System Time Clock –STC- de 90 kHz). A partir de los SCR se recupera la base de tiempos del sistema de 90 kHz.

Los decodificadores de audio y vídeo utilizan los DTS y PTS para determinar el momento en que deben llevar a cabo la decodificación.

La recomendación MPEG define un decodificador ideal llamado STD (System Target Decoder) utilizado como modelo de referencia. Una trama de sistema es correcta si puede decodificarse sin errores en el STD.

Estructura de la trama:

Está formada por una o más tramas elementales de audio, vídeo y datos multiplexadas.

Cada una de las tramas de audio y vídeo están formadas por "unidades de acceso" (UA) que contienen la información comprimida de una "unidad de presentación"(UP).

- UP vídeo: imagen (frame).

- UP audio: grupo de muestras (audio frame).

Cada trama elemental se divide en packets. Un packet incluye:

- Datos codificados.
- Cabecera formada por un código de identificación del packet (Packet Start Code) además de información de sincronización (time-stamps PTS y DTS).

Los packets se organizan en packs. Cada pack está formado por:

- Cabecera: Contiene un código de identificación del pack (Pack Start Code), un time-stamp SCR e información de velocidad binaria del pack.
- Cero o más packets de tramas individuales.

La sincronización del sistema se consigue mediante los SCR, DTS y PTS. La base de tiempos del sistema se recupera a partir de los SCR.

A continuación, los decodificadores de vídeo y audio utilizan los DTS y PTS contenidos en los packets para determinar el momento en que deben llevar a cabo la decodificación y la presentación de las UP de audio y vídeo.

3.2.6 Trama de vídeo MPEG-1

El algoritmo de codificación de vídeo MPEG es asimétrico ya que la mayor parte de la carga de procesamiento recae en el codificador, simplificando considerablemente la complejidad del decodificador. Se alcanza una alta tasa de compresión a costa de unas pérdidas de calidad aceptables basándose en dos técnicas de codificación:

1.- Reducción de la redundancia temporal.

Se definen 3 tipos de cuadros de imagen: Intra (I), predictivos (P) y bidireccionales o interpolados (B).

Los cuadros I son codificados sin referencia temporal alguna a cuadros anteriores o posteriores. Son puntos de acceso a búsquedas aleatorias. Los cuadros P son

codificados más eficientemente empleando predicción por compensación de movimiento sobre un cuadro I o P anterior, pudiendo ser referencia para futuros cuadros P.

Los cuadros B son los que presentan la mayor tasa de compresión. Emplean compensación de movimiento sobre cuadros I o P, tanto pasados como futuros (predicción bidireccional). No pueden emplearse como referencia.

Las compensaciones de movimiento se realizan sobre pequeñas zonas de imagen de 16x16 pixels (macrobloques), buscando en cuadros de referencias anteriores y posteriores el mismo patrón de la zona de imagen en cuestión y proporcionando finalmente los vectores de movimiento detectado.

2.- Reducción de la redundancia espacial.

La imagen diferencia entre la entrante y la predicha es codificada empleando la técnica de la transformada discreta del coseno (DCT) a nivel de macrobloque.

Es una transformación espacio-frecuencia que concentra gran parte de la información en los coeficientes de baja frecuencia.

A continuación se aplica un proceso de cuantificación, con pérdidas de información controladas por el nivel de llenado de los buffers de salida de datos codificados, y aplicado en menor medida a las zonas de baja frecuencia, en las que el ojo es más sensible.

Los coeficientes resultantes son transformados en otros de manera que los eventos más probables son codificados con un menor número de bits (codificación entrópica).

El modelo de sistema de codificación de vídeo es el siguiente:

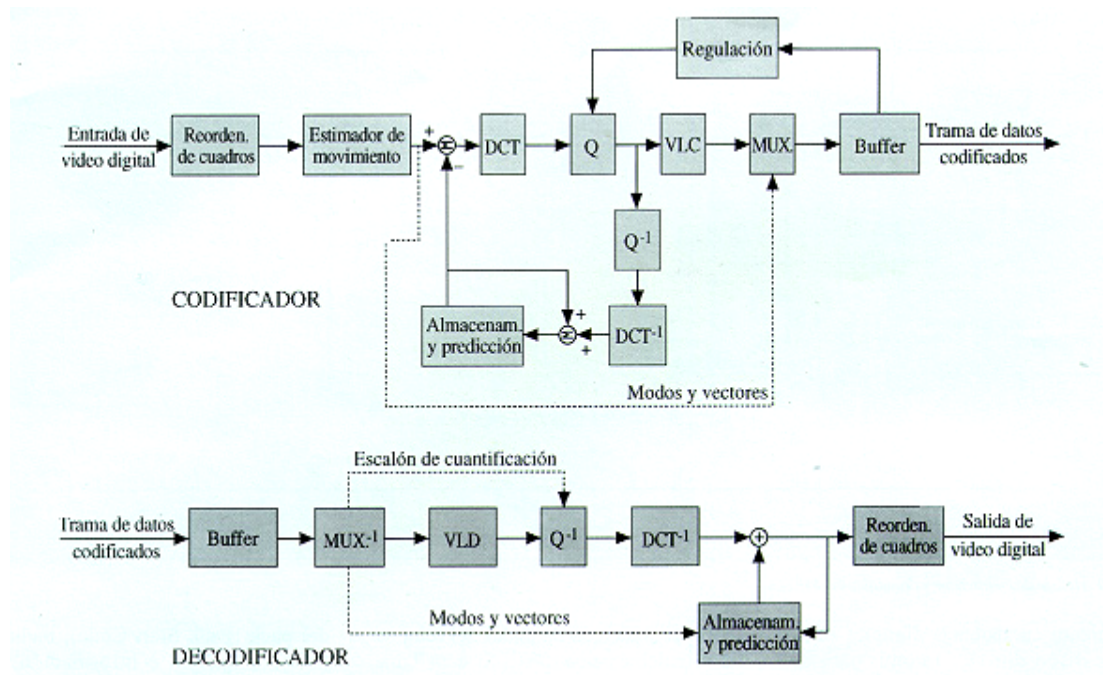


Fig. 27. modelo de sistema de codificación de vídeo.

En el codificador, la primera operación a realizar es la provisión de imágenes de vídeo con la resolución deseada y muestreando por separado las componentes de luma y de croma.

A continuación la secuencia de imágenes se reordena para permitir la codificación previa de las referencias futuras (I o P) de los cuadros B.

I1 B2 B3 P4 B5 B6 P7 B8 B9 I10 B11 B12...-> I1 P4 B2 B3 P7 B5 B6 I10 B8 B9 P13 B11...

A continuación se hace la estimación del movimiento, el cual genera un conjunto de vectores y una imagen estimada.

La diferencia entre el cuadro original y el estimado (error de estimación) es transformada mediante la DCT y posteriormente cuantificada (Q). Por último sufre un proceso de codificación entrópica antes de pasar al buffer de salida. La diferencia entre el nivel de

llenado y el de vaciado de esta buffer regula el tamaño del escalón de cuantificación por lo que también regula el factor de compresión. [6]

3.3 COMPRESIÓN DE VIDEO EN EL ESTANDAR MPEG-2

El interés en comprimir la información de la señal de video se aprecia no sólo en la reducción de la tasa binaria para su transmisión (como ha quedado ya demostrado) sino también en posibilitar el tratamiento y almacenamiento de un volumen de información tan grande.

Este interés ha motivado el desarrollo de técnicas sofisticadas de compresión que han derivado en varios estándares internacionales. Los estándares mas empleados son una familia desarrollada por el organismo llamado *Moving Pictures Expert Group* (MPEG).

La tarea básica de cualquier método MPEG es tomar las señales de audio y video y convertirlas en paquetes de información digital, de forma que puedan ser enviadas a través de redes de comunicaciones con mayor eficiencia de recursos. MPEG comprime las señales de audio y video, desechando gran parte de la información redundante de las mismas, consumiendo menos ancho de banda y manteniendo la calidad de la transmisión desde la generación de la señal hasta su descodificación y representación en el destino. MPEG determina la estructura que debe poseer la información de video digital, audio y datos asociados.

En TV digital se ha adoptado el estándar MPEG-2, que es la continuación del primer estándar emitido MPEG-1. La codificación establecida en el MPEG-1 reducía los estándares NTSC, PAL y SECAM a un formato único con muestreo 4:2:0. La aplicación fundamental de dicho estándar es el CD interactivo, con una tasa binaria de 1.5 Mbits/s. Anteriormente ya se había desarrollado un estándar para compresión de imágenes fijas (fotografía electrónica) conocido como JPEG (del *Joint Photographies Expert Group*).

MPEG-2 ésta optimizado para transmisión de TV porque consigue velocidades entre 1.5 y 6 Mbits/s con baja degradación en la calidad de imagen. Además, soporta diferentes relaciones de aspecto (4:3, 16:9), formatos de video (barrido entrelazado y progresivo), mejoras de la señal, etc.

La compresión de la información de video se realiza desde dos puntos de vista: *espacial* y *temporal*. El primero explota la existencia de información redundante dentro de una imagen fija cualquiera (elimina redundancias en zonas uniformes o con poco detalle) y la poca sensibilidad del ojo humano al color (elimina información no perceptible). El segundo aspecto de la compresión se basa en la alta correlación existente entre imágenes sucesivas. Combinando ambas filosofías se consigue un nivel de compresión muy alto y, en función del nivel de calidad exigido, se trata de tener una pérdida de información imperceptible. Las ideas básicas de ambas compresiones se exponen a continuación.

Compresión espacial

1. *Eliminación de información no perceptible*: Al igual que se hacía en los sistemas de TV analógicos para codificar la información de color, tanto los estándares MPEG como el JPEG aprovechan las características del sistema visual humano para eliminar aquella parte de la información que no podemos distinguir, como son las altas frecuencias (detalles) en color. Esto podrá hacerse en función del servicio final de las imágenes. En TV digital de calidad básica o convencional se elige un muestreo 4:2:0, que submuestra la crominancia.
2. *Eliminación de información redundante*: Cada píxel de la imagen tendrá la información de las tres componentes muestreadas (Y, CR y CB). Para esta compresión las tramas de video se dividen en bloques de 8×8 píxeles y, a su vez, cuatro bloques forman un macrobloque de 16×16 píxeles.

Los bloques se transforman al dominio de la frecuencia espacial aplicándole la transformada discreta del coseno (DCT) a cada uno de ellos. Esta transformada concentra la mayor parte de la información en un reducido número de coeficientes porque los bloques suelen ser bastante homogéneos (predominan las bajas frecuencias). Sobre estos coeficientes puede aplicarse una cuantificación y se eliminan los menos significativos, con lo que se reduce considerablemente la información a transmitir. El proceso se presenta de forma esquemática en la Figura 28.

En el destino el descodificador realiza una DCT inversa, restaurando los valores originales de forma muy aproximada.

Compresión temporal.

La compresión espacial reduce información en cada campo. Además, se sabe que los campos e imágenes sucesivos apenas cambian. Por tanto, se reduciría la cantidad de información a enviar si no se repitiera aquella parte de la información de la imagen que es común. Este hecho se explota empleando dos mecanismos diferentes.

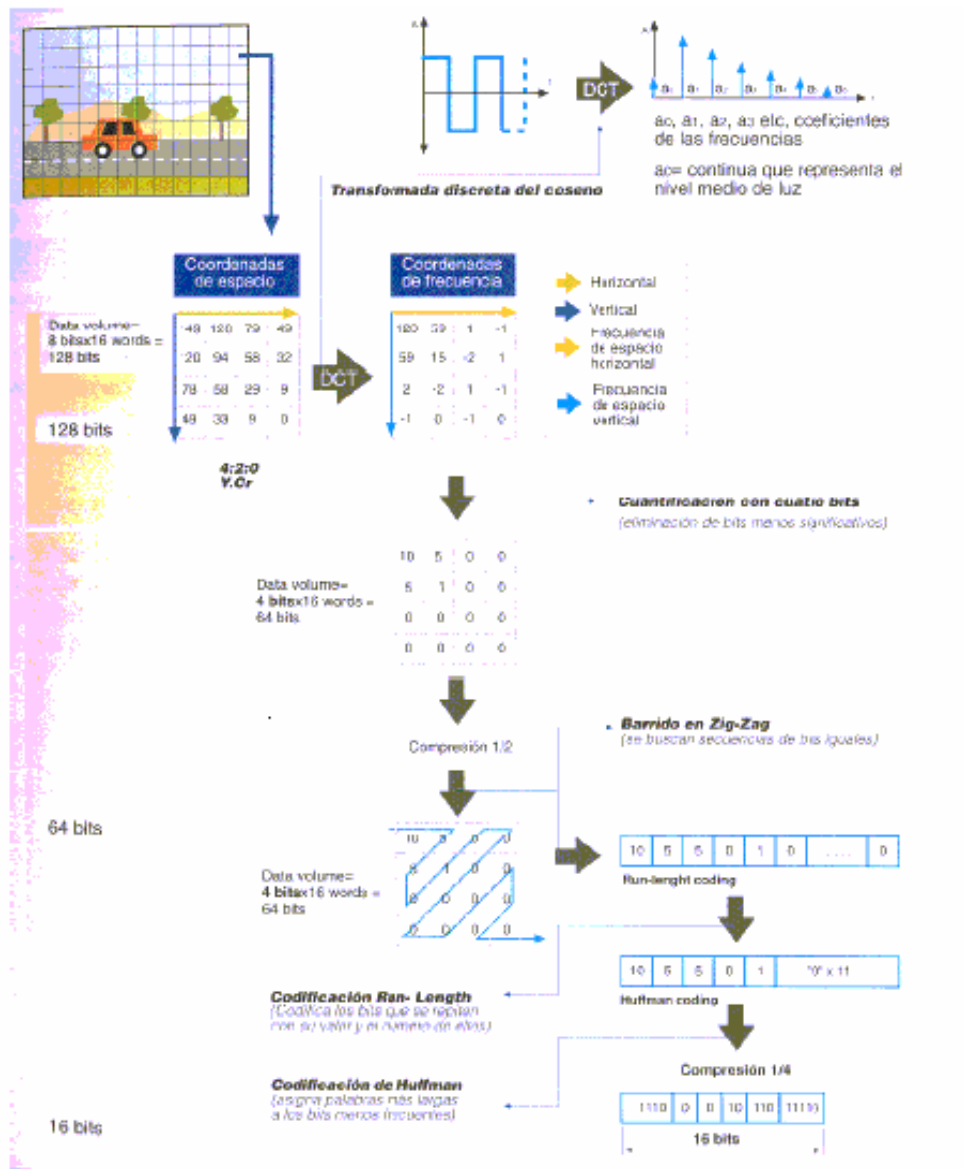


Fig. 28 ejemplo del uso de la DCT en la compresión espacial.

1. Los bloques y macrobloques que no cambian en campos sucesivos sólo se envían una vez (con cierta sincronización).
2. Basándonos en la uniformidad de los movimientos en la naturaleza, que los hace predecibles si no hay cambios bruscos de dirección, podemos reducir la información empleando los llamados *vectores de movimiento*. Esta técnica consiste en asociar a los bloques o macrobloques que se desplazan entre imágenes un vector que represente el desplazamiento sufrido.

En MPEG-2 se transmiten los vectores de movimiento y unas “imágenes” (comprimidas usando la DCT) que corresponden a la diferencia entre la predicción del movimiento (que nos darían esos vectores) y la imagen verdadera. Si los vectores están bien calculados, la imagen diferencia contiene poca información y la compresión es alta.

Para llevar a la práctica las compresiones anteriormente señaladas se emplean cuatro tipos de imágenes en MPEG-2:

Imágenes intraframe, I Se codifican como si fueran imágenes fijas utilizando la norma JPEG. Para descodificar una imagen de este tipo no hacen falta otras imágenes de la secuencia, sólo ella misma. Alcanzan una compresión moderada al explotar solamente la redundancia espacial. Su principal utilidad es que constituyen *puntos de acceso* a la trama binaria ya que permiten al descodificador iniciar su tarea sin necesidad de referenciarse a imágenes previas.

Imágenes interframe causales o imágenes predictivas, P Se codifican empleando algoritmos de predicción a partir de una imagen anterior, que puede ser I o P. Al eliminar redundancia espacial y temporal su tasa de compresión es mayor. Para descodificarlas en destino se necesita, además de ella misma, la I o P anterior.

Imágenes interframe bidireccionales, B Se codifican empleando la I o P anterior y la I o P siguiente. En ellas se alcanza la mayor tasa de compresión. Para permitir la predicción “hacia atrás” a partir de imágenes futuras, el codificador debe ordenar las imágenes de forma que la imágenes B se envíen después de las pasadas o futuras a las que estén referenciadas.

Imágenes intraframe de baja resolución, D Comparten características con las imágenes I, pero poseen menor resolución. Se emplean sólo en aplicaciones con bajos requerimientos de calidad, como el avance rápido en un magnetoscopio reproductor.

1.3.4 Perfiles y niveles en MPEG-2

El estándar MPEG-2 se ha definido como una norma flexible, capaz de soportar o ser aplicado a un gran número de servicios. Con este objetivo se han desarrollado también diferentes algoritmos para abarcar todas las aplicaciones concretas conocidas, desde el video en baja definición (para videoconferencia) a la TV de alta definición, siempre dentro del mismo estándar. La implementación final de los codificadores y descodificadores no se ha introducido en la norma, sino sólo lo que deben hacer, dejando así libertad a los mercados para el desarrollo de la tecnología.

Sin entrar en excesivo detalle, la jerarquía de las características o especificaciones para cada aplicación se compone de *niveles* y *perfiles*:

Nivel	Resolución horizontal (muestras/línea)	Resolución vertical (líneas/imagen)	Resolución temporal (imágenes/s)
Bajo	352	288	30
Principal	720	576	30
Alto 1440	1440	1152	60
Alto	1920	1152	60

TABLA 2 CARACTERISTICAS DE LOS PERFILES DE MPEG.

Perfil	Algoritmos
Simple	Video entrelazado, 4:2:0, sólo imágenes I y P
Principal	Añade imágenes tipo B
Escalable en S/N	Algoritmo de mejora de la S/N de la imagen
Escalable espacialmente	Algoritmo para obtener imágenes con diferentes resoluciones
Alto	Codificación del color en 4:2:2

TABLA 3 CARACTERISTICAS DE LOS PERFILES DE MPEG-2.

Niveles Rangos de parámetros soportados por la implementación de cada servicio. Fundamentalmente tratan la resolución temporal y espacial soportada. Los niveles bajos se corresponden a menores tasas binarias y los altos a mayores tasas.

Perfiles De uno a otro se introducen algoritmos para mejorar la codificación.

Las tablas 2,3 y 4 resumen las características de los diferentes perfiles y niveles, así como las combinaciones de ellos que se contemplan en la norma (no todas son posibles).

En la Figura 29 se presenta de forma esquemática el diagrama de bloques de un codificador.

A continuación también se resume la tarea de cada elemento del diagrama:

Reordenación de imágenes Coloca las imágenes en la secuencia necesaria para su transmisión (recordemos que la imágenes I y P de las que dependen las B deben llegar antes aunque sean posteriores en la señal original)

	Alto	Esc. espacial	Esc. S/N	Principal	Simple
Alto	100			80	
Alto 1440	80	60		60	
Principal	20		15	15	15
Bajo			4	4	

Tabla 4. Tasas binarias (Mbit/s) de las combinaciones posibles de nivel y perfil en MPEG-2.

Estimación de movimiento Cálculo de los vectores de movimiento entre dos imágenes.

Estos serán transmitidos y también empleados para calcular la predicción de la imagen actual

Almacén de imágenes y predicción Cálculo de la predicción de la imagen actual a partir de la imagen anterior (almacenada) y de los vectores de movimiento calculados para la imagen actual. La predicción hallada se resta a la imagen actual dando lugar al error de predicción (e) que es lo que se transmite

DCT Obtiene la transformada discreta de coseno del error de predicción

Cuantificación Cuantifica de forma uniforme la DCT hallada. El resultado, que es el error de predicción cuantificado, se pasa al siguiente elemento y se emplea también para hallar el error de predicción como se haría en recepción.

Codificación de longitud variable (CLV) Este bloque asigna un código de Huffman a los datos cuantificados. Este tipo de código emplea más bits para los símbolos que menos aparecen y viceversa. Con ello se disminuye la tasa binaria final

Buffer La salida del CLV se multiplexa con otros programas. Como la tasa binaria que procede del multiplexor (y del CLV) es variable, se necesita este buffer para entregar a la salida un régimen binario constante.

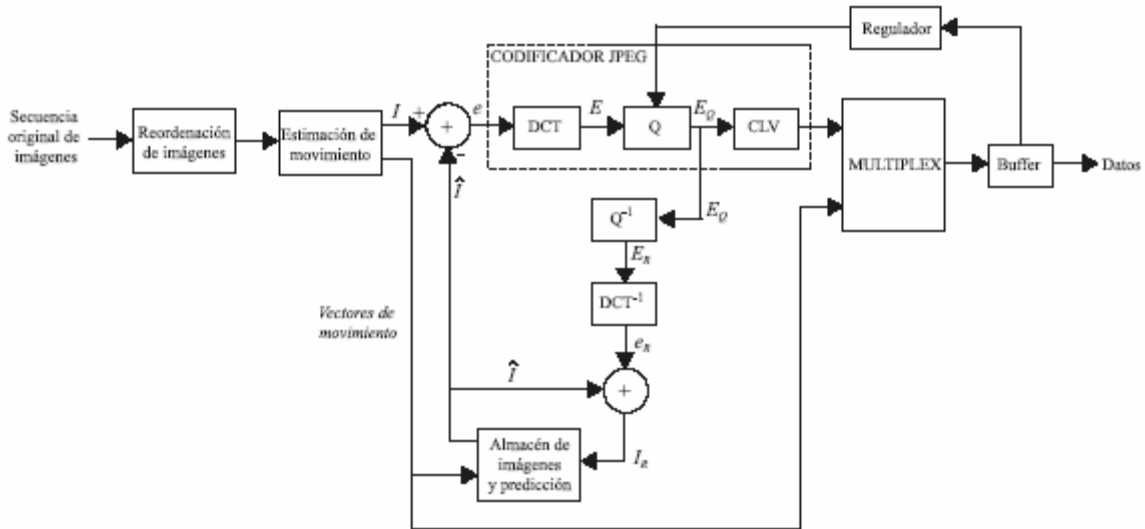


Fig. 29. diagrama de bloques de un calificador MPEG.

3.3.1 ESTÁNDAR MPEG-2 (APLICACIONES BROADCAST)

MPEG-2 puede describirse como una "caja de herramientas" de compresión más compleja que MPEG-1, por lo tanto, también puede ser considerada como una unidad superior: en efecto, toma todas las herramientas anteriores y le añade otras. Además, la norma prevé la compatibilidad ascendente, lo que significa que un decodificador MPEG-2 deberá decodificar trenes binarios elementales de la norma MPEG-1.

3.3.1.1 PERFILES Y NIVELES MPEG-2

MPEG-2 se puede utilizar en un vasto rango de aplicaciones, requiriendo diferentes grados de complejidad y desempeño.

Para un propósito práctico el estándar MPEG-2 es dividido en perfiles y cada perfil es subdividido en niveles (Ver la tabla 5). Un perfil es básicamente el grado de complejidad esperada en la codificación, mientras que un nivel describe el tamaño de la imagen, la resolución de esta o la velocidad de transferencia de bits usada en ese perfil. En principio, hay 24 combinaciones posibles, pero no todas están definidas. Un codificador MPEG cuando entrega un perfil y un nivel determinado, debe además ser capaz de decodificarlo a perfiles y niveles inferiores.

		PERFILES					
		Simple	Principal	4:2:2	SNR	Espacial	Alto
NIVELES	Alto		4:2:0 1920 x 1152 80Mb/s				4:2:0 o 4:2:2 1920 x 1152 100Mb/s
	Alto 1440		4:2:0 1440 x 1152 60Mb/s			4:2:0 1440 x 1152 60Mb/s	4:2:0 o 4:2:2 1440 x 1152 80Mb/s
	Principal	4:2:0 720 x 576 15Mb/s Sin B	4:2:0 720 x 576 15Mb/s	4:2:2 720 x 608 50Mb/s	4:2:0 720 x 576 15Mb/s		4:2:0 o 4:2:2 720 x 576 20 Mb/s
	Bajo		4:2:0 352 x 288 4Mb/s		4:2:0 352 x 288 4Mb/s		

Tabla 5. Niveles y perfiles de MPEG-2

Un perfil simple no soporta una codificación bidireccional y de este modo solo genera imágenes I y P. Esto reduce la tasa de compresión simplificando el codificador y el decodificador; permitiendo un sencillo hardware. Un perfil simple solamente está definido en el nivel main (principal) como (SP@ML). El perfil main (principal) corresponde actualmente al mejor compromiso entre calidad/tasa de compresión, utilizando los tres tipos de imágenes (I, P y B), a costa de un codificador y decodificador, más complejos.

Los perfiles escalables (código jerárquico) están previstos para operaciones posteriores y permitirán transmitir una imagen básica (base layer) en términos de resolución espacial (spatially scalable profile) o de cuantificación (SNR scalable profile), así como información suplementaria independiente (enhanced layer) que permite mejorar sus características, por ejemplo para transmitir la misma emisión en definición estándar y HD (High Definition), o permitir una recepción con calidad aceptable en caso de recepción difícil y de calidad óptima en buenas condiciones (por ejemplo, para la televisión digital terrestre).

La Figura 30 muestra un codificador MPEG convencional, con coeficientes de cuantificación de elevado peso; que al codificar una imagen la genera con una moderada razón señal a ruido. Después esta imagen al ser decodificada y sustraída de la imagen

original pixel a pixel da como resultado una imagen de "ruido de cuantificación". Esta imagen puede ser comprimida y transmitida como una imagen de ayuda. Un simple decodificador solo decodifica la imagen principal, con un flujo de bit con ruido de cuantificación, pero un decodificador más complejo puede decodificar ambas imágenes con diferentes flujos de bits y combinarlos para producir una imagen con bajo ruido. Este es el principio del perfil SNR escalable.

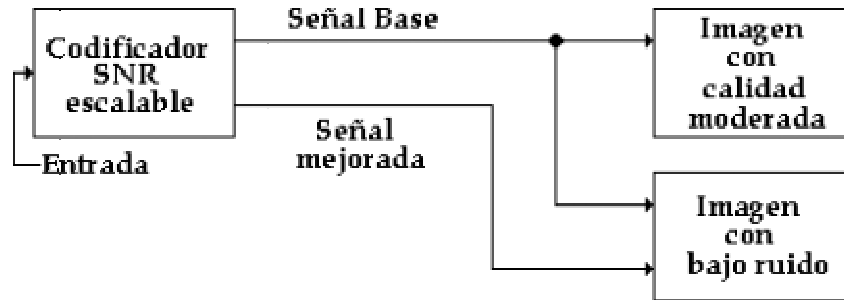


Figura 30. Codificación escalable SNR

Como otra alternativa, la Figura. 31 muestra que por solo codificar las frecuencias espaciales bajas en imágenes HDTV, parte del flujo de bits puede ser reconstruido por un receptor con codificador para SDTV (Standard Definition TeleVision). Si una imagen de baja definición es localmente decodificada y sustraída de la imagen original, se produce entonces una imagen de "realce de definición", esta imagen puede ser codificada en una señal de ayuda. Un decodificador de forma conveniente podría combinar las señales principales y de ayuda para recrear la imagen HDTV. Este es el principio del perfil de escalabilidad espacial.

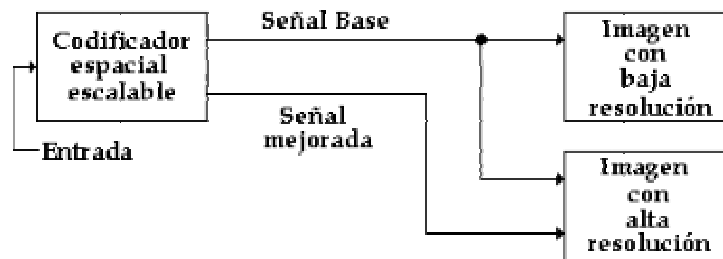


Figura 31. Codificador espacial escalable

El perfil high (alto) soporta tanto el SNR y la escalabilidad espacial como también la opción de muestreo 4:2:2.

El perfil 4:2:2 se ha desarrollado para proveer compatibilidad con los equipos de producción digital de televisión existentes. Este perfil admite trabajar con 4:2:2 sin requerir una complejidad adicional si se usa en el perfil high. Por ejemplo, un decodificador HP@ML debe soportar escalabilidad SNR que no es requerida en la producción de televisión digital. El perfil 4:2:2 tiene la misma libertad de escoger su estructura de GOP como en otros perfiles, pero en la práctica este usa comúnmente GOPs cortos de edición sencilla. La operación 4:2:2 requiere una mayor velocidad en la transmisión del bit que una operación 4:2:0, y el uso de pequeños GOPs requiere también de mayores velocidades de transferencia de bits para proporcionar calidad en sus imágenes.

- El nivel low (bajo) corresponde a la resolución SIF utilizada en el MPEG-1.
- El nivel main (principal) corresponde a la resolución 4:2:0 "normal" (de hasta 720 pixeles x 576 líneas).
- El nivel high-1440 (alto-1440) está destinado a la HDTV (de hasta 1440 pixeles x 1152 líneas).
- El nivel high (alto) está optimizado para la HDTV (de hasta 1920 pixeles x 1152 líneas).

Según el compromiso de calidad/flujo de bits perseguido y la naturaleza de las imágenes, el flujo de bits estará comprendido entre los 4 Mbits/s (calidad equivalente a la de una imagen codificada en PAL o SECAM) y los 9 Mbits/s (calidad próxima a la de una imagen de estudio CC1R-601).

Todo el proceso de codificación de las imágenes animadas descrito en el capítulo anterior para MPEG-1 se aplica a MPEG-2 (MP@ML), especialmente la jerarquía de capas (desde el bloque hasta la secuencia de la Figura 32).

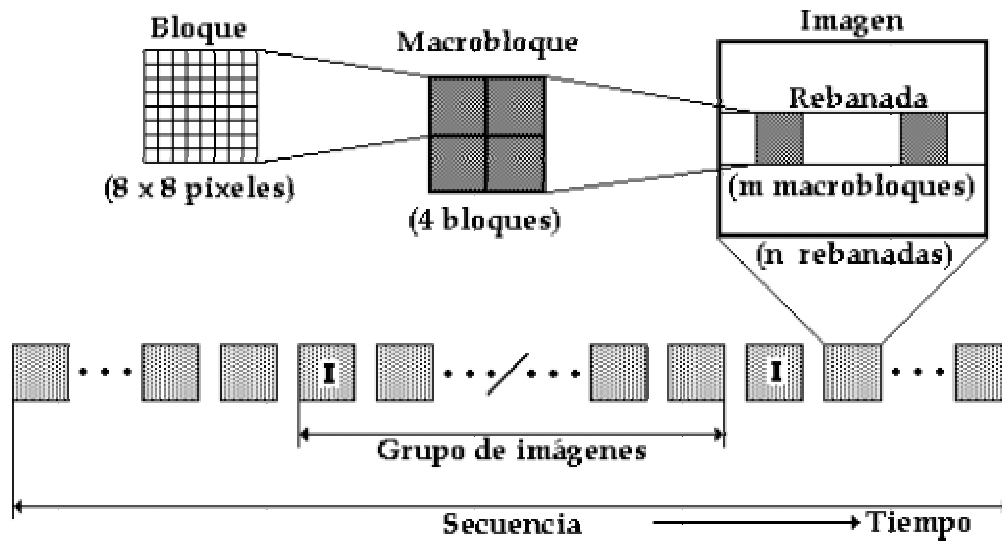


Figura 32. Jerarquía de capas de la secuencia

Una diferencia que hay que destacar para las slices, es que en MPEG-2 no necesariamente abarcan toda la imagen, y además deben estar compuestas únicamente de macrobloques contiguos situados en la misma línea horizontal. Ver Figura. 33.

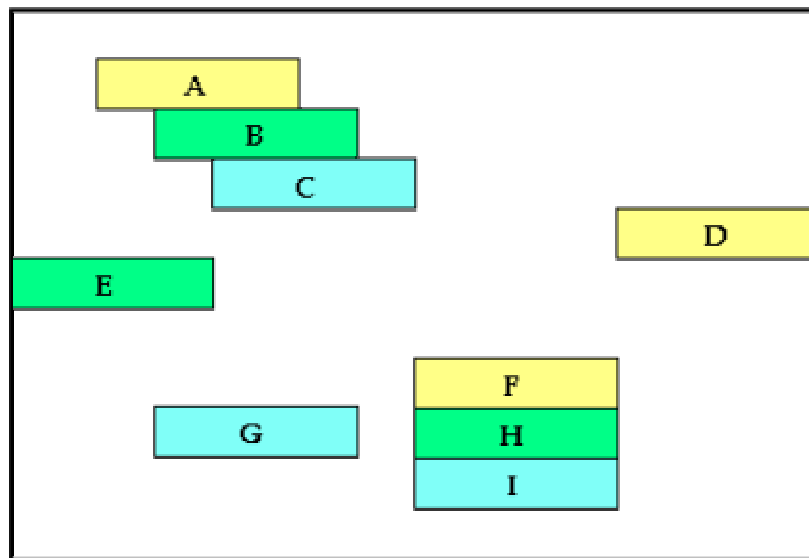


Figura 33. Rebanadas (slices) en MPEG-2

La principal novedad con respecto a MPEG-1, además de los perfiles y niveles, provienen del tratamiento de las imágenes entrelazadas.

3.3.2 MODOS DE PREDICCIÓN ESPECÍFICOS EN MPEG-2 (IMÁGENES ENTRELAZADAS)

Dependiendo del trabajo a realizar, estas pueden ser tratadas de manera diferente según la importancia de los movimientos entre los dos campos de una misma imagen (los casos extremos son, por un lado, cuando se transmiten películas cinematográficas por televisión "telecine" donde no hay movimiento entre los dos campos de TV, puesto que proceden de la exploración del mismo fotograma de la película, y por otro lado, las imágenes de video de acontecimientos deportivos, donde puede haber importantes movimientos entre los dos campos de una imagen).

La Figura 34 representa la secuencia temporal de la posición vertical de las líneas de los campos sucesivos en un sistema entrelazado.

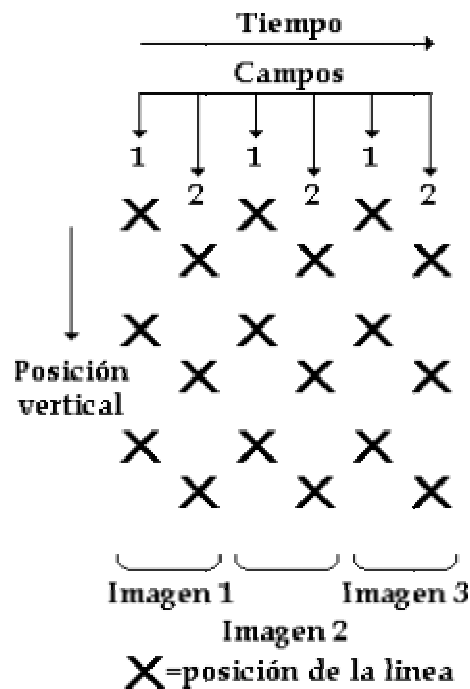


Figura 34. Posición de las líneas de campos sucesivos en un sistema entrelazado.

Para la codificación Intra de las imágenes entrelazadas, MPEG-2 permite elegir entre dos estructuras de imágenes llamadas frame (estructura "imagen") o field (estructura "campo").

3.3.3 LA ESTRUCTURA "FRAME"

También llamada "progresiva", es apropiada para los casos donde hay poco movimiento entre dos campos sucesivos. Los bloques y macrobloques se dividen en la imagen completa (Ver Figura. 35), y la DCT se efectúa, sobre puntos verticales que distan 20 ms en el tiempo, lo que no plantea problemas si los dos campos difieren poco.

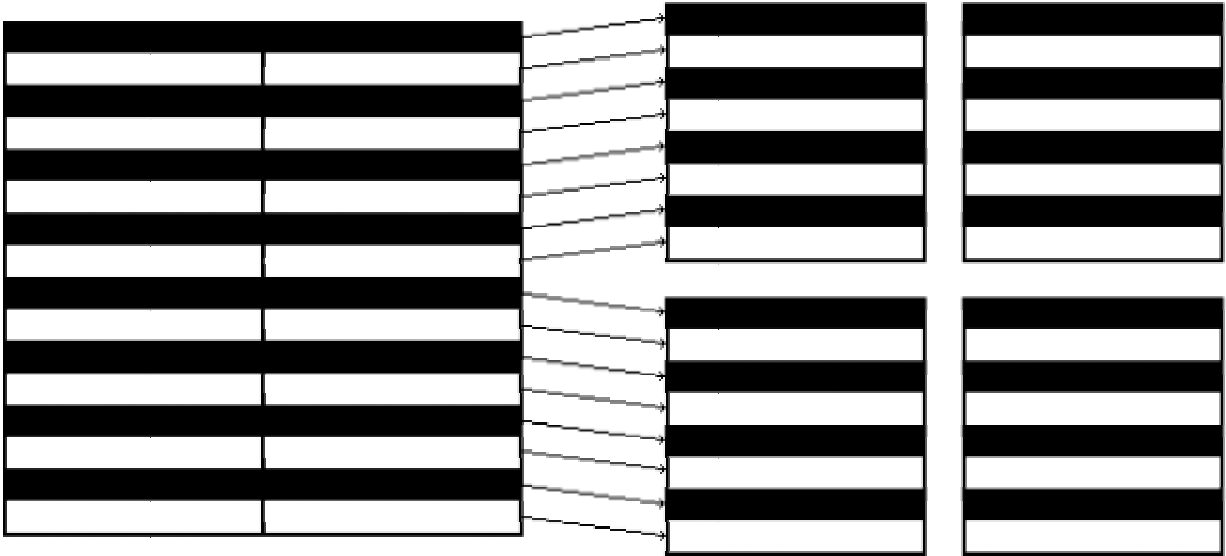


Figura 35. División de los macrobloques en bloques en modo imagen (frame)

En este caso, siempre es posible codificar los bloques de mayor animación en modo inter-campo, es decir, dividiéndoles en un campo.

3.3.4 LA ESTRUCTURA "FIELD"

También llamada "entrelazada", es preferible cuando el movimiento de un campo a otro es importante. En este caso, a fin de evitar un contenido en frecuencias verticales elevadas que reduciría la eficacia de la compresión tras efectuar la DTC, la división de los macrobloques se hace considerando cada uno de los campos como una imagen independiente en el interior del cual se toman los bloques. Ver Figura 36.

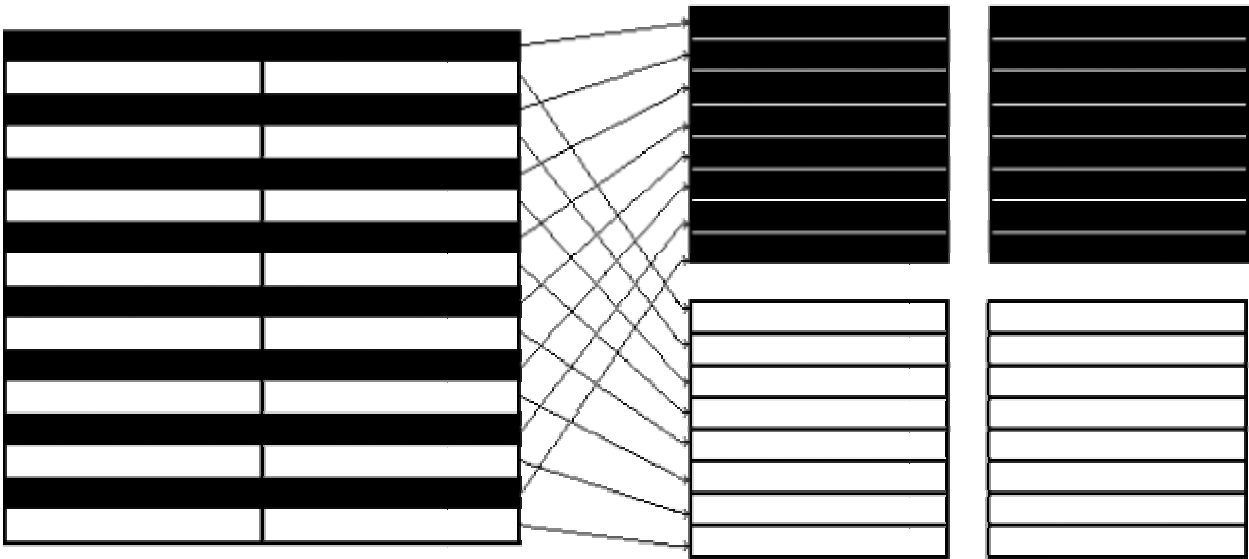


Figura 36. División de los macrobloques en bloques en modo campo (field)

En cuanto a la estimación de movimiento, también hay varios modos previstos:

Un macrobloque puede predecirse en modo "imagen", "campo" o "mixto"

- **El Modo "Frame"**

Un macrobloque formado en el campo impar sirve para predecir el bloque correspondiente del próximo campo impar, y lo mismo para los bloques del campo par, por tanto, la predicción se hace sobre un tiempo de 40 ms (2 campos).

- **El Modo "Field"**

La predicción de un bloque se efectúa a partir de un bloque del campo anterior, entonces aquí, los vectores de movimiento corresponden a un tiempo de 20ms.

- **El Modo "Mixto"**

Los bloques se predicen a partir de dos bloques que corresponden a dos campos.

entrelazada será sobre una diagonal de 67.5° . La Figura 38 muestra que esta forma de exploración, entrega primero las frecuencias espaciales verticales y luego las frecuencias espaciales horizontales.

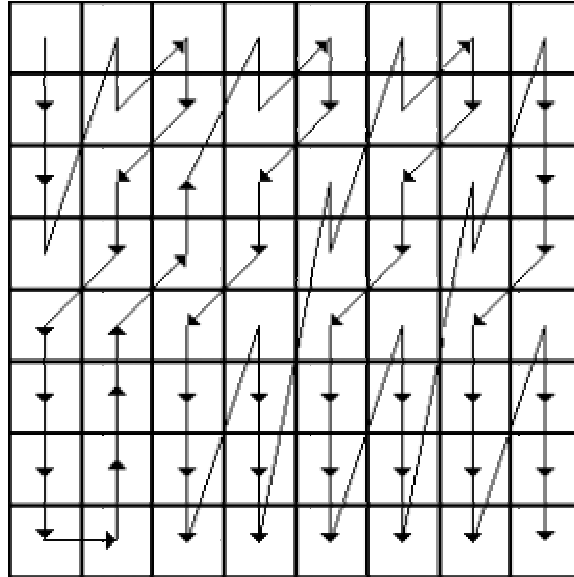


Figura 38. Exploración alternada, normalmente para campos (fields)

3.3.6 DESCRIPCIÓN DEL PROCESO DE CODIFICACIÓN MPEG-2

Al igual que MPEG-1, la norma no define explícitamente el método de codificación, sino únicamente la sintaxis que controla el tren binario a la salida del codificador, lo cual deja gran libertad a su diseñador.

El esquema de bloques MPEG-1, también se aplica al codificador MPEG-2. Ver Figura 39.

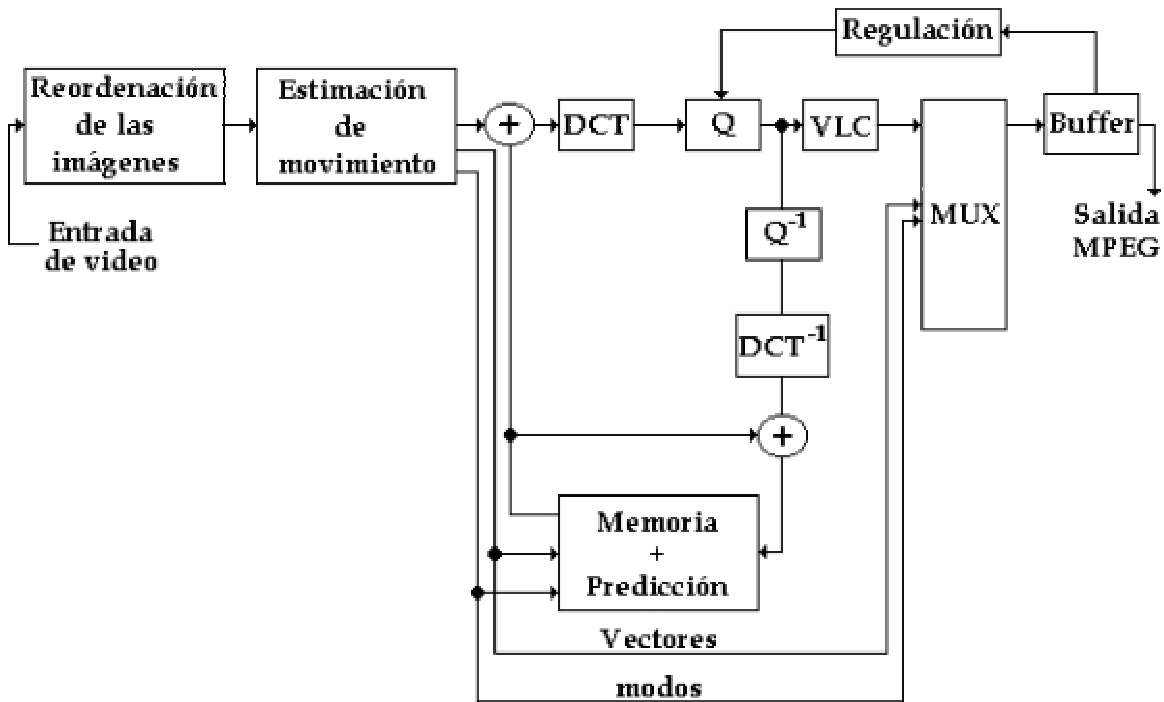


Figura 39. Esquema simplificado del codificador MPEG-2

A partir de la imagen digitalizada en formato 4:2:0 (caso del main profile), el codificador elige para cada imagen su tipo (I, P o B) y si esta debe ser codificada en modo frame (imagen) o field (campo). El codificador a continuación debe estimar los vectores de movimiento para cada macrobloque de 16x16 píxeles. El número de vectores depende del tipo de imagen y del modo de codificación escogido para cada bloque.

En el caso más general, donde el codificador es capaz de generar imágenes B (bidireccionales), deberá reordenar las imágenes antes de la codificación y la transmisión.

La unidad básica de codificación es el macrobloque, compuesto por 4 bloques de luminancia de 8x8 píxeles y (en el caso del formato 4:2:0) de 2 bloques de crominancia (un Cr y un Cb) de 8x8 píxeles que abarcan la misma zona de la imagen.

Todos los macrobloques de la imagen se codifican secuencialmente de izquierda a derecha y de arriba abajo, eligiéndose un modo de codificación independiente para cada uno de ellos.

Una vez que se ha elegido el modo de codificación, la predicción con compensación de movimiento del contenido del bloque se hace a partir de la imagen de referencia (I o P) pasada (caso de las imágenes P) y eventualmente futura (caso de las imágenes B). La predicción se elimina de los datos reales del macrobloque, lo que da la señal de error de predicción.

En una imagen con estructura frame, el codificador deberá elegir entre efectuar la DTC en modo frame o field. Esto depende principalmente de la amplitud del movimiento entre los campos de la imagen.

La señal de error se separa inmediatamente en bloques de 8x8, a los que se aplica la DTC. Cada bloque de coeficientes resultante se cuantifica y barre en zig-zag para formar una serie de coeficientes. Seguidamente, se codifica la información auxiliar necesaria para que el decodificador pueda reconstruir el bloque (modo de codificación, vectores de movimiento, etc.), codificando los coeficientes cuantificados con ayuda de una tabla VLC (codificación Huffman).

La unidad de control de flujo supervisa el estado de ocupación de la memoria intermedia de salida, utilizando esta información como retorno para controlar el número de bits que el codificador generará para los bloques siguientes, jugando principalmente con los coeficientes de cuantificación. Se obtiene entonces a la salida del codificador un tren binario completo, ya utilizable para un decodificador.

Para aumentar la calidad de la imagen decodificada, el propio codificador almacena y decodifica (decuantificación de los coeficientes después de la DTC inversa) las imágenes I y P, como referencia para reconstruir otras imágenes obtenidas por predicción con compensación de movimiento en el decodificador, y calcula una señal de error que se añade a la señal de predicción.

3.3.7 DESCRIPCIÓN DEL PROCESO DE DECODIFICACIÓN MPEG-2

Como ya se ha dicho, la decodificación es más sencilla que la codificación, ya que no tiene que efectuar alguna estimación de movimiento, que es una de las partes más complejas del codificador.

El esquema de bloques del decodificador de la Figura 40 es el que se va a analizar para MPEG-2.

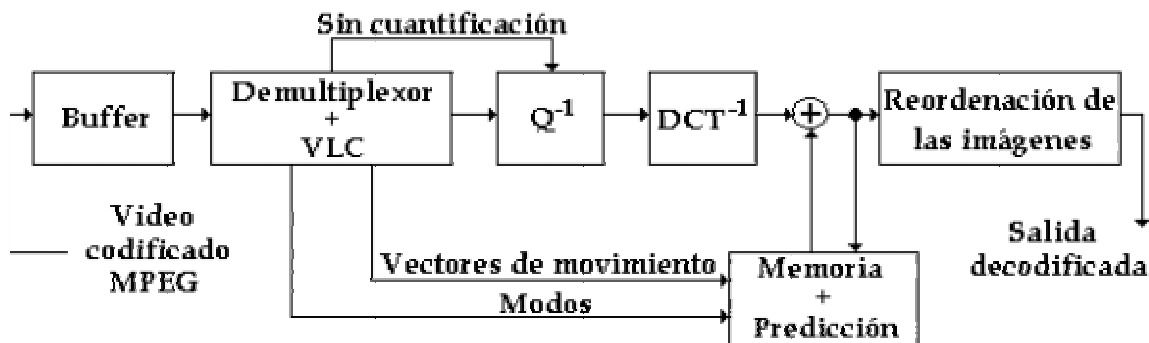


Figura 40. Esquema simplificado del decodificador MPEG-2

La memoria intermedia (buffer) de entrada recibe los datos del canal de transmisión, y el decodificador lee el tren binario hasta encontrar el principio de una imagen, su tipo (I, P o B) y su estructura (frame o field).

Empieza la decodificación con la primera imagen I, almacenándola en su memoria, así como la imagen P siguiente, para servir de referencia a las imágenes P o B que dependen de ella.

Para las imágenes I, la decodificación propiamente dicha consiste en aplicar a cada bloque la decodificación VLC, la decuantificación de los coeficientes y la transformación DTC inversa.

Para las imágenes P o B, este proceso consiste en construir la predicción de cada macrobloque a partir de su tipo, de los vectores de movimiento y de las imágenes de referencia memorizadas. El decodificador lee, decodifica y decuantifica los coeficientes

DTC del error de predicción transmitido para cada bloque de 8×8 píxeles, y, después de la transformada DTC inversa, añade el resultado a la predicción.

La reconstrucción de la imagen se efectúa cuando todos los macrobloques han sido tratados.

La última etapa de la decodificación es poner las imágenes en el orden inicial de visualización.

Como se vio anteriormente, la necesidad de memoria para el decodificador es de unas 3 imágenes (dos imágenes de referencia más la imagen en vía de reconstrucción), siendo para una imagen 4:2:0, de aproximadamente 16 Mbits.

3.3.8 CONSTITUCIÓN DEL PAQUETE DE TRANSPORTE MPEG-2

Un paquete "transport" de 188 bytes se compone de una cabecera de paquete (packet header), de 4 bytes y de una "carga útil" (payload) de 184 bytes como máximo, eventualmente precedida de un campo de adaptación (adaptation field), como se muestra en la Figura 41. La "carga útil" está formada por paquetes de trenes elementales (Packetized Elementary Stream, PES) que componen los programas de televisión transmitidos por el canal, así como cierto número de datos auxiliares que permiten al codificador no perderse por el tren de transporte MPEG-2.

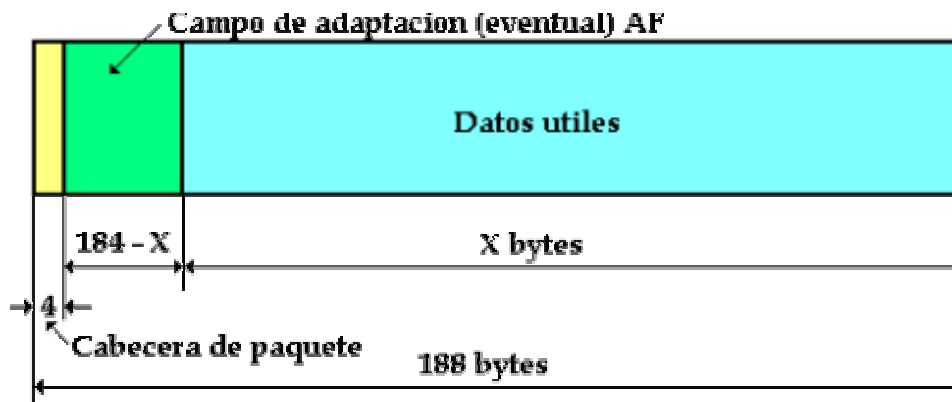


Figura 41. Constitución del paquete transport (caso general)

El formato de la cabecera del paquete de transporte se detalla en las Figuras 41, 42, 43 y el cuadro 6.

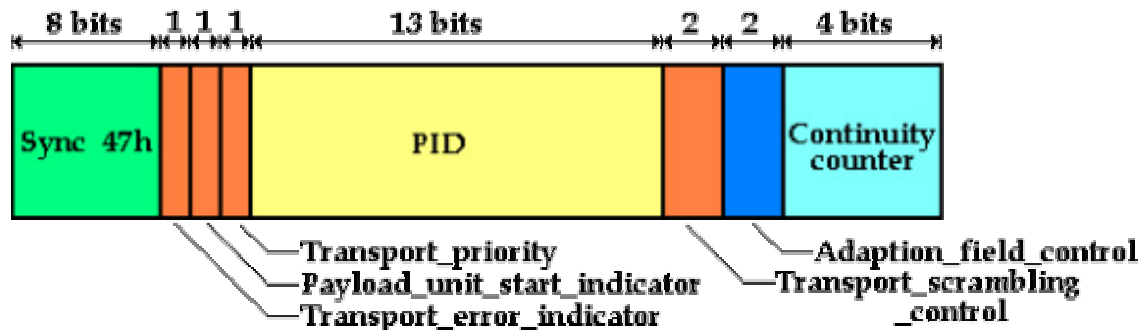


Figura 42. Detalle de la cabecera del paquete transporte

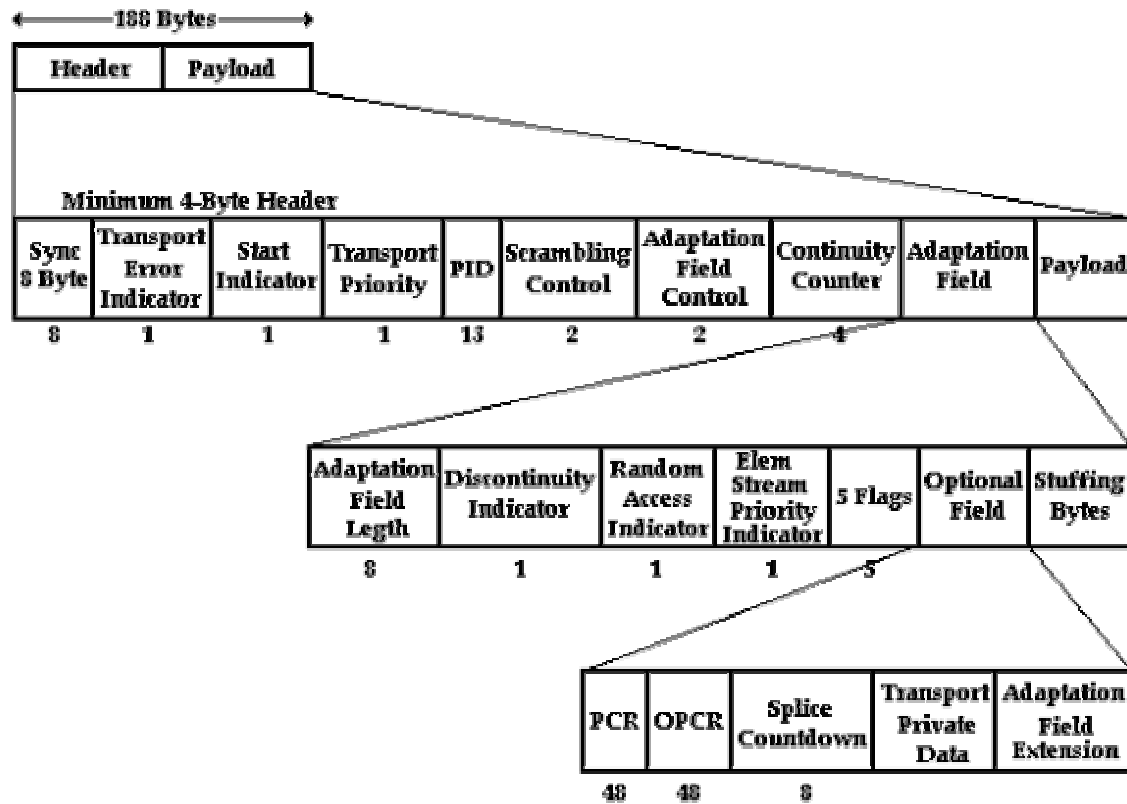


Figura 43. Estructura completa de un paquete de transporte

Campo	Definición (comentario)	N ^o de bits
sync_byte	byte de sincronización 1000 0111 (47 hex)	8
ei	transport_error_indicator (indica un error detectado mas atrás)	1
pusi	payload_unit_start_indicator (inicio de PES en el paquete)	1
tpriority	transport_priority (indicador de prioridad)	1
PID	Packet IDentifier (identificación del paquete)	13
scr_flags	transport_scrambling_flags (tipo de cifrado de transporte)	2
af	adaptation_field_flag (campo de adaptación en el paquete)	1
pf	payload_flag (datos útiles en el paquete)	1
Cc	continuity_counter (contador de continuidad entre trozos)	4

Cuadro 6. Estructura de una cabecera de paquete de transporte MPEG-2

La norma ISO/IEC 13818-1 especifica que un paquete de transporte dado sólo puede transportar datos procedentes de un sólo paquete PES, y que un paquete PES empieza obligatoriamente al principio de un paquete de transporte y se termina obligatoriamente al final de un paquete de transporte.

Debido a la longitud mucho más pequeña de los paquetes de transporte (184 bytes útiles) con respecto a los paquetes PES (por ejemplo, 2048 bytes), estos últimos deberán ser, pues, divididos en trozos de 184 bytes.

Como la longitud de los paquetes PES en general no es múltiplo de 184 bytes, el último paquete de transporte de un paquete PES deberá empezar por un campo de adaptación (Adaptation Field, AF), cuya longitud será el complemento a 184 del número de bytes que queden por transmitir para terminar este último paquete PES, como se muestra en la Figura

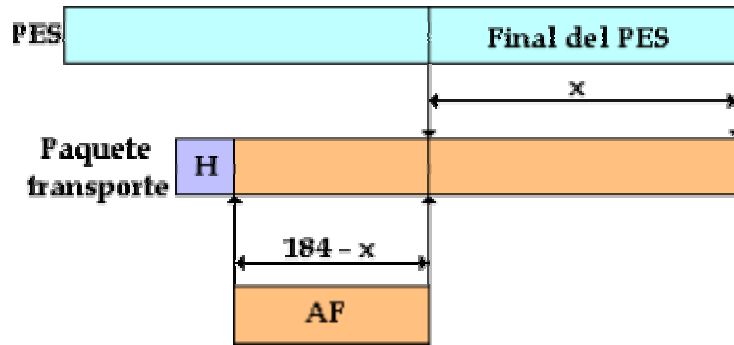


Figura 44. Constitución del último paquete de transporte de un PES

Además de esta función de relleno, el campo de adaptación se utilizara también para la transmisión del reloj de referencia del programa (Program Clock Reference, PCR), cuya cadencia de repetición mínima es de 10 por segundo, así como diversos tipos de datos opcionales.

Un paquete de transporte eventualmente podrá estar constituido únicamente por un campo de adaptación, de 184 bytes en este caso concreto (transporte de datos privados, PCR, etc.).

La Figura 45 ilustra el formato general del campo de adaptación, cuyo contenido se especifica en el Cuadro 7.

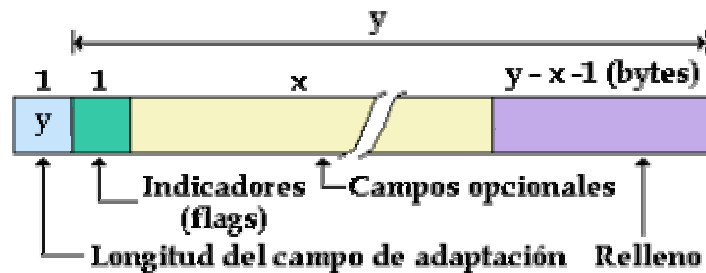


Figura 45. Detalle del campo de adaptación

Campo	Definición	Nº de bits
adaptation_field_length	longitud total en bytes, menos 1 (y)	8
flags	marcadores que indican la información transportada	8
optional_fields	campos opcionales de datos (x bytes)	$x \times 8$
stuffing	campo de relleno (y-1-x bytes de valor FF hex)	$(y-1-x) \times 8$

Cuadro 7. Constitución del campo de adaptación MPEG-2

3.3.9 ORGANIZACIÓN DEL MÚLTIPLEX DE TRANSPORTE

Como ya se ha visto, un múltiplex de transporte MPEG-2 puede transportar varios programas, cada uno compuesto por uno o varios trenes elementales de paquetes (PES).

A fin de que el decodificador pueda orientarse dentro de esta variedad de paquetes, MPEG-2 define cuatro tipos de tablas, que juntas constituyen la información específica de los programas (Program Specific Information, PSI).

3.3.9.1 Program association table (PAT). Esta tabla, cuya presencia es obligatoria, es transportada por los paquetes cuyo indicador lleve el número 0 (PID= 0x0000). Su función es la de indicar, para cada programa transportado por el múltiplex de transporte, la relación entre el número de programa (de 0 a 65535) y el PID de los paquetes que transportan una tabla que indica los datos que identifican el programa (Program Map Table, PMT). La PAT siempre se transmite "de forma clara", aunque todos los programas estén cifrados.

3.3.9.2 Program map table (PMT). Existe una tabla presente por cada programa en el múltiplex. Principalmente indica (sin cifrar) los PID de los trenes elementales que constituyen el programa y, opcionalmente puede ser cifrada.

Puede ser transportada por paquetes de PID arbitraria, definidos por el emisor en la PAT (excepto 0x0000 y 0x0001).

3.3.9.3 Conditional access table (CAT). Esta tabla debe estar presente a partir del momento en que al menos un programa sea de acceso condicional.

Se transporta por los paquetes de PID=0x0001 e indica los PID de los paquetes que transportan los EMM por uno o varios sistemas de control de acceso (una de las informaciones necesarias para el descifrado de los programas de acceso condicional).

La norma MPEG-2 prevé la posibilidad de añadir informaciones complementarias opcionales (DVB-SI, Service Information) que permiten que el receptor se configure automáticamente, al tiempo que ofrece al usuario la posibilidad de "navegar" entre los numerosos servicios que se ofrecen.

3.3.10 TABLAS PRINCIPALES DE DVB-SI

3.3.10.1 Network information table (NIT). Esta tabla de información de red, cuyo contenido no está definido por MPEG, transporta informaciones específicas relativas a una red formada por varios canales físicos (por tanto, varios trenes de transporte independientes), tales como las frecuencias, los números de los canales de red utilizados en la configuración del receptor.

Esta tabla, si está presente, constituye por definición el programa 0 del múltiplex y es considerada como datos privados.

3.3.10.2 Service description table (SDT). Esta tabla lista los nombres y otros parámetros asociados a cada servicio de un mismo múltiplex.

3.3.10.3 Event information table (EIT). Esta tabla se utiliza para transmitir información relativa a los acontecimientos en curso o que vendrán en el múltiplex MPEG recibido en la actualidad y, ocasionalmente, sobre unos múltiplex MPEG.

3.3.10.4 Time and date table (TDT). Esta tabla se utiliza para poner en hora el reloj interno del receptor (decodificador).

3.3.11 TABLAS OPCIONALES DE DVB-SI

3.3.11.1 Bouquet association table (BAT). Esta tabla se utiliza para agrupar la presentación al usuario de ramilletes (bouquets) de servicios asociados. Un servicio particular puede pertenecer a uno o varios ramilletes.

3.3.11.2 Running status table (RST). Esta tabla se transmite para la actualización rápida de uno o varios acontecimientos, una sola vez, en el momento en que se produce un cambio (a diferencia de otras tablas que se tramiten de forma repetitiva).

3.3.11.3 Stuffing tables (ST). Estas tablas de "relleno" se utilizan, por ejemplo para invalidar tablas que ya no sirven.

La frecuencia de repetición de las tablas no viene impuesta por la norma, no obstante, debe ser suficiente (de 10 a 50 veces por segundo) para que permita que el decodificador acceda con suficiente rapidez al programa buscado, especialmente, cuando se pone en marcha.

La información de algunas tablas (PAT, PMT) podrán eventualmente ser almacenadas en una memoria no volátil del decodificador para permitir un acceso más rápido tras la puesta en marcha. Posteriormente, esta memoria deberá actualizarse cuando se produzcan modificaciones esporádicas en el contenido del múltiplex.

Cada tabla está constituida, según su importancia, por una o varias secciones (256 como máximo, con una longitud máxima de 1024 bytes, salvo para la tabla private que puede alcanzar los 4026 bytes).

Los cuadros 8 a 11 muestran la estructura y los diferentes campos de las secciones de las 4 categorías de tablas (PAT, CAT, PMT Y PRIVATE) definidas por la norma MPEG-2.

Campo (PAT)	Comentario	N ^o de bits
table_id (00)	siempre a 0 para Program Allocation Table	8
section_syntax_indicator	siempre a "1"	1
"0"		1
reserved		2
section_length	2 bits de mayor peso a "0", valor máx. 1021	12
transport_stream_id	identificación del tren en una red	16
reserved		2
version_number	incrementando cada modificación del PAT	5
current_next_indicator	"1"=PAT actual, "0"=próximo	1
section_number	número de la sección en curso (1 ^a =00)	8
last_section_number	número de la última sección (N _{tot} -1)	8
program_number 0 ⁽¹⁾	Prg 0=tabla de información de red (NIT)	16
reserved		3
network_PID	PID de tabla de información de red	13
program_number1	número de programa (1 a 65535)	16
reserved		3
program_map_PID	PID de tabla de identificación de programa (PMT)	13
...,etc	4 bytes por programa suplementario	...
CRC_32	CRC sobre 32 bits	32
⁽¹⁾ el programa 0 (tabla de información de red, o NIT) es opcional		

Cuadro 8. Descripción de una sección de la tabla de asignación de programas (PAT)

Campo (CAT)	Comentario	N ^o de bits
table_id (01)	siempre a 01 para Control Access Table	8
section_syntax_indicator	siempre a "1"	1
"0"		1
reserved		2
section_length	2 bits de mayor peso a "0", valor máx. 1021	12
reserved		16
reserved		2
version_number	incrementando con cada modificación de CAT	5
current_next_indicator	"1"=CAT actual, "0"=próximo	1
section_number	número de sección en curso (1 ^a =00)	8
last_section_number	número de la última sección (Ntot-1)	8
descriptors	datos de control de acceso	máx. 1012 bytes
CRC_32	CRC sobre 32 bits	32

Cuadro 9. Descripción de una sección de la tabla de control de acceso (CAT)

Campo (PMT)	Comentario	N° de bits
table_id (02)	siempre a 02 para Program Map Table	8
section_syntax_indicator	siempre a "1"	1
"0"		1
reserved		2
section_length	2 bits de mayor peso a "0", valor máx. 1021	12
program_number	especifica el programa descrito a continuación	16
reserved		2
version_number	incrementando cada modificación de PMT	5
current_next_indicator	"1"=MPT actual, "0"=próximo	1
section_number	siempre 0 (una sola sección, identificada .../...	8
last_section_number	.../... por el número de programa)	8
reserved		3
PCR_PID	PID de la Program Clock Reference (PCR)	13
reserved		4
program_info_length	longitud total de los datos útiles (bytes)	12
stream_type ₁	naturaleza del tren elemental N° 1	8
reserved		3
elementary_PID ₁	PID del tren elemental N° 1	13
reserved		4
ES_info_length ₁	N ₁ = N° de bytes de los descriptores siguientes	12

descriptors ₁	datos complementarios	N ₁ bytes
stream_type ₂	naturaleza del tren elemental N ₁ 2	8
reserved		3
elementary_PID ₂	PID del tren elemental N ₁ 2	13
reserved		4
ES_info_length ₂	N ₂ = N ₁ de bytes de los descriptores siguientes	12
descriptors ₂	datos complementarios	N ₂ bytes
..., etc. (ES N ₁ x)		...
CRC_32	CRC sobre 32 bits	32

Cuadro 10. Descripción de una sección de la tabla de identificación de programa (PMT)

Campo (private)	Comentario	N ₁ de bits
table_id	libre (excepto de 00h a 3Fh y FFh)	8
section_syntax_indicator	0=formato libre despues de length, 1=estándar	1
private_indicator	marcador (uso definido por el usuario)	1
reserved		2
private_section_length	2 bits de mayor peso a "1", valor max. 4093	12
private_data_byte(1)	datos privados	max. 4093 bytes
(1) aquí, la zona de datos privados está representada en formato libre (syntax_indicator=0). Si syntax_indicator=1, el formato de datos privados es parecido al de la CAT (excepto en la longitud).		

Cuadro 11. Descripción de una sección de datos privados (private section)

La Figura 46 muestra, de qué forma se utilizan las tablas descritas anteriormente, constituyendo una PSI.

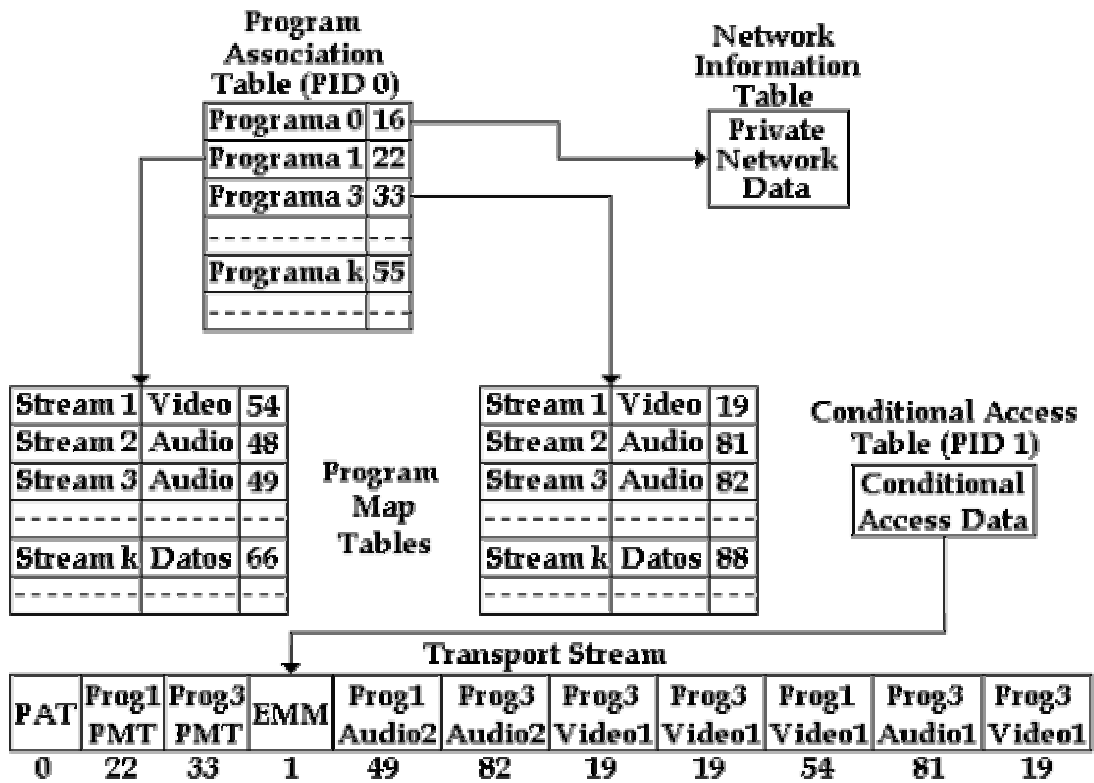


Figura 46. Estructura PSI

3.3.12 INSERCIÓN DE LAS SECCIONES EN LOS PAQUETES DE TRANSPORTE

Al contrario que en los PES, las secciones no empiezan y acaban forzosamente con un paquete de transporte. Cuando una sección o un PES empieza en un paquete, el indicador `payload_unit_start_indicator` (PUSI) se pone a "1".

Cuando se trata de una sección, el paquete puede empezar al final de otra sección, precedida o no de un campo de adaptación (`adaptation_field`). El primer byte de la "carga útil" (payload) es un indicador llamado `pointer_field` el que da el desplazamiento (offset) del comienzo de la nueva sección con respecto a este byte.

La Figura 47 ilustra el caso donde una sección empieza en un paquete de transporte, después de un campo de adaptación (AF) y el final de una sección anterior.

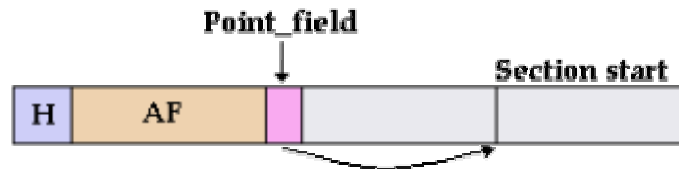


Figura 47. Caso general del comienzo de sección en un paquete de transporte

3.3.13 CÓMO DEMULTIPLEXAR MPEG-2

Las siguientes son las principales etapas que hay que seguir para encontrar un programa o servicio en un múltiplex de transporte MPEG-2, una vez que el múltiplex ha sido capturado (sintonizado en un canal).

A partir de la captura (sincronización) de un nuevo canal:

- Filtrar el PID 0 para capturar los paquetes que transportan las secciones PAT.
- Construir la tabla de asociación de programas (PAT) a partir de las secciones.
- Presentar el menú de programas disponibles al usuario.

Después de elegir el programa:

- Filtrar los PID correspondientes a este programa
- Construir la tabla de programa a partir de las secciones correspondientes
- Filtrar el paquete indicado por el campo PCR -PID; recuperar el PCR y sincronizar el reloj del sistema STC
- Si hay varios PID de audio o video para este programa, presentar las opciones al usuario.

Una vez hecha esta nueva elección.

- Filtrar los PID correspondientes; puede empezar la decodificación propiamente dicha.

La parte visible por el usuario de este proceso es la presentación interactiva de la "guía electrónica de programa" (Electronic Program Guide, EPG); generalmente asociada a la red

por medio de la información proporcionada por las tablas DVB-SI, para permitirle navegar fácilmente por los distintos programas y servicios que se le ofrecen. [7]

3.4 COMPRESIÓN DE VIDEO EN EL ESTANDAR MPEG-4

MPEG-4: el estándar de video para multimedia

El MPEG-4, cuyo nombre formal es ISO/IEC 44196, es un estándar desarrollado por el MPEG, el mismo grupo que desarrolló los estándares MPEG-1 y MPEG-2. El principal objetivo de este nuevo formato es ofrecer al usuario final un mayor grado de interactividad y control de los contenidos multimedia, por lo que en vez de basarse en el conjunto de la secuencia, el MPEG-4 se basa en el contenido. Así, mientras los estándares MPEG-1 y MPEG-2 codifican secuencias, el MPEG-4 es capaz de crear representaciones codificadas de los datos de audio y vídeo que la forman. Un ejemplo simple lo podemos ver en la figura 1. La figura 48 A representa el original sin codificar, mientras que la figura 48 B representa una de las muchas posibles presentaciones de la imagen original tras ser codificada en MPEG-4, decodificada y manipulada por el usuario. El MPEG-4 ha dividido la imagen original en diferentes capas: el tipo de fondo, los diferentes objetos y una capa de texto. Luego ha codificado estos objetos junto con algunos datos asociados de tal manera que permite al usuario decodificar por separado cada uno de ellos, reconstruir la secuencia con sus valores originales o, si lo desea, manipular el resultado, tal y como vemos en la figura 48 B.

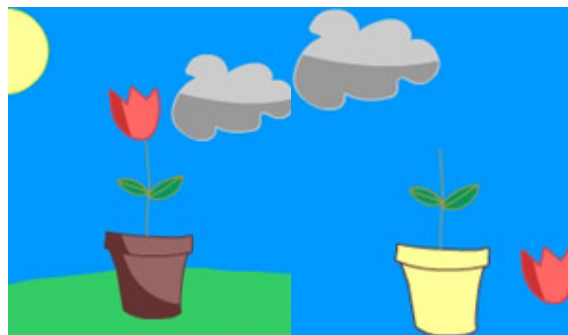


Fig. 48. A

48. B

El estándar de codificación de vídeo MPEG-4 codifica cada objeto en capas separadas. El contorno y transparencia de cada objeto, así como las coordenadas espaciales y otros parámetros adicionales como escala, localización, zoom, rotación o translación, son incluidos como datos asociados de cada objeto en su propia capa. El usuario puede reconstruir la secuencia original al decodificar todas las capas de objetos y visualizándolos sin modificar los parámetros asociados a los objetos. Como alternativa, el usuario tiene la posibilidad de manipular la secuencia realizando unas operaciones muy sencillas. Por ejemplo, en la Figura 48 B no se han decodificado ni usado algunos objetos para la reconstrucción, otros han sido representados con los mismos valores del original, mientras que otros objetos han visto modificados algunos de sus parámetros de escala, rotación o posición. Los parámetros de escala, rotación y translación usados para la manipulación de la secuencia de imagen pueden alterarse e incluirse durante el flujo de bits mediante operaciones de edición muy simples y sin la necesidad de tener que realizar nuevas decodificaciones. También es posible añadir objetos que no estén presentes en la secuencia original. Además, el MPEG-4 no hace distinción entre objetos naturales y sintéticos. Se entiende por objeto sintético cualquier recreación informática, como gráficos en tres dimensiones o incluso voz sintetizada.

3.4.1 Lo nuevo del MPEG-4

Comparándolo con sus dos antecesores, el MPEG-4 añade nuevas posibilidades. Principalmente ofrece nuevas y diferentes variedades de interactividad. Permite integrar los contenidos naturales y sintéticos en forma de objetos. Estos objetos, pues, pueden representar entidades "registradas" (una persona, una silla) o material sintético (una voz, una cara, un modelo de animación en 3D). Por lo tanto, también soporta contenidos en 2D y 3D. Ofrece mayor versatilidad en cuanto a relaciones de flujo, desde codificaciones con una relación de flujo muy baja (2 Kb/s para conversación, 5 Kb/s para vídeo) hasta relaciones muy altas (5 Mb/s para vídeo con calidad transparente, 64 Kb/s por canal para audio con calidad similar al CD). A todo ello, el MPEG-4 ofrece una administración y protección mejorada de la propiedad intelectual. En cuanto a la codificación, el MPEG-1 y MPEG-2 utilizan secciones de imagen rectangular para definir las partes en movimiento, pero el MPEG-4 es capaz de "perfilar"

estas secciones. Al definir el contorno de los objetos puede despegarlos de los otros objetos que participan en la secuencia original, así como del fondo. Una vez definidos los objetos y el fondo, permite la introducción de nuevos objetos naturales o sintéticos. En ambos casos el MPEG-4 seguirá conociendo la naturaleza de cada uno. La codificación se realiza de una manera muy similar a la codificación MPEG-1 y MPEG-2, incluyendo la predicción y compensación de movimiento seguido de una codificación de textura basada en DCT. Durante la decodificación MPEG-4 se recupera la información de cada objeto, junto con sus datos asociados, ofreciendo al usuario la posibilidad de modificar cualquier parámetro disponible (es el autor quien define qué parámetros pueden o no modificarse, lo que permite ofrecer diferentes grados de interacción) y sin la necesidad de nuevas decodificaciones.



Fig. 49. resolución del mpeg-4.

3.4.2 Resoluciones del MPEG-4

El MPEG-4 ha sido diseñado tanto para la teledifusión como para la difusión por la Web, mejorando la convergencia de ambos canales, ya que permite la integración de contenidos provenientes de ambos en la misma escena multimedia. Esta facilidad de difusión viene provista gracias a las diferentes relaciones de flujo que el estándar permite. Para la difusión de vídeo con flujos muy bajos (VLBV, very low bit rate video) se ofrecen diversos algoritmos y herramientas para aplicaciones que trabajen con flujos entre los 5 y 64 kbit/s. Un ejemplo es claramente la difusión por la Web.

Esta posibilidad soporta la secuencia de imágenes con resoluciones espaciales muy bajas (desde pocos píxeles por línea y filas hasta resolución CIF) y relación de cuadro baja (desde los 0 Hz para imágenes fijas hasta los 15 Hz). Las aplicaciones básicas que soportan esta aplicación pueden codificar secuencias de imagen rectangular con una alta eficiencia de

codificación y una alta resistencia a los errores, bajo tiempo de recuperación de datos y una baja complejidad para aplicaciones de comunicación multimedios a tiempo real. Además pueden ofrecer el acceso remoto, avance rápido y retroceso rápido para aplicaciones de almacenaje y acceso multimedios. En realidad, las mismas funciones que soporta el VLBV lo son para el HBV (higher bit rate video, vídeo de relación de bits alta), aunque con la posibilidad de ofrecer resoluciones muy cercanas a los parámetros R-601 (PAL estudio), utilizando los mismos algoritmos y herramientas de en el VLBV. El MPEG-4 soporta tanto la exploración progresiva como entrelazada.

3.4.3 Interactividad

Para permitir las funciones de interactividad basadas en el contenido, el estándar de vídeo MPEG-4 introduce el concepto de "planos de objetos de vídeo" VOP (Video Object Planes). Cada cuadro de la secuencia de vídeo a codificar es seccionado en un número de imágenes independientes VOP. En contraposición con el formato de vídeo de entrada usado con los estándares MPEG-1 y MPEG-2, el vídeo de entrada para ser codificado por el modelo de verificación del MPEG-4 no tiene que ser únicamente una imagen de vídeo rectangular. Puede ser también una región VOP con un contorno arbitrario, cuyos valores de contorno y localización pueden variar cuadro a cuadro. La sucesión de regiones VOP pertenecientes a un mismo objeto físico en una escena se conoce como objetos de vídeo VO (Video Objects). La información del contorno, movimiento y textura de cada VOP perteneciente a un objeto de vídeo es codificada y transmitida o codificada como una capa de objetos de vídeo VOL (Video Object Layer) independiente. Además se incluye la información relevante necesaria para identificar cada capa de objetos de vídeo VOL y su composición para reconstruir la secuencia original en el receptor. Esto permite la decodificación separada de cada plano de objetos de vídeo VOP y la posibilidad de manipulación.

Hemos visto que las imágenes MPEG-4 así como las secuencias de imagen son consideradas como contorneadas arbitrariamente, en contraste con las definiciones de los estándares MPEG-1 y MPEG-2 que codifican secuencias de imágenes rectangulares. Una codificación MPEG-4 de una secuencia de vídeo rectangular sin definición de contornos

ofrece una codificación muy similar al MPEG-1/2. En realidad, el MPEG-4 no especifica una única manera para codificar información de audio o vídeo, pero ofrece una herramienta para utilizar diferentes métodos de codificación, que pueden usarse para diferentes tipos de contenido. Cada uno tiene su codificador optimizado.

3.4.4 Contenido sintético y más posibilidades

Una de las aportaciones del MPEG-4 es la posibilidad de converger material registrado con material sintético. Un ejemplo es el audio estructurado, que permite aprovechar eficientemente el ancho de banda para la creación de contenidos de audio sintéticos de alta calidad. Está estrechamente relacionado con el MIDI, ya que permite la difusión de "órdenes musicales" en vez del propio audio, optimizando el ancho de banda (las partituras ocupan mucho menos espacio que el propio sonido). Sería como enviar por la red únicamente la partitura musical, que sería reproducida por nuestra propia banda de músicos a tiempo real. Otra posibilidad es la animación facial. Una de las versiones del MPEG-4 es la Interfaz Texto-para-locución, que permite la difusión de un texto para que el receptor lo lea textualmente y lo reproduzca utilizando una cabeza sintética inteligente. El teledifusor sólo debe emitir el texto, y el receptor utilizaría sus propias herramientas para crear una cara que interpretase el texto incluyendo los movimientos faciales, por ejemplo un presentador de noticias. La versión 2 del MPEG-4 incluye la animación corporal. También puede interpretar complejas estructuras lineales en 2D y 3D. De esta manera, junto con información sobre textura, el receptor puede crear cualquier objeto sintético tomando como referencia la estructura del esqueleto en 3D y la información de su textura, de la misma manera que un ordenador de animación 3D interpreta estos datos para la creación de una secuencia de imágenes generadas por computadora. [8]

3.5 ESTÁNDAR MPEG-7 Y SUS APLICACIONES

MPEG-7 es una representación estándar de la información audiovisual. Permite la descripción de contenidos por palabras clave y por significado semántico (quién, qué, cuándo, dónde) y estructural (formas, colores, texturas, movimientos, sonidos). El formato MPEG-7 se asocia de forma natural a los contenidos audiovisuales comprimidos por los

codificadores MPEG-1 (almacena y descarga archivos audiovisuales), MPEG-2 (televisión digital) y MPEG-4 (codifica audio y vídeo en forma de objetos), pero se ha diseñado para que sea independiente del formato del contenido.

El nuevo estándar ayuda a las herramientas de indexación a crear grandes bases de material audiovisual (imágenes fijas, gráficos, modelos tridimensionales, audio, discursos, vídeo e información sobre cómo esos elementos están combinados en una presentación multimedia) y buscar en estas bases de materiales manual o automáticamente.

Independiente de la tecnología es indudable la importancia en el mundo de hoy del material audiovisual en donde dado su natural relación con los sentidos hace mas completo su percepción por las personas, pero la recuperación de información sobre los sistemas multimediales es mucho mas compleja puesto que las personas consideran nuevos factores y asocian (o pueden asociar) nuevos conceptos en sus búsquedas.

Además el número de aplicaciones que crean y manipulan este tipo de contenido tiene tendencia a crecer, sobre todo los que involucran un procesamiento automático sin ser despreciable la cantidad de información existente en medios tradicionales (no digitales) en cintas y videos.

Los objetos audiovisuales incluyen imágenes estáticas, audio, video y modelos de profundidad entre otros, numerosas son las fuentes de contenido audiovisual: una película, un partido de fútbol una obra de teatro, un concierto, etc.

Con sus respectivos elementos interrelacionados en un escenario, un dialogo, objetos, personajes, movimientos, tonos de voz, etc, etc.

El conocer o definir la estructura de los “elementos” que podemos percibir “audiovisualmente” permite efectuar operaciones sobre estos objetos, desde reconocer las características faciales de una persona entre miles en un video, hasta encontrar patrones geológicos de un mineral en particular en una fotografía.

Son infinitas las ideas y proyectos que se puede uno imaginar para operar sobre el material audiovisual, si se conociese algo más sobre él:

- Obtener una lista de grabaciones y videoclips a partir de un extracto de la voz de Shakira.

- Filtrar un partido de fútbol sin noticias ni comerciales. Buscar escenas específicas tales como ... cuando el balón (objeto) se acerque a la portería (objeto) ...
- Mantener control de un sistema de seguridad mediante cámaras de vigilancia ..
- Buscar los discursos de un político en particular acordándose de su cara.

Existen grupos que se dedican a buscar nuevas aplicaciones potenciales para un sistema con estas características. Uno de ellos es el grupo MPEG, el que a su vez provee tecnologías para hacerlo realidad.

En este artículo se describirá someramente (dada la profundidad y complejidad del mismo) el MPEG-7 y sus aplicaciones, formalmente llamado “Multimedia Content Description Interface” o *interfaz para la descripción de contenido multimedia* es un estándar para describir el contenido multimedia para su posterior interpretación en términos de significado.

El MPEG-7 es un estándar ISO/IEC desarrollado por MPEG (Moving Picture Expert Group) comité que también desarrolló los reconocidos estándares MPEG-1, MPEG-2 y MPEG-4 los primeros dos hacen posible la distribución de video en CD-ROM y la televisión digital, el MPEG-4 fue el primer estándar real para multimedia permitiendo la codificación de audio y video en forma de objetos, pero el MPEG-7 va mas allá siendo diseñado para que sea independiente del formato del contenido.

El estandar fue desarrollado por expertos representando locutores, creadores de contenido, fabricante de aparatos electrónicos, editores, encargados de propiedad intelectual, proveedores de servicios de telecomunicaciones y universidades mediante un trabajo de desarrollo que comienza en 1998, en donde se realizó un llamado a propuestas de acuerdo a los objetivos y horizonte de lo que el estándar debía cubrir.

De las casi 400 propuestas recibidas se seleccionaron los elementos que fueron incorporados en un modelo común que evoluciona hasta el día de hoy.

3.5.1 Objetivos de MPEG-7

El contenido audiovisual puede incluir: imágenes estáticas, gráficos, modelos 3D, audio, video, y escenarios de los elementos anteriores combinados, la descripción admite distintos

niveles de granularidad y NO depende de la representación específica del material (MPEG1, JPEG, MPEG2, PCM, Etc...)

Las descripciones pueden encontrarse físicamente asociadas al material audiovisual o en cualquier otro lugar del mundo. Las descripciones pueden ser intercambiadas independientemente del material, pudiéndose procesar estos meta datos sin necesidad de poseer el objeto descrito, en ambas direcciones, Además las descripciones pueden ser procesadas “on-the-fly” lo que provee operaciones como las de filtrado.

El material puede ser descrito de distinta forma dependiendo de la necesidad, ya que las características descriptivas deben tener un significado en el contexto de la aplicación. Necesariamente estas descripciones serán diferentes para distintos dominios de usuarios y sistemas, por lo que no se pretende generar un sistema único para la descripción de contenidos, sino proveer un conjunto de métodos y herramientas para satisfacer los distintos puntos de vista que distintos usuarios poseen.

El material puede ser descrito utilizando diferentes niveles de abstracción, los trazos en un dibujo y los cambios de ritmo de una melodía pueden catalogarse como de bajo nivel de abstracción, mientras que la información semántica “Un gol a punto de anotarse...” cae en un nivel más alto. Mientras mayor es el nivel de abstracción, mayor es la dificultad para realizar un procesamiento automático.

Además de disponer de la descripción relacionada con el contenido, también es necesario incluir otro tipo de información descriptiva, como es la información sobre su creación, información sobre el formato utilizado, derechos de autor, links hacia otro material relacionado.

Un diagrama simple de cómo es el funcionamiento o ciclo de este estándar:

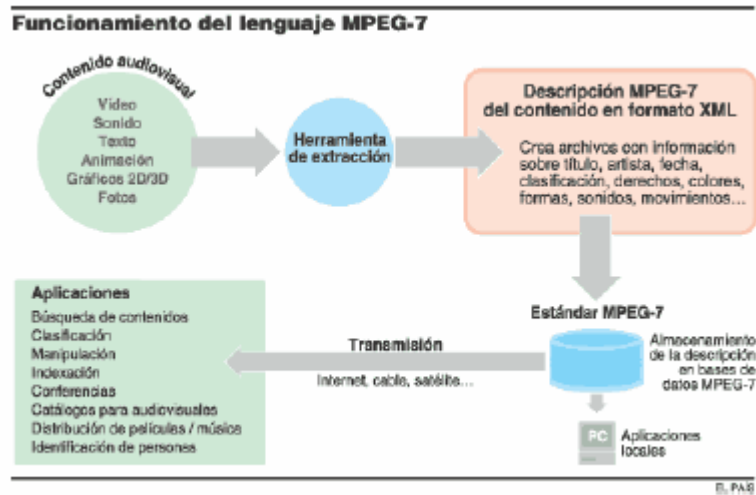


Fig. 50. funcionamiento del lenguaje MPEG-7.

3.5.2 Principales Elementos del estándar MPEG-7.

El estándar MPEG-7 consiste de las siguientes partes:

1. MPEG-7 sistema.
2. MPEG-7 DDL.
3. MPEG-7 Visual.
4. MPEG-7 Audio.
5. MPEG-7 Esquemas de descripción multimedia.
6. MPEG-7 software de referencia.
7. MPEG-7 pruebas de conformancia.
8. MPEG-7 Extracción y uso de descripciones.

Para tener un mejor entendimiento de la terminología que se mencionará mas adelante veamos esta gráfica:

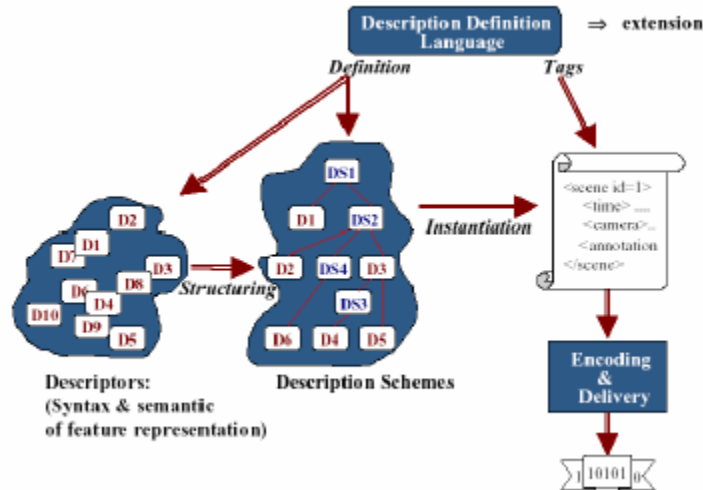


Fig. 51. terminología mpeg-7.

En detalle:

1. Sistema MPEG-7.

El sistema MPEG-7 incluye las herramientas que se necesitan para preparar descripciones MPEG-7 de manera eficiente para transporte y almacenamiento (BIM formato binario) y la normatividad y arquitectura de sus interfaces.

2. Un 'Description Definition Language' (DDL)

Para definir las relaciones estructurales entre descriptores y permitir la creación de nuevos esquemas de descripción. MPEG7 adopta como su DDL el XML Schema, extendido para satisfacer algunos requerimientos específicos como polimorfismo y nuevos tipos de datos espaciotemporales y audio-visuales.

Sin embargo, una descripción en XML puede ser muy voluminosa. Es un problema para las aplicaciones en las que el espacio de almacenamiento o el ancho de banda de transmisión son insuficientes (discos con capacidad limitada, transmisión por módem, etc.) Para estos casos se ha desarrollado el compresor BIM (Binary Format for MPEG-7).

En un escenario típico, una aplicación genera la descripción MPEG-7 del contenido de, por ejemplo un millón de películas; luego se pasa al formato XML, se almacena en servidores o si se tiene que mandar en un canal de transmisión con poco ancho de banda, el XML se compacta en un espacio hasta 100 veces menor con el codificador BIM. Al final del proceso se puede decodificar otra vez en XML y ya se pueden utilizar estos datos, además el BIM es más robusto que el XML frente a los errores de transmisión.

El DDL puede ser subdividido en los siguientes componentes normativos lógicos:

- Componentes de lenguaje para el esquema estructural XML.
- Componentes de lenguaje para el esquema de datos XML.
- Extensiones específicas MPEG-7.

3. MPEG-7 Visual

Consiste en las estructuras básicas y descriptores que cubren distintas características visuales: color, textura, forma, movimiento, localización y reconocimiento de caras. Los 'visual descriptors' son el componente que especifica la estructura y semántica de las relaciones entre sus componentes. (Para filtrado y recuperación de la información)

4. MPEG-7 Audio

Son un conjunto de descriptores de bajo nivel para el contenido de audio, se describen mediante ellos las características espectrales, paramétricas y temporales de una señal. También son utilizados descriptores de alto nivel que incluyen reconocimiento de sonido general y herramientas para la indexación de descriptores, herramientas para descripción de timbres instrumentales, herramientas para contenido hablado, un esquema para firma de audio y herramientas para la descripción de melodías.

5. MPEG-7 Esquemas de descripción multimedia.

Se optó por utilizar múltiples 'Multimedia Description Schemes' en vez de un gran sistema de descripción. Estos esquemas son estructuras que son definidas utilizando el 'Description Definition Language'(DDL) de MPEG7 y tienen muy variadas funciones. La generación de estas descripciones pueden generarse manualmente o procesadas automáticamente. Algunos ejemplos de estos esquemas:

Descripción del contenido: representación de la información percibible del mundo real, la semántica asociada a las acciones eventos y objetos.

Administración del contenido: información sobre las características, creación y uso del contenido audiovisual (Dublin core)

Acceso y navegación: especificación de resúmenes sobre las variaciones del material audiovisual a lo largo del tiempo.

Interacción del usuario: Descripción de las preferencias del usuario y su historial de uso referente al consumo de material multimedia

6. MPEG-7 software de referencia.

MPEG-7 Dispone además de modelos, procedimientos y software para la simulación de una plataforma sobre 'Descriptors', 'Description Schemes', 'Coding Schemes' y la 'Description Definition Language'. Útiles a la hora de desarrollar.

7. MPEG-7 pruebas de conformidad.

Estándar que se encuentra bajo desarrollo todavía se trata de guías y procedimientos para probar la conformidad con el estándar de implementaciones MPEG-7.

8. MPEG-7 Extracción y uso de descripciones.

El reporte técnico para la extracción y uso de descripciones que se encuentra bajo desarrollo incluirá material informativo acerca de la extracción y uso de algunas herramientas para dar adicional soporte a MPEG-7.

3.5.3 Áreas de aplicación del MPEG-7

Los elementos que MPEG-7 estandariza provee soporte para un amplio rango de aplicaciones (por ejemplo: Librerías digitales multimedia, edición multimedia, elementos de entretenimiento casero, etc.) también permite mejorar las búsquedas Web de material de contenido multimedia que hoy sólo se puede encontrar por medio de texto, todo esto aplica a grandes contenidos de información, catálogos multimedia etc.

Todo el dominio de aplicaciones multimedia puede beneficiarse de MPEG-7, considerando de que hoy es difícil encontrar un campo que no use multimedia, algunas ideas:

- Arquitectura, Diseños de interiores (Ej. Buscando ideas)
- Selección de medios de transmisión (broadcast) (Ej. Canales de radio, canales de TV)
- Servicios culturales (Museos de historia, galerías de Arte, etc.)
- Librerías digitales (Ej. Catálogos de imágenes, música, catálogos de imágenes biomédicas, archivos de video, etc.)
- Comercio Electrónico (Ej. Publicidad personalizada, catálogos en línea, etc.)
- Educación (Ej. Repositorios de cursos en multimedia, búsqueda de información y soporte).
- Hogar (sistemas para el manejo de las colecciones personales, videos, fotos, etc.)
- Periodismo (búsqueda de material por sonidos, rostros, etc.)

- Servicios de directorio multimedia (páginas amarillas, guías de turismo, información geográfica, etc.)

- Supervisión (Ej. Control de tráfico, sistemas de seguridad, etc.)

La forma como los datos MPEG-7 pueda ser usada al responder a operaciones de preguntas o filtros por parte del usuario esta fuera del alcance del estándar. El tipo de contenido y la pregunta tampoco tiene que ser del mismo tipo, por ejemplo material visual puede ser solicitado o filtrado usando otro contenido visual, música, voz, etc. Es responsabilidad del agente que busca y/o filtra el hacer coincidir los datos con la descripción MPEG-7.

Unos pocos ejemplos de preguntas son:

- Ejecutar unas pocas notas en el teclado y recuperar piezas musicales similares.

- Dibujar unas pocas líneas en la pantalla y encontrar un conjunto de imágenes que contengan graficas similares, logos, diagramas...

- Definir objetos, incluyendo partes de color, textura, y recuperar objetos que contengan parte del diseño propuesto.

- Describir acciones y obtener escenarios conteniendo tales acciones.

- Usando un extracto de la voz de Shaquira, obtener lista de discos de ella, video clips donde ella cante, y material fotográfico donde aparezca.

Un pequeño ejemplo de cómo puede ser el análisis de por ejemplo un video.

En este ejemplo se descompone y organiza el material en estructuras que describen las relaciones espacio-temporales, con el objeto de representar los aspectos lógicos y temporales del contenido audiovisual. La descomposición se manifiesta en muestras de audio, en regiones de las imágenes, un conjunto de cuadros del video o combinaciones de ellos, la descomposición puede materializarse en un árbol u otra estructura dependiendo del alcance de la descripción.

Los segmentos son subdivididos recursivamente. Un segmento puede ser descompuesto en varios niveles de las escenas, tomas, varias pistas de audio o puntos de vista de muchas cámaras.

Ejemplo de un partido de fútbol, donde las relaciones entre tres regiones en movimiento y una estática son descritas mediante un grafo:

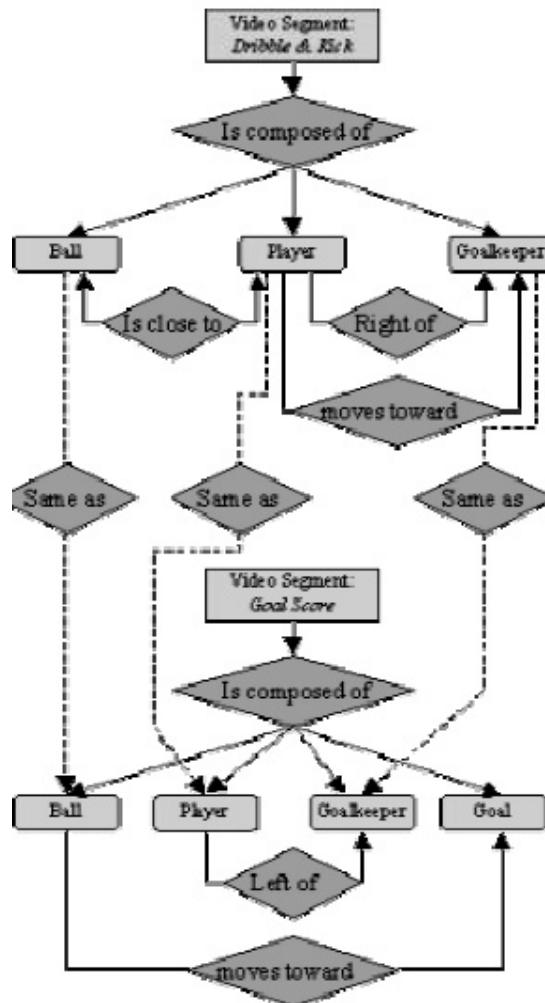
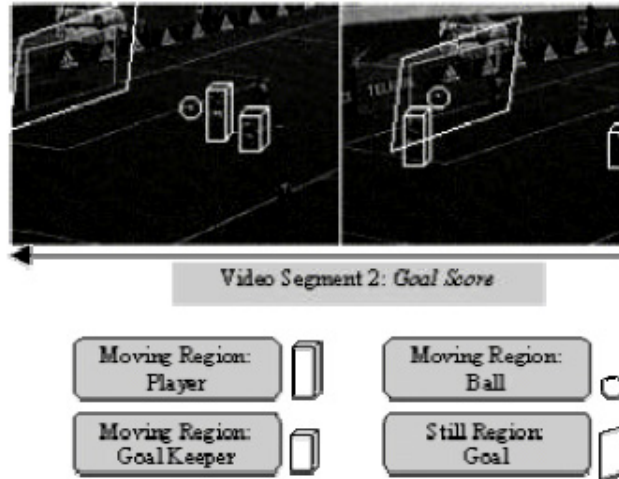


Fig. 52. tres regiones en movimiento y una estática son descritas mediante un grafo.

3.5.4 El PostScript del multimedia

MPEG-7 es al multimedia lo que PostScript es al papel. Este último describe a un programa de textos el formato que debe tener la página; MPEG-7 hace lo mismo, pero sobre el contenido audiovisual. El MPEG-7 se basa en el popular lenguaje de metadatos XML, "que dominará Internet dentro de poco; habrá un montón de herramientas y eso favorecerá la interoperabilidad y la creación de aplicaciones", afirma Philippe Salembier, investigador de la Universidad Politécnica de Cataluña, que ha participado en el proyecto MPEG-7 durante los últimos tres años.

Sin embargo, una descripción en XML puede ser muy voluminosa. Es un problema para las aplicaciones en las que el espacio de almacenamiento o el ancho de banda de transmisión son insuficientes (discos con capacidad limitada, transmisión por módem, etcétera). Para estos casos se ha desarrollado el compresor BIM (Binary Format for MPEG-7).

En un escenario tipo, una aplicación genera la descripción MPEG-7 del contenido de, pongamos, un millón de películas; luego se pasa al formato XML, se almacena en servidores con discos de gran capacidad y ya está listo para su uso. Si la información de la descripción es demasiado grande para los servidores o si se tiene que mandar en un canal de transmisión con poco ancho de banda, el XML se compacta en un espacio hasta 100 veces menor con el codificador BIM. Al final del proceso se puede decodificar otra vez en XML y ya se pueden utilizar esos datos. "Además, el BIM es más robusto que el XML frente a los errores de transmisión", añade Salembier.

MPEG-7 incluye, además de la descripción de los contenidos, información sobre el tipo de compresión utilizada (JPEG, en dibujos; MPEG-2, en imágenes), las condiciones para acceder (derechos, precio), clasificación (adultos, por ejemplo), enlaces a otros materiales relevantes (para acelerar la búsqueda) y el contexto (final de los 200 metros femeninos de los Juegos Olímpicos de Verano de 2000).

El lenguaje MPEG-7 puede generar grandes ingresos al sector audiovisual (producción, posproducción distribución y consumo). A juicio de Goldman Sachs, el mercado de herramientas informáticas para la gestión de contenidos pasará de los 378 millones de

dólares (7.100 millones de pesetas) del año 2000 a los 4.500 millones de dólares (850.000 millones de pesetas) en 2005.

La industria también prepara el estándar MPEG-21, que propone la creación de un marco tecnológico para todo el proceso multimedia, explica Jose M. Martínez, científico de la Universidad Politécnica de Madrid, que también ha participado en el desarrollo del MPEG-7. El MPEG-21 debe abarcar desde los terminales utilizados al tipo de redes. También será importante para proteger legalmente los contenidos para todas las variedades de uso. En la primera fase, intentará definir las diferentes opciones tecnológicas de extremo a extremo en la cadena multimedia. Por ejemplo, se enviará un esquema MPEG-7 adecuado a un portátil.

¿Quién ganará la partida audiovisual? Seguramente, el creador de aplicaciones que tengan mayor capacidad de actuar entre los diversos agentes del proceso industrial. "La industria tiene ahora la última palabra", afirma José Martínez. Puede mirarse en el espejo del MPEG-2, formato digital adoptado por las cadenas de televisión.

3.5.5 las primeras soluciones con el lenguaje

Las primeras aplicaciones del MPEG-7 están viendo la luz. Serán importantes en sectores como el entretenimiento, la medicina, la educación, la industria, la seguridad o la arquitectura. La biblioteca del Congreso de Estados Unidos es un gran cliente potencial: recibe 10.000 entradas multimedia a la semana y necesita clasificarlas, archivarlas y ponerlas fácilmente a disposición de los ciudadanos en formato digital.

El Fraunhofer Institute, creador del MP3, ha desarrollado el AudioID, un sistema para la identificación automática de una canción en cinco segundos entre más de 30.000, afirma Eric Allamanche, investigador del instituto alemán. Utilidades: venta por catálogo o control de los derechos; incluso podrá servir para buscar desde el móvil piezas en las que el piano sea el instrumento protagonista. A finales de 2001 aparecerá la licencia de la tecnología, posiblemente para la empresa Thomson.

Telematica Institut e IBM prueban en Holanda servicios personalizados de vídeo sobre IP (educación y negocios), y LG Electronics, un buscador de información audiovisual para

proveedores de contenidos y usuarios; Ricoh experimenta un software para la recuperación y distribución de videoclips a través de una interfaz Web; NEC ha desarrollado un equipo de identificación de vídeos en tiempo real, y Expway, el codificador Bin-XML y TV Anytime, en su proyecto de televisión interactiva. [9]

3.6 MPEG-21

MPEG-21 proporciona un marco de intercambio de contenidos multimedia, respetando los derechos de autor, de copia y de distribución, adecuando los contenidos a las capacidades de los usuarios en cada momento.

Dentro de la estandarización de contenido multimedia un nombre nos viene inmediatamente a la cabeza: *MPEG*. El "*Moving Picture Experts Group*", o de forma abreviada, *MPEG* es uno de los grupos de trabajo de *ISO* (International Organization for Standardization). Es el encargado de la estandarización para la representación digital de audio y vídeo. Fundado en 1988, el grupo ha producido los siguientes estándares:

MPEG-1: Utilizado para los vídeo-Cds, fue el primer estándar promovido por el grupo. El formato de audio MP3 forma parte de la especificación de la capa de sonido del MPEG-1.

MPEG-2: Este formato es para vídeo de alta calidad, como la transmisión digital de televisión o los DVD.

MPEG-4: Algo menos conocido, busca ser el estándar de los contenidos multimedia en las redes móviles y fijas de intercambio de información.

Las siguientes dos iniciativas en las que el grupo trabaja se alejan de la codificación y compresión directa del audio y vídeo, y se vuelcan en mayor medida a la gestión de estos contenidos multimedia. En primer lugar aparece MPEG-7, encargado de la estandarización de sistemas de descripción y búsqueda para contenido multimedia. Por último llega, en Junio de 2000, MPEG-21, el nuevo "marco digital de intercambio de contenidos".

Para muchos de nosotros raro es el día en que no descargamos una foto, un vídeo, una canción o cualquier otro documento digital de Internet. El consumo de este tipo de contenidos es cada vez más alto. Consultar la cartelera, un mapa, las páginas amarillas, escuchar música,..., todo ello desde nuestros ordenadores, agendas electrónicas e incluso en

los últimos modelos de teléfonos móviles. Parece que el mundo digital ha entrado, por fin, en nuestras vidas con intención de quedarse.

A medida que la demanda aumenta, la creación de este tipo de registros multimedia crece, y este crecimiento es cada vez más rápido, alcanzando un ritmo exponencial. Los proveedores de materiales se multiplican, aparecen nuevos servicios casi a diario, nacen nuevos formatos como el MP3, OGG o DivX, que han impulsado la aceptación del público de este tipo de materiales.

Las nuevas tecnologías e Internet nos dan acceso a unos medios tan amplios que a día de hoy cualquiera puede poner a disposición de los demás las fotos que sacó en sus vacaciones, una canción que compuso el fin de semana con los amigos o un pequeño ensayo que está escribiendo en los ratos libres.

Todos estos cambios, con las nuevas posibilidades que brindan, traen consigo nuevos retos. No se trata tan solo de retos tecnológicos sino mucho más profundos. Se requieren cambios sociales y culturales para poder absorber las capacidades que trae consigo el mundo digital. El mercado digital ya está aquí pero ¿estaremos preparados para él?

Los grandes creadores de contenidos ven en este mercado un problema potencial muy grave.

Sus peores pesadillas se hacen realidad cuando iniciativas como las redes punto a punto, como KaZaA o Edonkey, proliferan en todas las partes del mundo. En ellas sus usuarios intercambian cualquier clase de contenidos, infringiendo frecuentemente, las limitaciones de copia o distribución. Canciones o incluso películas se intercambian a diario en este tipo de redes, lo que preocupa sobremanera a estas compañías. Los productores ven en el intercambio digital de contenidos un problema para el desarrollo de sus negocios y para su entrada en la venta de contenidos por Internet. Por ello existe una necesidad imperiosa para ellos de crear herramientas para poder acceder a contenidos de *calidad* con todas las seguridades que necesitan.

3.6.1 Definición MPEG-21

Con todas estas limitaciones parece que se ha creado una nueva necesidad, la de controlar la distribución de copias y distribución de registros digitales. Para ello se deben desarrollar herramientas que permitan tratar de forma eficiente los registros multimedia, ajustándolos a nuestras necesidades en cada momento. A la vez es necesario crear un marco protegido dentro de los actuales sistemas de transmisión de contenidos digitales. Su fin no es otro que poder trasladar los actuales modelos de negocio y ventas a los que estamos acostumbrados en el mundo real al mundo virtual de las redes de datos. Las limitaciones buscadas son claras: impedir la copia y distribución no autorizada de materiales que tengan derechos de autor, y derechos derivados de este, restringidos.

Según MPEG para que un mercado digital sea viable dentro de la red, es necesario, para determinados tipos de contenidos, poder construir un entorno con estas mismas limitaciones, en los que la copia esté sujeta a restricciones, para así garantizar el desarrollo de la propiedad intelectual.

El propósito de MPEG-21 es establecer un marco claro para la realización de transacciones dentro de un mercado digital: el acceso, envío, gestión y protección de estos contenidos de forma integrada, totalmente transparente para el usuario. Dentro de este marco de intercambio de contenido, cualquier usuario podrá descargarse registros multimedia, adecuados a las capacidades del dispositivo que esté empleando en ese momento.

Para ello, MPEG-21 proporciona una visión del escenario en el que va a trabajar, incluyendo las estructuras, tecnologías y ejemplos de uso, y se encargará de facilitar la integración de componentes y estándares para armonizar las tecnologías necesarias.

El primer punto con el que tratará el estándar es el propio bien con el que se realizarán transacciones. Dada la importancia de este punto no es raro que una parte muy importante de la estandarización de MPEG-21 se esté dedicando a la definición de lo que es un *objeto digital*.

En la celebración de un acontecimiento familiar contratamos los servicios de un fotógrafo. Durante toda la fiesta se recogen imágenes, vídeos o incluso grabaciones de sonido. Tras tomar todas las fotos, se puede enviar los contenidos a un experto para que haga un

procesado eliminando, por ejemplo, los ojos rojos y publique el álbum, basándose en algún criterio estético.

Cualquiera puede personalizar la presentación de este álbum, por ejemplo, incluyendo música de otros artistas en modalidad de pago por escuchar. Los usuarios que tengan los permisos adecuados, es decir, los invitados, podrán conectarse y ver el álbum, escuchar la música recomendada o ordenar copias en papel de las fotos al fotógrafo. También podrían hacerse otras producciones como salva-pantallas, CD, camisetas o tazas, cada una de estas gestionada por una empresa distinta elegida por los clientes, pero todas ellas podrían tener acceso al mismo álbum. Todo este proceso sería totalmente transparente a los usuarios finales, que podrían ver las fotos, vídeos, etc, desde cualquier sitio. Además la intimidad está garantizada porque nadie podría sacar las fotos para un uso incorrecto.

La ventaja que presenta MPEG-21 es que en todo momento se tiene constancia de quién es el autor de cada una de las partes del objeto digital, en este caso el álbum con fotos, la música o los servicios añadidos. De esta forma cada uno de los autores puede obtener una compensación económica cada vez que alguien emplee uno de sus trabajos, todo ello gestionado por MPEG-21.

3.6.2 EJEMPLO DE USO DE MPEG-21

Dentro de MPEG-21 ambos serán, simplemente, *usuarios*, sin distinción alguna entre ellos. Cualquiera podrá servir y recibir indistintamente cualquier contenido en un momento dado. También se estandariza la forma en la que se intercambian objetos digitales entre ellos, incluyendo las limitaciones a las que podrían verse sometidos éstos. Si yo he comprado una canción para poderla usar en mi reproductor MP3, el terminal MPEG-21 debe asegurar que únicamente la voy a usar en mi reproductor y que no la voy a enviar a otras personas. Ningún contenido que esté dentro de la red MPEG-21 debe salir de ella, puesto que fuera de la misma no se podría asegurar la protección de la propiedad intelectual y los derechos de copia de la obra. El *Lenguaje de Expresión de Derechos* pretende proteger a los contenidos y a los usuarios contra estas infracciones y es otro de los grandes pilares de la definición de MPEG-21.

A día de hoy todos estos puntos no son más que guías generales para el desarrollo de MPEG-21. Este estándar todavía se encuentra abierto, en proceso de desarrollo. Los calendarios del proyecto tienen previsto prolongarse hasta mediados de 2004.

3.6.3 Objetos Digitales

El MPEG-21 está basado en la definición de una unidad básica de contenido con la que realizar operaciones: el *objeto digital*. Definir qué es un objeto digital, y que no lo es, resulta ser una tarea harto complicada. Cada vez es más difusa la frontera de separación entre archivos de vídeo, audio, texto o espacios sintéticos y todos ellos son sujetos de ser objetos digitales, pero cada uno debería tener propiedades diferentes. También es necesario tener en cuenta cuales eran las intenciones del autor a la hora de crear un contenido. Según cuales fueran el objeto digital puede ser bastante diferente en un caso u otro.

Tomemos, a modo de ejemplo, una herramienta que presente una página web diferente en función del idioma del usuario. A la hora de definir el objeto digital existe la ambigüedad de si el propósito del autor era considerar el objeto formado por todas las páginas de diferentes idiomas, que cada una de ellas se trate como un objeto digital independiente, o que, por el contrario, el objeto sea la herramienta que crea las páginas.

Dado que las dos vertientes están sujetas de ser transferidas dentro de la red MPEG-21 como objetos digitales, debe ser labor del autor especificar qué contiene el objeto a la hora de crearlo. La herramienta para solucionar esta ambigüedad es la declaración de objeto digital. Su objetivo es definir una serie de términos abstractos que sirvan para expresar que es un objeto digital, entendiendo este como la *representación digital de la realización de un trabajo*. La declaración de objeto digital se basa en la utilización de un modelo, que debe ser tan flexible y general como sea posible. El modelo no define un lenguaje en particular, simplemente es una ayuda para proveer de un conjunto abstracto de términos y conceptos que pueden usarse para definir un esquema o realizar una equivalencia entre diferentes esquemas con el fin de realizar comparaciones.

Para trabajar con objetos digitales es necesario identificar cada uno de manera única. También se debe identificar cuál es el dueño de la propiedad intelectual asociada a un

objeto determinado, cuál es el esquema de descripción que se está empleando y que restricciones tiene asociadas. MPEG-21, al igual que MPEG-7, no tiene ánimo de reinventar la rueda sino todo lo contrario. La filosofía de su trabajo es emplear aquellas tecnologías que estén funcionando y hayan demostrado su utilidad para emplearlas como partes del estándar.

Siguiendo esta filosofía se permite que cualquier identificador con forma de *URI (Uniform Resource Identifier)* pueda emplearse como identificador para un objeto digital. La especificación así mismo proporciona la posibilidad de tener sistemas de registro de identificadores a través de autoridades de registro. También se recoge el caso en el que usuarios diferentes empleen esquemas de identificación distintos para identificar sus contenidos, empleando el mecanismo de espacios de nombre de XML como solución.

Propiedad Intelectual

El objetivo de MPEG-21 es salvaguardar los derechos de autor y de copia para poder establecer un mercado digital en el que los productores de contenidos puedan confiar, en el que sus obras no puedan ser utilizadas sin la debida autorización o el correspondiente pago. Para ello existe una parte del estándar denominado *gestión y protección de la propiedad intelectual*.

Aunque estos sistemas ya se empezaron a desarrollar con MPEG-4, aparecieron problemas para su utilización. Podía darse el caso de que dispositivos creados por diferentes fabricantes podían resultar incompatibles entre si. En MPEG-21 se decidió comenzar un nuevo sistema de control más interoperable. La estandarización incluye los sistemas para recibir información de localizaciones remotas, intercambiar mensajes entre herramientas de gestión de propiedad intelectual y entre las herramientas y los terminales. Otro mecanismo que se incluye en la *gestión y protección de la propiedad intelectual* es la integración de las herramientas con el *Lenguaje de Expresión de Derechos*.

Una de las quejas más frecuentes por parte de los usuarios es la complicación del lenguaje legal a la hora de acceder a un contenido o descargar un programa. Las páginas y páginas de licencia, en general, son vistas como un incordio. El *Lenguaje de Expresión de Derechos* tiene dos objetivos. En primer lugar simplificar la comprensión de las licencias de recursos

digitales en su publicación, distribución y consumo. Por otra parte, intenta proporcionar mecanismos flexibles y interoperables para conseguir la gestión de la propiedad intelectual de forma eficiente y totalmente transparente al usuario.

El modelo que utiliza este lenguaje consta de cuatro entidades básicas, relacionadas entre sí. La relación entre estas entidades presentan las cláusulas que rigen un acuerdo entre las dos partes involucradas en una transacción. Las cuatro entidades son:

Principal: a quien pertenece un determinado derecho.

Derecho: que se especifica en la cláusula.

Recurso: al que está asociado el derecho.

Condición: que debe cumplirse antes de que se pueda ejecutar el derecho.

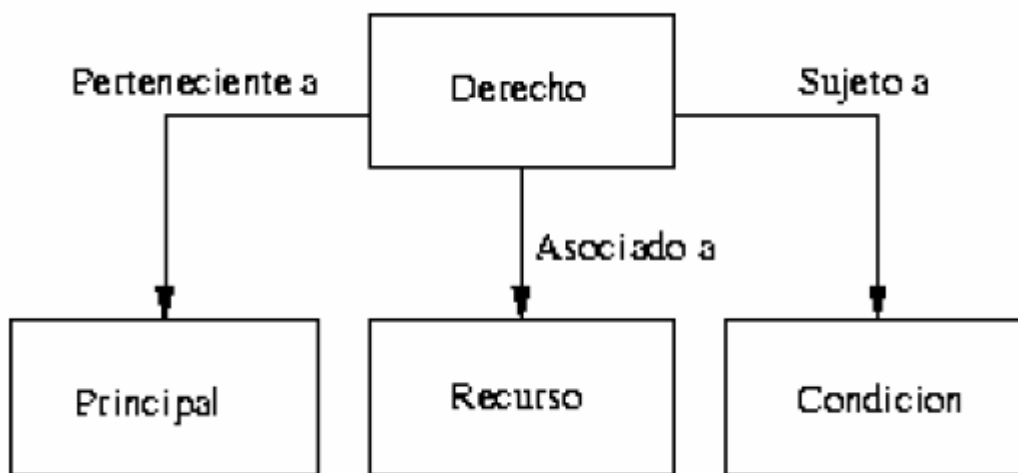


Figura 53: Relación entre las entidades que define el Lenguaje de Expresión de Derechos.

Normalmente los *derechos* especifican acciones que se pueden llevar a cabo o no por el *principal* con un determinado contenido. Las *condiciones* simples habitualmente indican un intervalo de tiempo durante el cual el *principal* puede ejercitar un *derecho* un determinado *recurso*.

Condiciones más elaboradas pueden incluir la existencia de determinados prerrequisitos, como un pago, para la obtener un *derecho*.

Con la aplicación de estas cláusulas se pretende proteger los contenidos de una incorrecta utilización. Cuando un usuario acceda a un registro se comprobará si cumple todas las

condiciones para ejercitar cada una de las acciones, o *derechos* que intente. Los dispositivos MPEG-21 se encargarán de limitar el acceso en caso de ser necesario.

Al igual que en MPEG-4 y MPEG-7, MPEG-21 emplea XML como herramienta para definir partes del estándar. El *Lenguaje de Expresión de Derechos* está basado en XML. Para más información en las referencias se incluye un enlace al *working draft* del mismo.

Como complemento al lenguaje de expresión de derechos, se está trabajando a su vez en el *Diccionario de Términos sobre Derechos*. Éste se compone de un conjunto claro, consistente y estructurado de términos sobre los que se soporta el Lenguaje. Con el las referencias a determinados términos dejarán de ser ambiguas, habilitando un método sencillo para desarrollar cláusulas de uso que acompañarán a los contenidos. Con la estandarización de términos y entidades, se facilita tanto la redacción de condiciones como la comprensión y aceptación de las mismas.

Adaptación de Objetos Digitales

La multiplicidad de dispositivos lleva a MPEG-21 a buscar la máxima transparencia para los usuarios a la hora de acceder a la red. Los contenidos que quiere consumir un determinado usuario deben llegar hasta él adecuados a las peculiaridades de su situación, tanto en prestaciones del método de presentación (limitaciones de pantalla, etc.) como de ancho de banda. El objetivo principal de los terminales y las redes de comunicación es alcanzar acceso transparente a contenido multimedia avanzado y distribuido, aislando a los usuarios de la instalación, gestión e implementación de los sistemas.

La adaptación de los contenidos pasa por la adecuación de los objetos digitales según las capacidades disponibles por cada usuario en cada momento o según las calidades contratadas.

MPEG-21 no especifica de manera directa cuales deben ser los métodos de adaptación, únicamente los descriptores y los mecanismos, independientes de formatos digitales particulares, que proporcionan soporte a la adaptación de objetos digitales en términos de recursos, descriptores y/ o gestión de la calidad de servicio. [10]

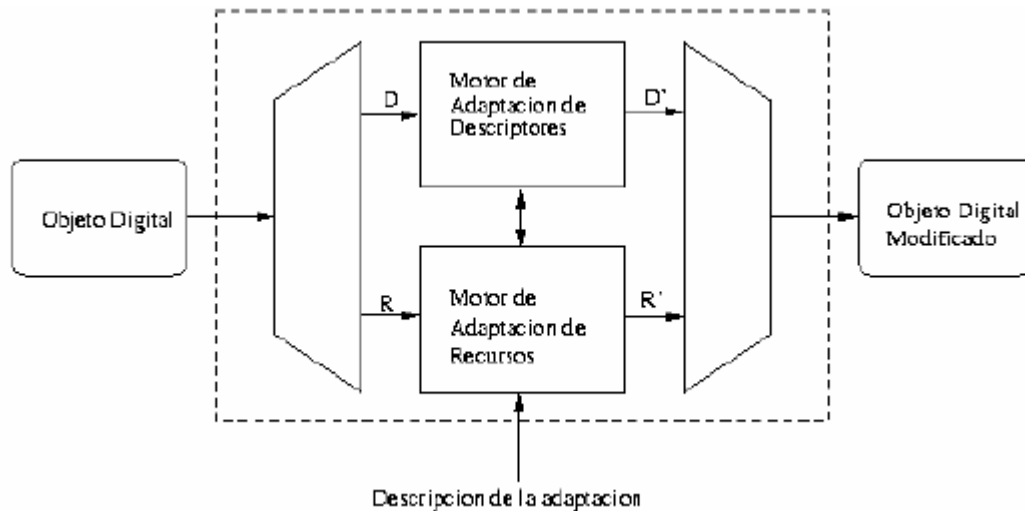


Figura 54: Concepto de la adaptación de objetos digitales.

3.7 MPEG AUDIO

3.7.1 El estándar MPEG audio.

El estándar MPEG Audio contempla tres niveles diferentes de codificación-decodificación de la señal de audio, de los cuales sólo el primero está totalmente terminado. Los otros dos son aplicables, y de hecho se utilizan habitualmente, pero siguen abiertos a ampliaciones. Estos tres niveles son:

- MPEG-1: "Codificación de imágenes en movimiento y audio asociado para medios de almacenamiento digital hasta 1'5 Mbit/s"
- MPEG-2: "Codificación genérica de imágenes en movimiento e información de audio asociada"
- MPEG-3: la planificación original contemplaba su aplicación a sistemas HDTV; finalmente fue incluido dentro de MPEG-2.
- MPEG-4: "Codificación de objetos audiovisuales"

A su vez, MPEG describe tres esquemas de codificación de audio denominados esquema-1, esquema-2 y esquema-3. Del primero al tercero aumentan tanto la complejidad del codificador como la calidad del sonido. Los tres son compatibles jerárquicamente, esto es, el decodificador esquema-i es capaz de interpretar información producida por un

codificador esquema-i y todos los niveles por debajo del i. Así, un decodificador esquema-3 acepta los tres niveles de codificación, mientras el esquema-2 sólo acepta el 1 y el 2.

MPEG define, para cada esquema, el formato del bitstream y el decodificador (que puede ser implementado de diferentes maneras). Con vistas a admitir futuras mejoras no se define el codificador, pero en un apartado informativo se da un ejemplo de codificador para cada uno de los esquemas. Hay que decir que tanto MPEG-1 como MPEG-2 emplean estos tres esquemas, pero este último añade nuevas características.

3.7.2 Introducción al sistema MPEG-1.

Este es el sistema que describe la norma ISO en lo referente al sistema MPEG-1:

Codificación: el codificador procesa la señal de audio digital y produce el bitstream empaquetado para su almacenamiento y/o transmisión. El algoritmo de codificación no está determinado, y puede utilizar enmascaramiento, cuantización variable y escalado. Sin embargo, debe ajustarse a las especificaciones del decodificador.

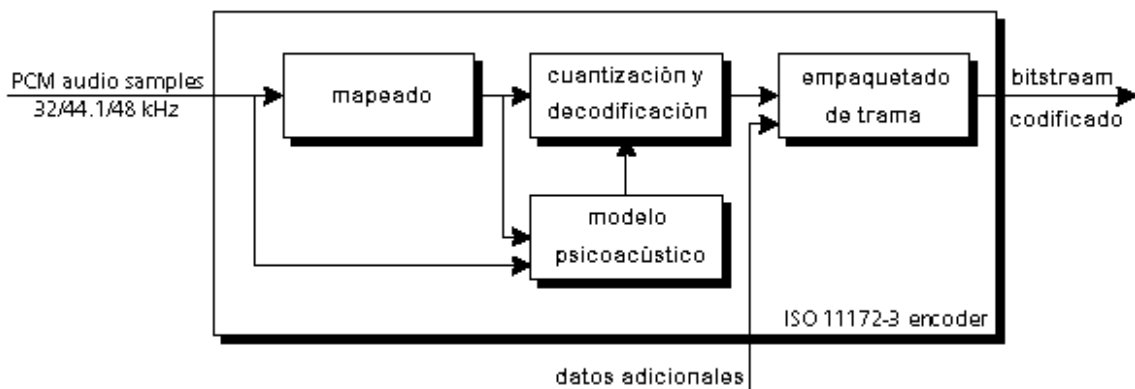


Figura 55: codificador según la norma ISO 11172-3

Las muestras se introducen en el codificador y a continuación el mapeador crea una representación filtrada y submuestreada de la señal de entrada. Las muestras mapeadas se

denominan tanto muestras de subbanda (esquemas 1 y 2) como muestras de subbanda transformadas (esquema 3). El modelo psicoacústico crea una serie de datos (dependiendo de la implementación del codificador) que sirven para controlar la cuantización y codificación. Este último bloque crea a su vez su propia serie de datos, de nuevo dependiendo de la implementación. Por último, el bloque de empaquetamiento de trama se encarga de agrupar como corresponde todos los datos, pudiendo añadir algunos más, llamados datos adicionales, como por ejemplo CRC o información del usuario.

Decodificación: el decodificador debe procesar el bitstream para reconstruir la señal de audio digital. La especificación de este elemento sí esta totalmente definida y debe seguirse en todos sus puntos. La figura ilustra el esquema del decodificador.

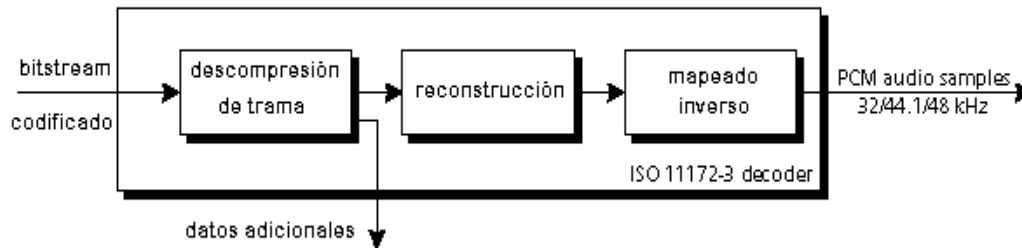


Figura 56: decodificador según la norma ISO 11172-3

Los datos del bitstream son desempaquetados para recuperar las diversas partes de la información. El bloque de reconstrucción recompone la versión cuantizada de la serie de muestras mapeadas. El mapeador inverso transforma estas muestras de nuevo a PCM.

Esquemas:

1. Incluye la división del mapeado básico de la señal de audio digital en 32 subbandas, segmentación para el formateo de los datos, modelo psicoacústico y cuantización fija. El retraso mínimo teórico es de 19 ms.
2. Incluye codificación adicional, factores de escala y diferente composición de trama. El retraso mínimo teórico es de 35 ms.
3. Incluye incremento de la resolución en frecuencia, basado en el uso de un banco de filtros híbrido. Cuantización no uniforme, segmentación adaptativa y codificación entrópica de los valores cuantizados. El retraso mínimo teórico es de 59 ms.

Esquema	Objetivo	Compresión	Calidad 64 kbps	Calidad 128 kbps	Retardo
Esquema-1	192 kbps	4 a 1			19 ms
Esquema-2	128 kbps	6 a 1	2'1 a 2'6	Más de 4	35 ms
Esquema-3	64 kbps	12 a 1	3'6 a 3'8	Más de 4	59 ms

Tabla 12: resumen de datos de los tres esquemas

La calidad viene dada del 1 al 5, siendo el 5 la superior. Hay que señalar que pese a los números de la norma ISO, el retraso típico acostumbra a ser tres veces mayor en la práctica.

Modos: hay cuatro modos de funcionamiento para cualquiera de estos tres esquemas.

- a. single channel o canal único: una señal en un bitstream.
- b. dual channel o canal doble: dos señales independientes en un mismo bitstream.
- c. stereo: como el anterior, perteneciendo las señales al canal izquierdo y derecho de una señal estéreo original.
- d. joint stereo: como el anterior, aprovechando ciertas características del estéreo como irrelevancia y redundancia de datos para reducir la tasa de bits.

3.7.3 MPEG-1 en detalle.

Tras haber visto la introducción que figura en los documentos ISO, podemos pasar a analizar en detalle el funcionamiento del sistema. A continuación haremos hincapié en las características y diferencias entre los tres esquemas de MPEG-1.

La codificación:

1. El banco de filtros: realiza el mapeado del dominio del tiempo al de la frecuencia. Existen dos tipos: el polifase y el híbrido polifase/MDCT. Estos bancos proporcionan tanto la separación en frecuencia para el codificador como los filtros de reconstrucción del decodificador. Las muestras de salida del banco están cuantizadas.
2. El modelo psicoacústico: calcula el nivel a partir del cual el ruido comienza a ser perceptible, para cada banda. Este nivel se utiliza en el bloque de asignación de bit/ruido para determinar la cuantización y sus niveles. De nuevo, se utilizan dos diferentes. En ambos, los datos de salida forman el SMR (signal-to-mask ratio) para cada banda o grupo de bandas.
3. Asignación de bit/ruido: examina tanto las muestras de salida del banco de filtros como el SMR proporcionado por el modelo psicoacústico, y ajusta la asignación de bit o ruido, según el esquema utilizado, para satisfacer simultáneamente los requisitos de tasa de bits y de enmascaramiento.
4. El formateador de bitstream: toma las muestras cuantizadas del banco de filtros, junto a los datos de asignación de bit/ruido y otra información lateral para formar la trama.

Los tres esquemas utilizan diferentes algoritmos para cumplir estas especificaciones:

Esquema I:

- El mapeado tiempo-frecuencia se realiza con un banco de filtros polifase con 32 subbandas. Los filtros polifase consisten en un conjunto de filtros con el mismo ancho de banda con interrelaciones de fase especiales que ofrecen

una implementación eficiente del filtro subbanda. Se denomina filtro subbanda al que cubre todo el rango de frecuencias deseado. En general, los filtros polifase combinan una baja complejidad de computación con un diseño flexible y múltiples opciones de implementación.

- El modelo psicoacústico utiliza una FFT (Fast Fourier Transform) de 512 puntos para obtener información espectral detallada de la señal. El resultado de la aplicación de la FFT se utiliza para determinar los enmascaramientos en la señal, cada uno de los cuales produce un nivel de enmascaramiento, según la frecuencia, intensidad y tono. Para cada subbanda, los niveles individuales se combinan y forman uno global, que se compara con el máximo nivel de señal en la banda, produciendo el SMR que se introduce en el cuantizador.
- El bloque de cuantización y codificación examina las muestras de cada subbanda, encuentra el máximo valor absoluto y lo cuantiza con 6 bits. Este valor es el factor de escala de la subbanda. A continuación se determina la asignación de bits para cada subbanda minimizando el NMR (noise-to-mask ratio) total. Es posible que algunas subbandas con un gran enmascaramiento terminen con cero bits, es decir, no se codificará ninguna muestra. Por último las muestras de subbanda se cuantizan linealmente según el número de bits asignados a dicha subbanda concreta.
- El trabajo del empaquetador de trama es sencillo. La trama, según la definición ISO, es la menor parte del bitstream decodificable por sí misma. Cada trama empieza con una cabecera para sincronización y diferenciación, así como 16 bits opcionales de CRC para detección y corrección de errores. Se emplean, para cada subbanda, 4 bits para describir la asignación de bits y otros 6 para el factor de escala. El resto de bits en la trama se utilizan para la información de samples, 384 en total, y con la opción de añadir cierta información adicional. A 48 Khz, cada trama lleva 8 ms de sonido.

Esquema II:

- El mapeado de tiempo-frecuencia es idéntico al del esquema I.
- El modelo psicoacústico es similar, salvo que utiliza una FFT de 1024 puntos para obtener mayor resolución espectral. En los demás aspectos, es idéntico.
- El bloque de cuantización y codificación también es similar, generando factores de escala de 6 bits para cada subbanda. Sin embargo, las tramas del esquema II son tres veces más largas que las del esquema I, de forma que se concede a cada subbanda tres factores de escala, y el codificador utiliza uno, dos o los tres, según la diferencia que haya entre ellos. La asignación de bits es similar a la del esquema I.

El formateador de trama: la definición ISO de trama es la misma que en el punto anterior. Utiliza la misma cabecera y estructura de CRC que el esquema I. El número de bits que utilizan para describir la asignación de bits varía con las subbandas: 4 bits para las inferiores, 3 para las medias y dos para las superiores, adecuándose a las bandas críticas. Los factores de escala se codifican junto a un número de dos bits que indica si se utilizan uno, dos o los tres. Las muestras de subbanda se cuantizan y a continuación se asocian en grupos de tres, llamados gránulos. Cada uno se codifica con una palabra clave, lo que permite interceptar mucha más información redundante que en el esquema I. Cada trama contiene, pues, 1152 muestras PCM. A 48 KHz. cada trama lleva 24 ms de sonido.

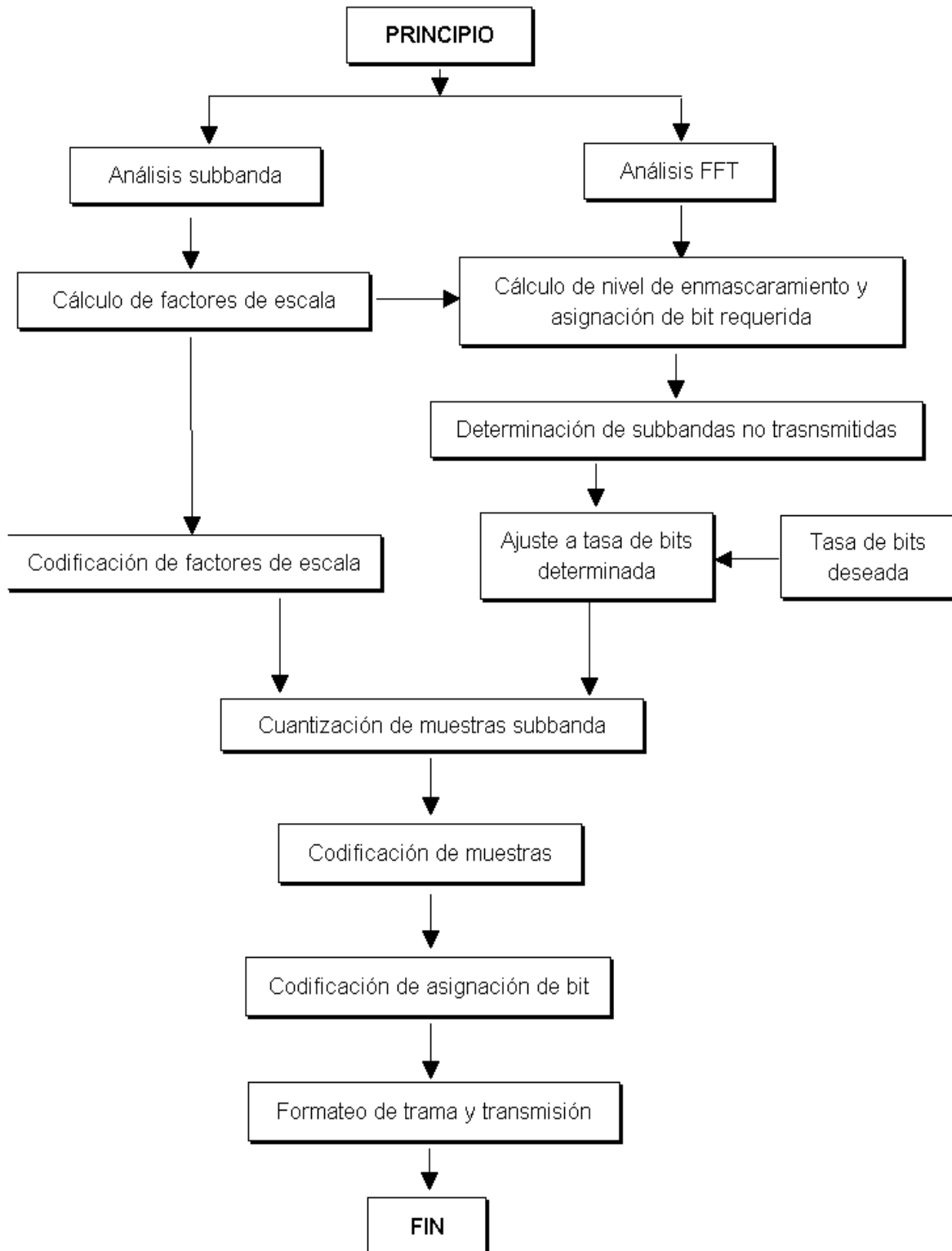


Diagrama 57: diagrama de flujos del codificador para esquema-1 y esquema-2 según ISO 11172-3

Esquema III:

- El esquema III es sustancialmente más complicado que los dos anteriores e incluye una serie de mejoras cuyo análisis resultaría desbordante, de manera que no entraremos en tantos detalles. Su diagrama de flujos es conceptualmente semejante al visto para los otros dos esquemas, salvo que se realizan múltiples iteraciones para procesar los datos con el mayor nivel de calidad en un cierto tiempo, lo cual complica su diseño hasta el punto de que los diagramas ISO ocupan decenas de páginas.
- El mapeado de tiempo-frecuencia añade un nuevo banco de filtros, el DCT (Discrete Cosine Transform), que con el polifase forman el denominado filtro híbrido. Proporciona una resolución en frecuencia variable, 6x32 o 18x32 subbandas, ajustándose mucho mejor a las bandas críticas de las diferentes frecuencias.
- El modelo psicoacústico es una modificación del empleado en el esquema II, y utiliza un método denominado predicción polinómica. Incluye los efectos del enmascaramiento temporal.
- El bloque de cuantización y codificación también emplea algoritmos muy sofisticados que permiten tramas de longitud variable. La gran diferencia con los otros dos esquemas es que la variable controlada es el ruido, a través de bucles iterativos que lo reducen al mínimo posible en cada paso.
- El formateador de trama: la definición de trama para este esquema según ISO varía respecto de la de los niveles anteriores: "mínima parte del bitstream decodificable mediante el uso de información principal adquirida previamente". Las tramas contienen información de 1152 muestras y empiezan con la misma cabecera de sincronización y diferenciación, pero la información perteneciente a una misma trama no se encuentra generalmente entre dos cabeceras. La longitud de la trama puede variarse en caso de necesidad. Además de tratar con esta información, el esquema III incluye

codificación Huffman de longitud variable, un método de codificación entrópica que sin pérdida de información elimina redundancia. Los métodos de longitud variable se caracterizan, en general, por asignar palabras cortas a los eventos más frecuentes, dejando las largas para los más infrecuentes.

La decodificación:

Es mucho más sencilla que la codificación, de manera que con lo ya comentado en partes anteriores basta para seguir los siguientes diagramas ISO que incluyen algunas notas aclaratorias al margen que no forman parte de las figuras originales de la norma. [11]

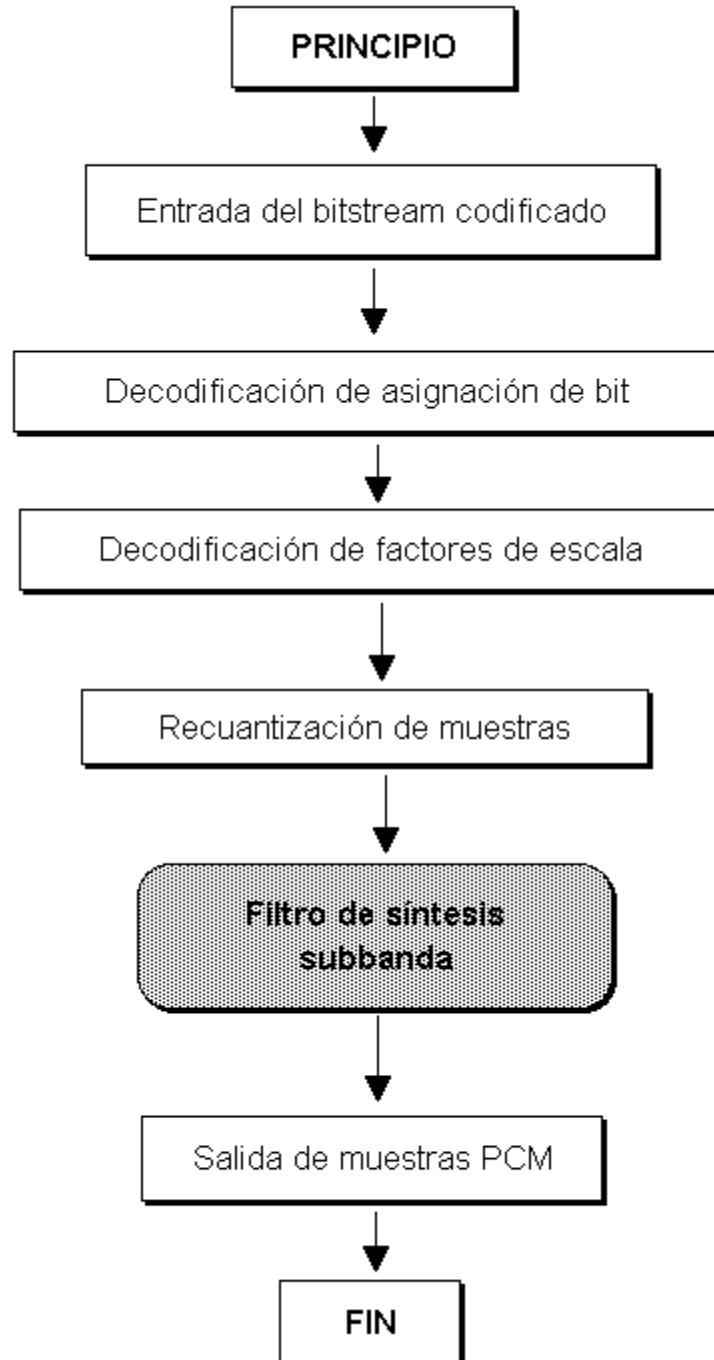


Diagrama 58: diagrama de flujos del decodificador para esquema-1 y esquema-2 según ISO 11172-3

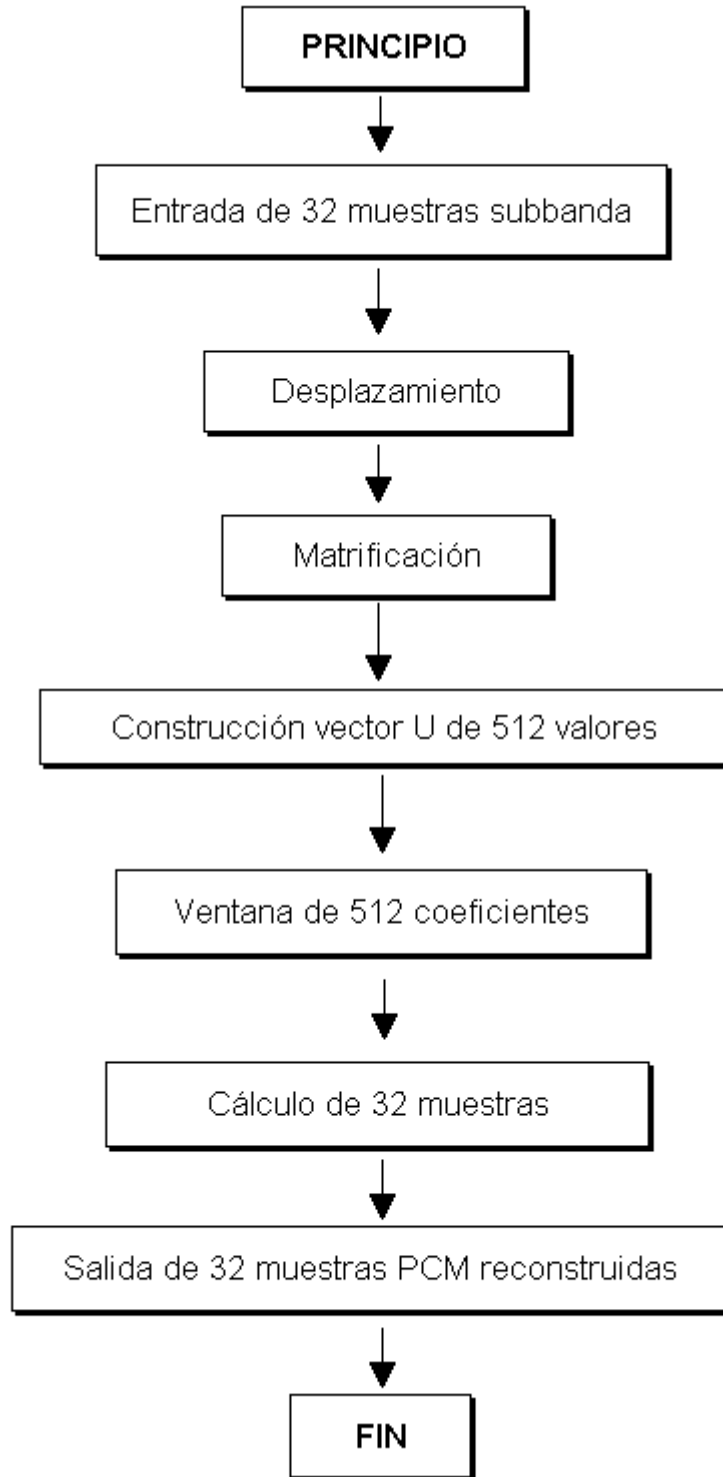


Diagrama 59: detalle del filtro del diagrama 2

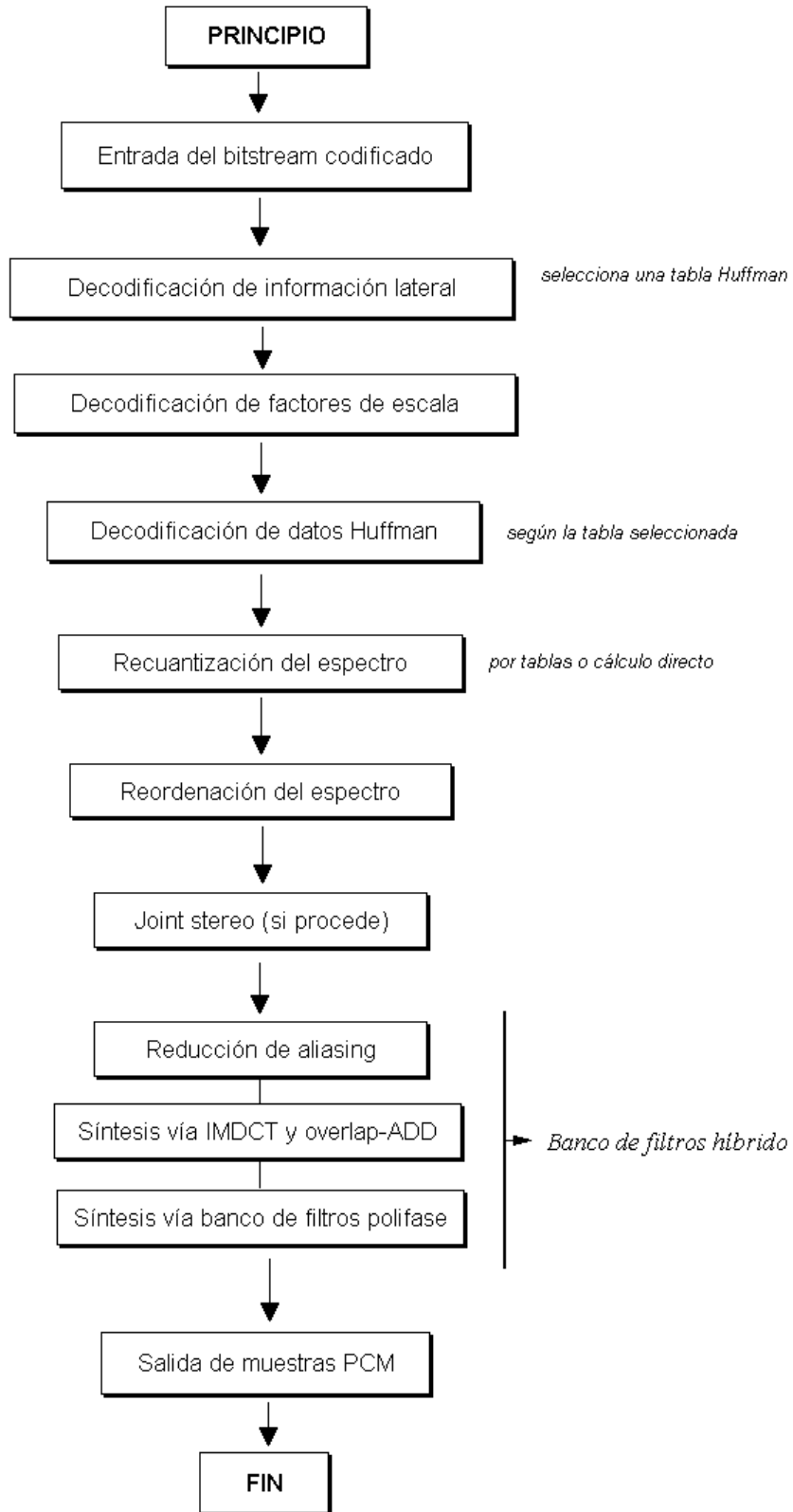


Diagrama60:diagrama de flujos del decodificador para esquema-3 según ISO 11172-3

3.7.4 TRAMA DE AUDIO

Modelo de sistema de codificación de audio MPEG:

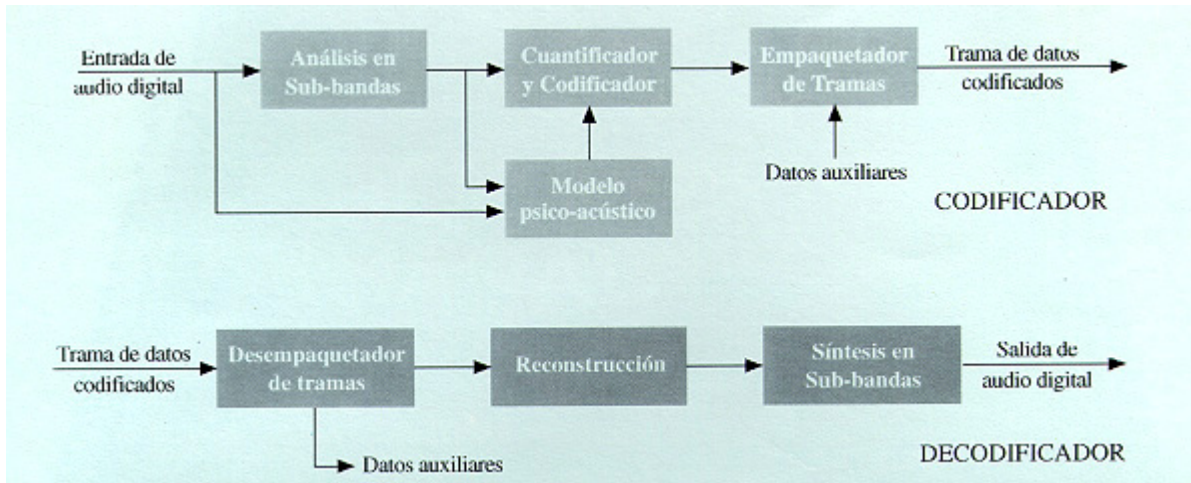


Fig. 61. Modelo de sistema de codificación de audio MPEG.

Trama de audio MPEG-1

La codificación de audio MPEG-1 proporciona un método de compresión de alta calidad. El algoritmo no aprovecha la redundancia de información de la señal de audio, sino que se basa en un modelo de percepción del oído humano. Se determinan qué señales están enmascaradas, para poder eliminarlas sin que se produzca una pérdida apreciable en la calidad de la señal acústica percibida.

Principales características del algoritmo MPEG-1:

Frecuencias de muestreo de 32, 44.1 y 48 kHz, con 16 bits por muestra.

Uno o dos canales de audio con cuatro modos de funcionamiento: *mono*, *dual*, *estéreo* y *join estéreo*.

Velocidades binarias de salida de 32 a 192 Kb/s por canal.

Hay tres capas (*Layer I*, *II* y *III*) en el algoritmo, que ofrecen, en orden creciente, una mayor calidad a costa de una mayor complejidad de los procesos de codificación y decodificación.

La trama puede incluir *datos auxiliares* además de la información de audio.

En el diagrama de bloques del modelo de sistema de codificación de audio MPEG-1 se puede ver el proceso básico de codificación/decodificación de la señal de audio digital. Esta señal se divide en 32 subbandas de frecuencia con un banco de filtros. Simultáneamente se calcula para cada subbanda, en el modelo psicoacústico, los coeficientes entre los niveles de señal y umbral de enmascaramiento (los sonidos no se distinguen por debajo de este umbral). A cada banda se le asigna un número de bits para reducir el ruido total de cuantificación, esto se realiza en el bloque de asignación de bits y cuantificación. El último bloque empaqueta la información anterior y añade los datos auxiliares en la trama.

En el proceso de decodificación de la señal de audio MPEG-1 se realizan las funciones inversas a las del proceso de codificación.

Trama de audio MPEG-2

Algunas de las funciones nuevas aportadas son:

Audio multicanal. Se introduce el *sonido envolvente* (surround) análogo al de las salas cinematográficas. El sistema está formado por 3 canales frontales (izquierdo, central y derecho), 2 canales traseros (izquierdo y derecho) y un canal de refuerzo de graves.

Soporte para *audio multilingüe*. Se pueden introducir en la trama hasta 8 canales de voz.

Nuevas frecuencias de muestreo (16, 22.05 y 24 kHz) y valores más bajos de flujo binario (hasta 16 Kb/s). Así se puede utilizar para servicios que requieran menos calidad de audio.

La trama MPEG-2 se construye en torno a dos canales principales (izquierdo y derecho), para que sea compatible con la trama MPEG-1, el resto de información necesaria para los canales surround o los canales multilingües se incluye en la zona de datos auxiliares. Así un decodificador que siga la norma MPEG-1 obtendrá los canales izquierdo y derecho, mientras que un decodificador MPEG-2 puede obtener todos los canales.

3.7.5 MÁS SOBRE MPEG-1

El estándar combina partes de los algoritmos de codificación MUSICIAN y ASPEC. La codificación de audio de MPEG-1 ofrece una reproducción con calidad subjetiva equivalente a la del CD (16-bit PCM). Y gracias al gran rango dinámico con el que trabaja MPEG-1, tiene suficiente potencial para mejorar la calidad del CD.

Estructura

La estructura básica de los codificadores de audio MPEG-1 siguen la de los codificadores basados en la percepción. En primer lugar, la señal de audio es convertida a componentes espectrales por medio de un banco de filtros de análisis; los niveles I y II hacen uso de un banco de filtros para una banda determinada y el nivel III utiliza un banco de filtros híbrido. Cada componente espectral es cuantificada y codificada con el fin de conservar un ruido de cuantificación por debajo de cierto límite. El número de bits de cada subbanda y el factor de escala son determinados por bloques. Cada bloque tiene 12 (nivel I) o 36 (nivel II y III) muestras de la subbanda. El número de bits cuantificados es obtenido a partir de un 'algoritmo de asignación de bits' (niveles I y II) que es controlado por un 'modelo psicoacústico'. Las palabras código, el factor de escala y la información de asignación de bits, son multiplexados en una trama, a la que se añade una cabecera y datos opcionales. En el decodificador, el banco de filtros reconstruye un bloque de 32 muestras de audio a partir de la trama demultiplexada.

MPEG-1/Audio soporta frecuencias de muestreo de 32, 44.1 y 48 kHz y velocidades binarias entre 32 kb/s (mono) y 448 kb/s, 348 kb/s y 320 kb/s (estéreo y niveles I, II, y III, respectivamente).

Niveles y modelos de operacion

El estándar consta de tres niveles (I, II y III). Desde el punto de vista de hardware y software, el más alto de los niveles incorpora el principal bloque para la construcción de los niveles más bajos. Un decodificador de audio MPEG-1 es capaz

de decodificar tramas a partir de los tres niveles. Los más típicos son los decodificadores MPEG-1/Audio nivel X (X=I, II o III).

Codificación estereo

Mpeg-1/Audio soporta cuatro modos: mono, estéreo, dual con dos canales diferentes (útil para programas bilingües) y '*joint*' estéreo.

En el modo '*joint*' estéreo opcional, las dependencias entre canales son explotadas para reducir la velocidad binaria mediante una técnica de reducción llamada 'estéreo de intensidad'. Es sabido que, por encima de los 2 kHz y en una banda crítica alrededor de esta frecuencia, el sistema auditivo humano basa su percepción del estéreo más en el entorno temporal de la señal de audio que en su estructura temporal fina. Por eso, el algoritmo de compresión de MPEG audio soporta un código de redundancia para modo estéreo llamado codificación de 'estéreo de intensidad'. Lo cual reduce la velocidad binaria total sin violar la integridad espacial de la señal estereofónica.

En este modo, el codificador codifica alguna salidas de la subbanda de alta frecuencia con la señal L+R (o una combinación lineal de ellas) en vez de enviar independientemente las señales L (canal izquierdo) y R (canal derecho). El decodificador reconstruye los canales izquierdo y derecho basándose sólo en la señal L+R y en la independencia de los factores de escala de los canales izquierdo y derecho. La forma espectral de las salidas izquierda y derecha es la misma alrededor de cada banda de codificación de intensidad, pero las magnitudes son diferentes. El opcional modo '*joint*' estéreo sólo será efectivo si la velocidad binaria requerida excede la disponible, y será aplicada a bandas correspondientes a frecuencias alrededor de los 2 kHz y superiores.

El nivel III tiene una opción adicional: en el modo mono/estéreo las señales de los canales izquierdo y derecho son codificadas como canales (L+R) y (L-R). Este último modo puede ser combinado con el modo '*joint*' estéreo.

Modelo psicoacústico

Este modelo controla el algoritmo de asignación de bit. Utiliza la parte baja del espectro del bloque de audio que va a ser codificado (la cual puede obtenerse siguiendo diferentes caminos), y el conocimiento de la máscara de ruido. El modelo sólo es necesario en la codificación, lo cual hace el decodificador menos complejo; esta asimetría es un hecho deseable para las aplicaciones de *play-back* y difusión de audio.

La parte normativa del estándar describe el decodificador y el significado de la codificación de la trama, pero el codificador no está estandarizado.

Trama y estructura del múltiplex

- Estructura de la trama.

Cada trama tiene una cabecera; su primera parte contiene 12 bits de sincronización, 20 de información del sistema y 16 opcionales de redundancia cíclica. Su segunda parte contiene la información sobre la asignación de bits y el factor de escala. Como información principal, una trama lleva un total de 32x12 muestras (correspondiendo a 384 muestras de entrada codificadas en PCM- equivalente a 8 ms para una velocidad binaria de 48 kHz) para el nivel I, y un total de 32x36 muestras en el nivel II (correspondiendo a 1152 muestras de entrada codificadas en PCM- equivalente a 24 ms para una velocidad binaria de 48 kHz).

Cabe destacar que las tramas son autónomas: cada trama contiene toda la información necesaria para su decodificación, por tanto, cada una de ellas puede ser decodificada independientemente de las tramas anteriores (esto define un punto de partida para aplicaciones de almacenamiento y edición de audio). La longitud de las tramas tampoco es fija, sino que depende de:

- la longitud del campo principal de información, el cual depende de la velocidad binaria y de la frecuencia de muestreo

- la parte del campo de información que varía en el nivel II
- una parte del campo de datos, cuya longitud no está especificada.
- Estructura múltiplex.

El campo variable de las tramas MPEG es fragmentado en paquetes. La estructura de los paquetes consta de 188 bytes, de los cuales 4 son de cabecera, y los 184 restantes de carga. La cabecera incluye un byte de sincronismo, y un campo de 13 bits llamado identificador de paquete, el cual informa al decodificador sobre el tipo de datos, e información adicional.

La combinación de audio, vídeo y la trama de datos no necesita ser determinada. La combinación debe cambiar dinámicamente; y los servicios pueden ser proporcionados de modo flexible. Si se requiere información adicional en la cabecera como sincronización periódica de audio y vídeo, puede ser usado un sistema de longitud variable llamado adaptación de cabecera como parte del campo de 184 bytes transportado.

Aunque la longitud de las tramas no es fija, el intervalo entre las cabeceras de estas es constante mediante el uso de protección de bytes. La especificación del sistema MPEG describe la compresión de tramas de datos de audio y vídeo; cuya información tiene que ser multiplexada junta para formar una trama simple de datos.

3.7.6 MÁS SOBRE MPEG-2

La segunda fase del MPEG, llamado MPEG-2, incluye en su parte de audio dos estándares de codificación de audio multicanal; ambos son compatibles con MPEG-1, uno hacia delante y otro hacia atrás. La compatibilidad hacia delante significa que un decodificador multicanal en MPEG-2 es capaz de decodificar señales MPEG-1 mono o estereofónicas; la compatibilidad hacia atrás significa que existen decodificadores estéreo MPEG-1, los cuales sólo manejan dos canales de audio y son capaces de reproducir una señal estéreo básica desde una trama multicanal MPEG-2 para servir la necesidad de los usuarios con equipos mono o estéreo. Los codificadores multicanal no compatibles hacia atrás no serán

capaces de alimentar una trama de un decodificador estéreo MPEG-1. Por otro lado, estos tienen más libertad en la reproducción de alta calidad de las señales de audio.

Con la compatibilidad hacia atrás es posible introducir multicanales audio al mismo tiempo sin la necesidad de que existan decodificadores de dos canales estéreo, los cuales están ya obsoletos. Un importante ejemplo es el sistema de difusión digital de audio europeo, el cual requerirá decodificadores estéreo MPEG-1 en la primera generación pero puede ofrecer audio multicanal en una posterior aplicación.

Compatibilidad hacia atrás

La compatibilidad hacia atrás implica el uso de la compatibilidad de matrices. Un grupo de cinco canales desarrolla una señal estéreo básica 2/0, consistente en un canal de la derecha y otro de la izquierda, L0 y R0, respectivamente. Estos son transmitidos en formato MPEG-1 a través de dos de los canales. Los otros tres forman la señal de extensión multicanal, y deben ser elegidos para que el decodificador pueda reconstruir la señal multicanal 3/2-estéreo completa. Las redundancias entre canales y los efectos de la máscara son tomados en cuenta para encontrar la elección óptima. En MPEG-2 las matrices pueden ser hechas de diferentes modos, e incluso dependientes con el tiempo. Sin embargo, el contenido de la señal extendida de audio está ya desarrollado en la trama de audio MPEG-1; esta redundancia reduce la tasa de compresión.

La 'matricidad' es necesaria para proporcionar compatibilidad hacia atrás; sin embargo, si es usada con la codificación perceptiva, puede aparecer un ruido de cuantificación no enmascarable. Ello puede producirse en el proceso inverso a la construcción de matrices, cuando se forman las señales suma y diferencia. En ciertas situaciones cada componente de estas señales puede desaparecer en uno de los canales. Si esta componente fuese incluida en la máscara del ruido de cuantificación del e canal, este ruido puede llegar a ser audible. Como medida adicional, el modo de velocidad binaria variable opcional puede ser evocado para codificar el contenido problemático para altas velocidades binarias.

Los decodificadores MPEG-1 tienen limitación de velocidad (384 kb/s en el nivel II). Para superar esta limitación, el estándar MPEG-2 permite el uso de una segunda trama, lo cual proporciona compatibilidad multicanal audio a altas velocidades.

Codificación MPEG-2 audio avanzado

Existe un segundo estándar que soporta aplicaciones que no requieren compatibilidad con el formato MPEG-1 estéreo. De este modo, la codificación y decodificación de matrices no es necesaria.

- Herramientas:

El estándar MPEG-2 AAC emplea filtros de alta resolución, técnicas de predicción, y codificación de bajo ruido. Todo esto está basado en evaluaciones recientes de herramientas (o módulos), las cuales han sido seleccionadas a partir de un número de propuestas. El propio contenido de estas incluye un pre-procesado opcional, un banco de filtros, un modelo de percepción, un modelado temporal del ruido, una codificación de intensidad multicanal, predicción, codificación estéreo, cuantificación, codificación sin ruido, y un multiplexor. El banco de filtros es la transformada del coseno discreto modificado en 1024 líneas, y el modelo de percepción está tomado de MPEG-1. La herramienta 'forma temporal del ruido' controla la dependencia con el tiempo del ruido de cuantificación, la intensidad, y la codificación mono/estéreo; y el predictor adaptativo hacia atrás de segundo orden incrementa la eficiencia de la codificación.

Éste reduce la velocidad binaria para la codificación de muestras en una banda dada, y basa su predicción en el espectro cuantificado de bloques previos, el cual sólo está disponible en el decodificador (en ausencia de errores en el canal). Finalmente, para la cuantificación y la codificación sin ruido, se utiliza un método iterativo para conservar el ruido de cuantificación, en las bandas críticas, por debajo del límite de la máscara global.

- Perfiles:

Para servir diferentes necesidades, el estándar proporciona tres perfiles:

- el perfil principal, que ofrece alta calidad
- el perfil de baja complejidad, que trabaja sin predicción. Éste no emplea modelado temporal del ruido ni predictor en el dominio del tiempo (la predicción añade una complejidad significativa al sistema).
- el perfil de frecuencia de muestreo variable, que ofrece la más baja complejidad. Éste utiliza un banco de filtros híbrido.

MPEG-2 AAC soporta 46 canales de varias configuraciones ruidosas multicanal, además de otras aplicaciones; por defecto, estas configuraciones son el canal monofónico, el canal estereofónico, y un sistema compuesto por otros cinco canales y el canal LFE.

La lista seleccionada de módulos o herramientas anteriormente descrita define el estándar MPEG-2 AAC audio, el cual se convirtió en un estándar internacional en Abril de 1997 como extensión de MPEG-2. El estándar ofrece alta calidad a las más bajas velocidades posibles, entre 320 y 384 kb/s para cinco canales; éste encontrará su aplicación tanto en usos profesionales como en el ámbito de consumidor. [12]

3.7.7 COMPRESIÓN DE AUDIO

Toda compresión de datos de audio se basa en la comprensión del mecanismo auditivo, por lo que constituye una forma de codificación perceptual. El oído es sólo capaz de extraer una cierta proporción de la información contenida en un determinado sonido. A esto se le puede denominar entropía perceptual, siendo redundante el sonido adicional. Un sistema ideal debe eliminar toda redundancia, dejando únicamente la entropía.

Existen muchos tipos diferentes de compresión de audio y cada uno permite un factor de compresión diferente. Algunas aplicaciones como DCC (Digital Compact Cassette) y DAB (Digital Audio Broadcasting) requieren un valor de 0.25. Para el Minidisco, es de 0.2. La

transmisión de audio por RDSI requiere más compresión todavía, que sólo puede realizarse empleando técnicas sofisticadas.

La codificación de la sub-banda imita el mecanismo de análisis en frecuencia del oído humano y divide el espectro de audio en un gran número de bandas de frecuencia diferentes con el fin de poder explotar el hecho de que la mayoría de las bandas contienen señales cuyo nivel es inferior al de la señal más alta. Las señales en estas bandas pueden ser entonces cuantificadas independientemente. El error de cuantificación que resulta es confinado a los límites de frecuencia de la banda y así este puede arreglarse para ser enmascarado por el material del programa. Las técnicas usadas en las capas 1 y 2 de MPEG audio son basadas en la codificación de la sub-banda como son aquéllas usadas en el DCC.

En la codificación por transformación la forma de onda de audio en el dominio del tiempo es convertida a una representación en el dominio de la frecuencia como una transformada de Fourier, Discreta del Coseno o Wavelet. La codificación por transformación toma ventaja del hecho de que la amplitud o cubierta de una señal de audio cambia relativamente despacio y para que así los coeficientes de la transformada puedan transmitirse relativamente con poca frecuencia. Claramente tal aproximación se estropea en presencia de transitorios y se requieren en la práctica sistemas adaptables. Los transitorios causan que los coeficientes puedan ser frecuentemente actualizados mientras que en las partes estacionarias de la señal como las notas sostenidas la tasa de actualización puede reducirse. La codificación por Transformada Discreta del Coseno se usa en la capa 3 (layer 3) de MPEG audio y en el sistema de compresión del Minidisco.

3.7.8 CODIFICACIÓN DE LA SUB-BANDA

La compresión de los datos de la sub-banda aprovecha el hecho de que los sonidos reales no tienen una energía espectral uniforme. La longitud de la palabra del audio PCM está basada en el rango dinámico requerido. Cuando una señal con un espectro no uniforme es transmitida por PCM, todo el rango dinámico es ocupado únicamente por la componente espectral más alta, y todas las demás componentes son codificadas con excesivo headroom (área entre el nivel normal de funcionamiento y el nivel de recorte). En su forma más simple, la codificación de la sub-banda funciona dividiendo la señal de audio en un número

de bandas de frecuencia y comprimiendo y expandiendo cada banda de acuerdo con su propio nivel. Las bandas en las que hay poca energía dan como resultado amplitudes pequeñas que pueden transmitirse con una longitud de palabra corta. Por tanto, cada banda se traduce en muestras de longitud variable, pero la suma de todas las longitudes de palabra de la muestra es inferior a la de la PCM, pudiendo obtenerse así una ganancia de codificación.

El número de sub-bandas que pueden utilizarse depende de qué otra técnica de compresión se vaya a combinar con la codificación de la sub-banda. Si se tiene la intención de utilizar la compresión basada en el enmascaramiento auditivo, es preferible que las sub-bandas sean más estrechas que las bandas críticas del oído y, por tanto, se requerirá un gran número; así, por ejemplo en MPEG y PASC (Precision Adaptive Subband Coding) se emplean 32 sub-bandas. La Figura 62 muestra la condición crítica en la que el tono del enmascaramiento se encuentra en el límite superior de la sub-banda. Se observará que cuanto más estrecha es la sub-banda, mayor es el ruido de recuantificación que puede enmascarse. No obstante, la utilización de un número excesivo de sub-bandas acentúa la complejidad y el retardo de codificación, y también arriesga el pre-eco en los transitorios que exceden el enmascaramiento temporal.

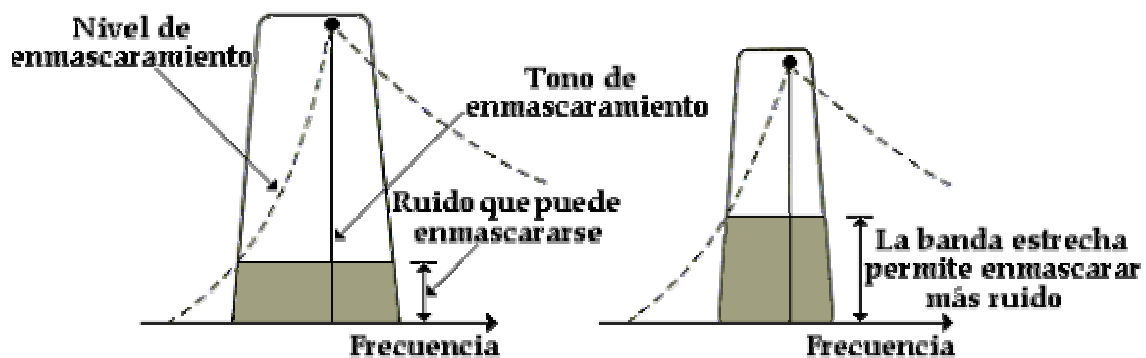


Figura 62. En la codificación de la sub-banda, el caso más desfavorable tiene lugar cuando el tono de enmascaramiento se encuentra en el límite superior de la sub-banda

Las Figuras 63 y 64 muestran los diagramas de bloques de un codificador y un decodificador de sub-bandas, respectivamente. En la entrada, el rango de frecuencias es dividido en sub-bandas mediante un banco de filtros tal como un filtro especular en cuadratura. Los datos descompuestos de la sub-banda se organizan en bloques de tamaño

fijo, antes del proceso de reducción. Aunque todas las sub-bandas pueden utilizar bloques de la misma longitud, algunos codificadores pueden utilizar bloques que se hacen más largos a medida que disminuye la frecuencia de la sub-banda. Los bloques de las sub-bandas también se denominan bins de frecuencia.

La ganancia de codificación se obtiene cuando la forma de onda de cada banda pasa a través de un recuantificador. La recuantificación se consigue multiplicando los valores de las muestras por una constante y redondeando el resultado por exceso o por defecto de acuerdo con la longitud de palabra requerida. Por ejemplo, si en una sub-banda determinada la forma de onda es de 36 dB o menos a fondo de escala, al menos habrá 6 bits en cada muestra que sólo reproduce el bit de signo. Multiplicando por 64, entrarán en uso los bits de orden superior de la muestra, permitiendo que se pierdan bits en el extremo más bajo al redondear a una longitud de palabra menor. Cuanto menor es la longitud de palabra, mayor es la ganancia de codificación, pero más toscos resultan los escalones de cuantificación y, por tanto, el nivel de error de cuantificación.

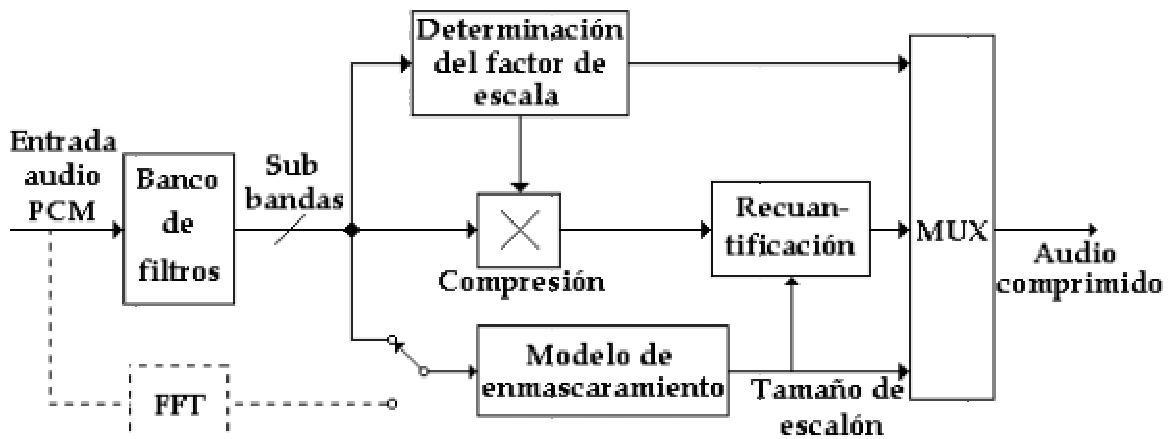


Figura 63. Diagrama de bloques de un codificador de sub-bandas

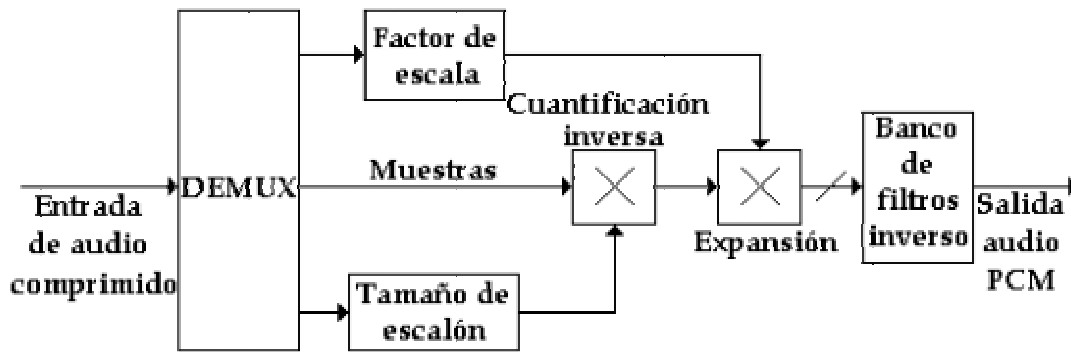


Figura 64. Diagrama de bloques de un decodificador de sub-bandas

3.7.9 CODIFICACIÓN POR TRANSFORMACIÓN

El análisis de Fourier permite representar cualquier forma de onda mediante un conjunto de componentes armónicamente relacionados de amplitud y fase adecuadas. La transformada de una forma de onda de audio típica varía de manera relativamente lenta. La lenta señal sonora procedente del tubo de un órgano o de la cuerda de un violín, o el lento decrecimiento de la mayoría de los sonidos musicales, permite la reducción de la frecuencia a la que la transformada es muestreada, obteniéndose una ganancia de codificación. Es posible obtener una ganancia de codificación adicional si las componentes que experimentarán el enmascaramiento se cuantifican de manera más rudimentaria.

Las transformadas prácticas requieren bloques de muestras en lugar de cadenas interminables. La solución está en cortar la forma de onda en cortos segmentos solapados y, seguidamente, transformar cada uno de ellos individualmente tal como indica la Figura 65. De este modo, cada muestra de entrada aparece en sólo dos transformadas, pero con una ponderación variable dependiendo de su posición en el eje temporal.



Figura 65. La codificación por transformación se realiza en la práctica en bloques cortos

La DFT (Discrete Frequency Transform o Transformada de Frecuencia Discreta) requiere gran número de cálculos, debido al requisito de tener que utilizar una aritmética compleja para obtener la fase de las componentes, así como la amplitud. Una alternativa consiste en

emplear la Transformada Discreta del Coseno (DCT). Esta presenta una ventaja cuando se utiliza con ventanas solapadas. En la Transformada Discreta del Coseno Modificada (MDCT), se usan ventanas con un solapamiento del 50%. De este modo, se obtiene el doble de coeficientes necesarios, que se submuestran por un factor de dos para obtener una transformada muestreada críticamente, lo cual tiene como resultado un efecto un aliasing potencial en el dominio de la frecuencia. Sin embargo, variando levemente la transformada, los productos de aliasing en la segunda mitad de una determinada ventana son iguales en tamaño, pero de polaridad opuesta a los productos de aliasing de la primera mitad de la siguiente ventana, por lo que se eliminarán en su reconstrucción. Éste es el principio de la eliminación del aliasing en el dominio temporal (TDAC, Time Domain Aliasing Cancellation).

La recuantificación realizada en el codificador eleva el ruido de cuantificación en el bin de la frecuencia, pero lo hace durante todo el tiempo que dura el bloque. La Figura 66 muestra que, si se produce un transitorio hacia el extremo final de un bloque, el decodificador reproduce la forma de onda correctamente, pero el ruido de cuantificación comenzará al principio del bloque y puede dar lugar a un pre-eco en el que el ruido se oye antes que el transitorio.

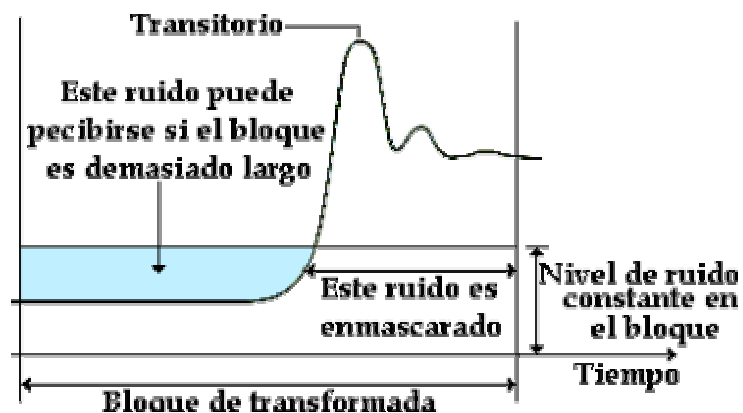


Figura 66. Transitorio en el final de un bloque de una transformada

La solución es utilizar una ventana de tiempo variable de acuerdo con el contenido del transitorio de la forma de onda de audio. Cuando se producen transitorios musicales, se necesitan bloques cortos, por lo que la resolución de la frecuencia y, por tanto, la ganancia de codificación serán bajas. En otras ocasiones, los bloques pueden hacerse más grandes,

mejorando así la resolución de la frecuencia de la transformada y obteniéndose una mayor ganancia de codificación.

3.7.10 CODIFICACIÓN MPEG DE AUDIO

Las normas MPEG de audio definen tres capas (layers) de codificación, que se distinguen por su tasa de compresión para una calidad de audio percibida dada.

La norma de televisión digital DVB prescribe para el sonido la utilización de las capas 1 y 2 de la especificación MPEG-1 de audio, que prevé cuatro modos principales de transmisión:

- Stereo: los canales izquierdo y derecho se codifican de manera completamente independiente
- Joint_stereo: aprovechamiento de la redundancia entre los canales izquierdo y derecho a fin de reducir el flujo (con dos codificaciones posibles; intensity_stereo o MS_stereo)
- Dual_channel: los dos canales son independientes (sonido bilingüe, por ejemplo)
- Mono: un solo canal de sonido

Capa 1. También llamada "pre-MUSICAM" utiliza el algoritmo PASC, desarrollado por PHILIPS para su casete de audio digital (DCC). Utiliza una velocidad fija entre las 14 posibles (de 32 a 448 Kbits/s); la calidad Hi-Fi necesita 192 Kbits/s por canal de audio (384 Kbits/s en estéreo). Su principal ventaja es la relativa sencillez para implementar el codificador y el decodificador.

La cuantificación de los coeficientes de sub-banda está definida para toda la duración de la trama por un número de 4 bits, permitiendo una codificación de 0 a 15 bits para cada sub-banda, así como el factor de escala sobre 6 bits.

Capa 2. Su algoritmo se conoce bajo el nombre de MUSICAM, es el estándar adoptado para la radio (DAB) y televisión (DVB) digitales europeas. Permite obtener una calidad equivalente con un flujo menor (reducción del 30% al 50%) que el de la capa 1, a costa de un incremento moderado de la complejidad tanto del codificador como del decodificador.

El flujo, constante, puede escogerse entre 32 y 192 Kbits/s por canal, la calidad subjetiva Hi-Fi se obtiene a partir de 128 Kbits/s por canal, es decir, 256 Kbits/s en estéreo.

El modelo psicoacústico utilizado es el mismo que para la capa 1, pero la trama tiene el triple de duración, lo que reduce la proporción de bits de sistema, haciendo que la cuantificación de los coeficientes de sub-banda tenga una resolución decreciente (cuantificación definida sobre 4 bits para las bandas bajas, 3 bits para las medias, 2 bits para las elevadas) en lugar del formato uniforme sobre 4 bits de la capa 1. Por otro lado, 3 muestreos de sub-banda consecutivos pueden ser eventualmente reagrupados en "gránulos" para ser codificados por un sólo coeficiente, de ahí la reducción del flujo.

Capa 3. Es de desarrollo más reciente y utiliza un modelo psicoacústico diferente (llamado modelo 2), una codificación Huffman y un análisis de la señal basado en la DCT en vez de en la codificación en sub-bandas de las capas 2 y 3. Están permitidos los dos tipos de codificación joint_stereo.

Permite un flujo variable y una tasa de compresión aproximadamente dos veces más elevada que la capa 2, a costa de una complejidad claramente mayor del codificador y del decodificador, así como de un tiempo de codificación/decodificación más largo. La calidad Hi-Fi se obtiene a partir de los 64 Kbits/s por canal (128 Kbits/s en estéreo). Está destinada principalmente a aplicaciones de redes de baja velocidad (por ejemplo INTERNET).

Las capas MPEG de audio soportan compatibilidad ascendente entre ellas, es decir, que un decodificador de capa 3 decodificará también las capas 1 y 2, y que un decodificador de la capa 2, normalmente decodificará la capa 1.

Los diagramas de bloques de un codificador y un decodificador MPEG de audio se representan en las Figuras 67 y 68 respectivamente.

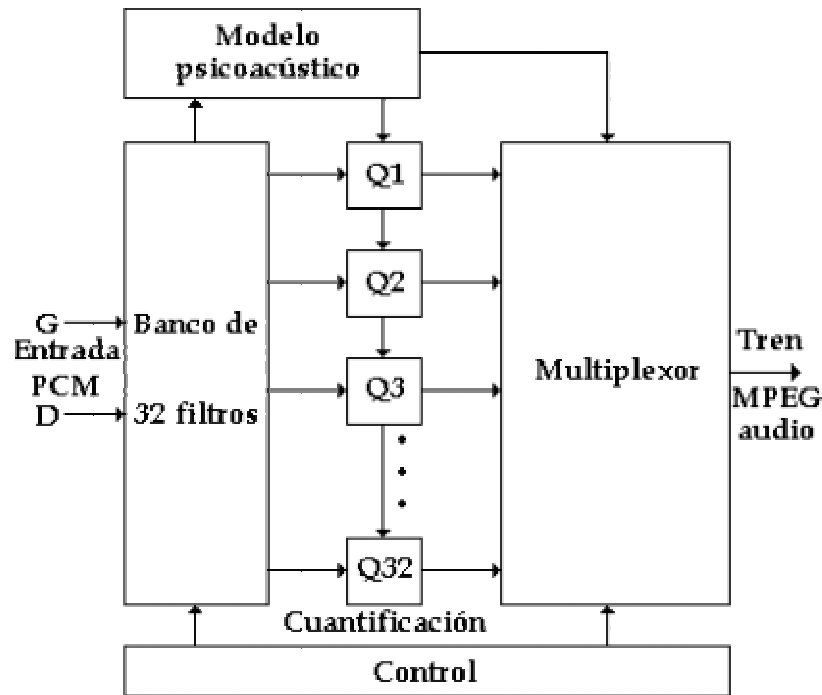


Figura 67. Diagrama de bloques de un codificador MPEG de audio

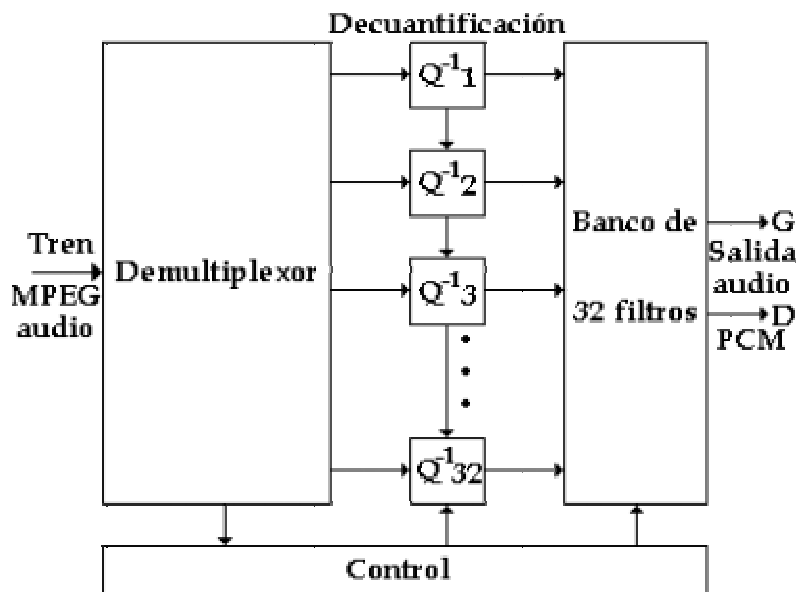


Figura 68. Diagrama de bloques de un decodificador MPEG de audio

3.7.11 FORMATO DE LA TRAMA MPEG DE AUDIO

La trama constituye la unidad de acceso elemental para una frecuencia de audio MPEG.

Una trama (capa 1, 2 o 3), se descompone en 4 partes:

- Cabecera de 32 bits (header)
- Paridad sobre 16 bits (CRC)
- Datos de audio (AUDIO), longitud variable
- Datos auxiliares (AD, ancillary data)

Capa 1. La trama MPEG de audio de la capa 1 se compone de 384 muestreos PCM de audio de entrada. Cuando el número de muestreos PCM es independiente de la frecuencia de muestreo, la duración de la trama es inversamente proporcional a la frecuencia de muestreo. Esta es de:

- 12 ms a 32 KHz
- 8.7 ms a 44.1 KHz
- 8 ms a 48 KHz

En la Figura 69 se puede ver una estructura de una trama de la capa 1.



Figura 69. Representación de la estructura de una trama MPEG de audio, capa 1.

Observaciones:

- La cabecera transporta la sincronización y la información de sistema, detalladas en el cuadro 3.1.
- La utilización de paridad (CRC) es optativa.
- El campo de asignación de los bits/MT-SB (32 enteros codificados sobre 4 bits) define la resolución de codificación (de 0 a 15 bits) de los muestreos de cada una de las 32 sub-bandas.

- El campo factor de escala (32 enteros codificados sobre 6 bits) indica para cada sub-banda el factor multiplicador de los muestreos de esta forma cuantificados.

Capa 2. La trama se compone en este caso de 12 gránulos de $3 \times 32 = 96$ muestreos de audio PCM, es decir:

- 36 ms a 32 KHz
- 26.1 ms a 44.1 KHz
- 24 ms a 48 KHz

La estructura de la parte de audio difiere de la capa 1 debido a una asignación de bits más compleja, motivada por la mayor cantidad de opciones de codificación. En la Figura 70 se puede observar una estructura de una trama de la capa 2.

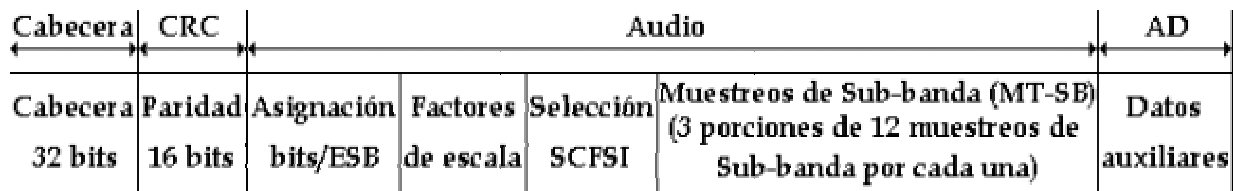


Figura 70. Representación de la estructura de una trama MPEG de audio, capa 2.

Observaciones:

- La cabecera transporta la sincronización y la información del sistema. (Ver cuadro 13).
- La utilización de la paridad (CRC) es optativa.
- El campo asignación de los bits/MT-SB (32 enteros codificados sobre 2 o 4 bits, según la sub-banda) define la resolución de codificación de los muestreos de cada una de las sub-bandas y si están o no agrupadas en 3.
- El campo SCFSI (Scale Factor Selection Information) (32 enteros codificados sobre 2 bits) indica si el factor de escala de sub-banda se aplica a toda la trama o si hay 2 o 3 factores de escala.
- El campo factor de escala indica el factor multiplicador de los muestreos de esta forma cuantificados para la porción de trama definida por SCFSI. [13]

Campo	Comentario	Nº de bits
syncword	tren 1111 1111 1111 (FFF hex)	12
ID	siempre a "1" para MPEG-1 de audio	1
layer	11=1, 10=11, 01=111, 00 reservado (capa)	2
protection_bit	0 si se añade redundancia, 1 si no	1
bitrate_index	15 valores (0000=flujo libre, 1111=prohibido)	4
sampling_frequency	00=44.1 KHz, 01=48, 10=32, 11=reservado	2
padding_bit	1=ajuste (necesario para Fmuestreo=44.1 KHz)	1
private_bit	no especificado, uso libre	1
mode	00=stereo, 01=joint, 10=dual, 11=mono	2
mode_extension	margen de las sub-bandas en intensity_stereo	2
copyright	1=copyright, 0=libre	1
original/copy	1=original, 0=copia	1
emphasis	00=no, 01=50/75 μ s, 10=reservado, 11=J17	2

Cuadro 13. Campos de la cabecera de una trama MPEG-1 de audio

CAPTITULO 4 TEMAS RELACIONADOS CON EL VIDEO DIGITAL

4.1 HDTV (TELEVISIÓN DE ALTA DEFINICIÓN ó TELEVISIÓN DIGITAL)

La televisión en Alta Definición promete ser la innovación más importante desde la introducción del color en 1960. Proporcionará una calidad de imagen cercana al cine de 35mm con sonido multicanal. La HDTV representa una verdadera revolución en la resolución para televisión.

La resolución es el número de líneas y el número de pixels por línea de una imagen de televisión, y en la HDTV se ha incrementado sustancialmente. Los pixels son los elementos individuales de la imagen dispuestos en filas y columnas que conforman la imagen de televisión.

Mientras la imagen de televisión estándar SD (Standard Definition) se transmite con una resolución de 720 x 576 pixels, la imagen de alta definición HD (High Definition) tiene un tamaño de hasta 1920 x 1080 pixels. Por tanto, el número de elementos de la imagen por segundo se multiplica por un factor de 5.

El resultado es una imagen con mucha claridad, nitidez y detalle que para aprovecharlas al máximo, es preciso visionarlas en pantallas o monitores HDTV de alta calidad o a través de proyectores.

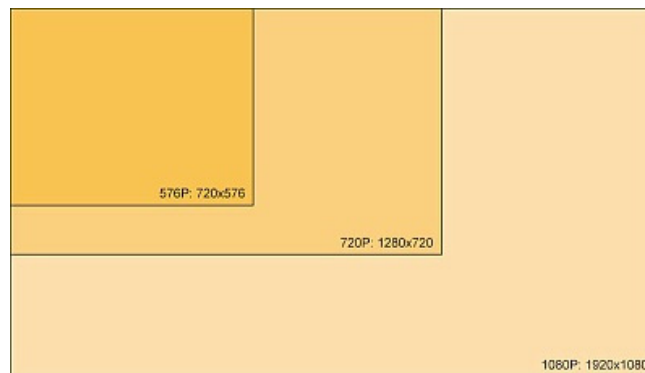


Fig. 71 resolución en pixeles de la HDTV.

Dolby Digital

La HDTV no sería lo mismo si no fuera acompañada de un gran sonido. El Dolby Digital puede proporcionar 5.1 canales de sonido envolvente para complementar perfectamente la gran calidad ofrecida por la HDTV. Una emisión en Dolby Digital 5.1 puede captar nuestra atención de tal forma, que no lo haría una emisión en estéreo, proporcionándonos una experiencia tipo cinematográfica en una película, o poniéndonos en el centro de un acontecimiento deportivo. El Dolby Digital es lo que la audiencia demanda.

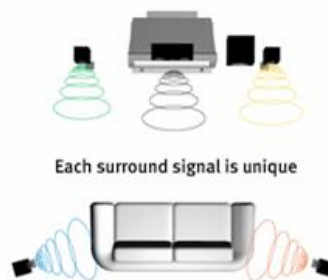


Fig. 72. Dolby digital.

El Dolby Digital es el formato mayoritario para las películas y está presente en más de 40,000 salas de cine de todo el mundo, aparte de ser un estándar en los reproductores de DVD domésticos, existiendo millones de unidades en todo el mundo.

El Dolby Digital es el formato que proporciona sonido envolvente para las películas, deportes y dramáticos, pero su flexibilidad permite emisiones desde sonido monofónico a multicanal y sin problemas de compatibilidad. Esto quiere decir que si una emisión se hace en multicanal, la podremos escuchar con todo su esplendor en nuestro equipo de home cinema, en nuestra televisión estéreo o en el pequeño receptor monofónico de 14" de la cocina o del dormitorio, ya que el decodificador de Dolby Digital permite realizar el mezclado hacia abajo (downmixing) para compatibilizarlo con todo tipo de equipos.

El Dolby Digital es fácil de usar, ya que sólo se necesita una conexión digital entre el decodificador y nuestro sistema multicanal. Los sistemas actuales son muy fáciles de

instalar, tanto en las conexiones como en los menús de configuración de los equipos, llevándonos sólo algunos minutos la colocación de los altavoces tipo satélite.

Los descodificadores poseen una salida analógica estéreo para los equipos que no dispongan de sonido multicanal. Esta salida ofrece la mezcla compatible con sonido estéreo o monofónico.

La información contenida en cada uno de los canales es distinta e independiente (llamado Discrete) Estos seis canales llamados “sistema 5.1” corresponden a 5 canales con todo el espectro de audio (izquierdo, central, derecho, posterior izquierdo y posterior derecho) y a un canal únicamente para sonidos de bajas frecuencias comprendidas entre 3 y 120 Hz que requiere de un altavoz especial llamado subwoofer.

Dolby Digital Plus

Los laboratorios Dolby ya han presentado el Dolby Digital Plus que se adapta a las nuevas tecnologías. Permite más variedad de flujos digitales, sobre todo para usarse con codecs de baja tasa de datos.

Codificación de video

Todos los sistemas de HDTV del mundo, incluido el sistema DVB, están basados en la compresión de video MPEG-2. El consorcio DVB está estudiando la implantación de nuevos codecs de audio/video, tales como el H.264 (MPEG-4 AVC) y el Windows Media 9 de Microsoft, que permiten reducir los flujos binarios manteniendo la misma calidad audiovisual.

Cuando aparezcan los nuevos equipos, éstos serán compatibles con las normas anteriores, tanto en HDTV como en SDTV (MPEG-2 MP@HL).



Formatos de video

Existen dos formatos de video mayoritarios para usar en emisiones HDTV. Estos formatos se abrevian como 720P (donde la **P** indica progresivo) y 1080i (donde la **i** indica entrelazado).

1080i es un formato definido en la ITU-R Recomendación BT.709-5.

Consta de 1080 líneas activas por imagen y 1920 muestras por línea activa. La relación de aspecto es 16:9 y usa pixel cuadrado.

720P es un formato progresivo definido por el estándar televisivo SMPTE 96M-2001. Este formato proporciona 720 líneas activas por imagen y 1280 muestras por línea. La relación de aspecto es 16:9 y usa pixels cuadrados.

En otro artículo encontrarás más detalles sobre esto y el número de imágenes por segundo.

Entrelazado versus Progresivo.

La televisión tradicional ha usado desde siempre el escaneado entrelazado para conservar el ancho de banda. En el escaneado entrelazado, cada cuadro se muestra en la pantalla dos veces. Durante la primera vez, se muestran las líneas impares durante 1/50 de segundo y, en el segundo pase, se muestran las líneas pares durante otro 1/50th de segundo. La imagen completa se muestra 25 veces por segundo.

Por otro lado, en el escaneado progresivo, se muestra toda la imagen completa durante esa secuencia (cada 1/50 de segundo).

Existen muchas discusiones sobre las ventajas e inconvenientes de un sistema u otro. Aquí hay algunos:

- El formato progresivo es fácil de comprimir y permite bajos flujos de datos.
- Las imágenes fijas son mejores con 720P50; el escaneado entrelazado puede introducir artefactos en la imagen durante los movimientos rápidos, sobre todo cuando se visualizan en monitores progresivos (LCD, Plasma,..)

- 720P50 presenta menos artefactos que el 1080i
- La producción en 1080i es más fácil ya que hay más equipos disponibles. El 1080i está más asentado a nivel mundial que el 720P.

Modulación

DVB-S2 es el sucesor natural del estándar DVB-S. Combinado con codificadores avanzados de audio y video, proporcionará la solución óptima para las emisiones HDTV en pocos años. DVB-S2 proporciona una capacidad mayor del 30 % bajo las mismas condiciones actuales de transmisión, por tanto, puede mejorar la eficacia del espectro radioeléctrico.

El nuevo Variable Coding and Modulation (VCM) proporciona diferentes niveles de protección de errores aprovechando el incremento de los diferentes servicios HD que el DBV-S2 puede suministrar.



El HDMI

El HDMI puede que sea el sucesor del Euroconector o SCART. Es una conexión reciente totalmente digital basada, fundamentalmente, en la tecnología DVI y disfruta del soporte de la industria audiovisual.

Las ventajas del HDMI son:

- Su ancho de banda es muy elevado, asegurándose aplicaciones futuras tipo HDTV.
- Un simple cable puede transferir información de audio y video sin comprimir desde un decodificador hasta un monitor o pantalla.
- El hecho que transporte video sin comprimir, garantiza la más alta calidad de imagen.
- Elimina las conversiones analógico-digitales y digital-analógicas.
- Sustituye todos los actuales cables de audio y video entre diferentes componentes audiovisuales, por sólo un cable: esto simplifica las conexiones y elimina errores de cableado y “limpia” la parte posterior de los muebles donde se alojan los equipos audiovisuales.
- Proporciona una protección llamada HDCP (High-bandwidth Digital Content Protection) protegiendo el enlace entre diferentes componentes HDMI. (Así lo quiso la industria audiovisual)
- Soporta el Consumer Electronics Control Protocol (CEC) que permite controlar todos los componentes audiovisuales con un solo mando a distancia.
- Es totalmente compatible con el DVI-HDCP y permite a los equipos DVI conectarse a monitores HDMI a través de un simple cable adaptador pasivo. Los equipos HDMI también pueden conectarse a monitores DVI-HDCP.
- Es un pequeño conector por donde, por los pines internos, pasa la misma información que por un conector DVI, pero por donde también circulan datos de audio digital multicanal. [14]



4.1.1 USO DE LA ALTA DEFINICION

La tecnología de la alta definición es la que mejores cambios ha proporcionado en el mundo del cine y la televisión. La captura de imágenes en alta definición y la visualización, combinada con las nuevas técnicas de codificación y transmisión digital, pueden proporcionar imágenes con una resolución de hasta cuatro veces mejor que la televisión estándar. La calidad de la imagen es mejor ya que tiene más resolución al tener mayor número de pixels. La televisión estándar se basa en una pantalla que contiene 450.000 pixels aproximadamente, visualizados a 25 o 30 cuadros por segundo. La alta definición puede tener hasta 2 millones de pixels vistos a 24, 25 o 30 cuadros por segundo, o 1 millón vistos a 50 o 60 cuadros por segundo.

El mundo del cine crea y muestra imágenes y sonidos de tal forma que impactan profundamente en la audiencia. Visto en buenas condiciones de oscuridad y con toda la atención centrada en una gran pantalla, la audiencia puede sentir nuevas sensaciones.

La televisión, hasta la fecha, es básicamente un medio de información visual de pantalla pequeña. Mientras algunos programadores producen, con imaginación, las imágenes, las tradicionales cámaras de televisión están diseñadas para replicar lo que vemos con los ojos.

Las nuevas cámaras de video de alta definición, ofrecen muchas facilidades para crear efectos cinematográficos más parecidos al cine que a la televisión. Los proyectores de alta definición se han desarrollado para proporcionar unas extraordinarias imágenes de gran calidad sobre grandes pantallas de cine. Así nace la cinematografía digital.

4.1.2 TELEVISION DE ALTA DEFINICION

Emisión en HDTV

La emisión de programas en alta definición está limitada, a fecha de hoy (2004) a Estados Unidos y Australia, aunque otros países están decidiendo usar, en un futuro no muy lejano,

su espectro radioeléctrico para este tipo de emisiones. No hay emisiones regulares de alta definición en Europa, ya que los programadores han optado por usar el espectro radioeléctrico para difundir multitud de programas en definición estándar. Tan sólo existe una cadena, Euro 1080, que emite regularmente en alta definición a través del satélite Astra. También existen algunos canales promocionales de alta definición en Eutelsat (2) en Astra (promo de Pioneer) y en Hispasat.

Programación en HDTV

Aunque los programadores de contenidos son internacionales, algunos productores europeos optan por hacer sus programas en alta definición, obteniendo un acceso más fácil a los mercados internacionales de alta definición. Los programas de alta definición pueden ser convertidos en definición estándar de forma muy fácil, manteniendo un nivel de calidad más elevado que si se hubiera realizado en definición estándar. Por otro lado, las producciones de buena calidad pero realizadas en definición estándar, también se pueden convertir a alta definición para venderse a nivel internacional.

Los programas de alta definición pueden usar dos técnicas diferentes y de mayor uso:

Usando una resolución de 1920 x 1080, proporcionando 2,07 millones de pixels, con una relación de aspecto 16:9 y filmado en modo cine a 24 o 25 cuadros por segundo con escaneado progresivo. Este término se llama 1080/25p ó 1080p (1080/24p) El escaneado progresivo de la imagen a 24 o 25 cuadros proporciona una calidad similar al cine y algunos productores la están empezando a usar para sustituir a la película.

La misma resolución pero a 25 o 30 cuadros por segundo en modo entrelazado, llamado 1080/25i ó 1080i (1080/30i) derivado de la televisión. Las imágenes entrelazadas proporcionan una buena calidad en imágenes en movimiento, sobre todo en deportes.

Existe una tercera opción que es 1280 x 720 proporcionando casi un millón de pixels y con una relación de aspecto 16:9. Se utiliza el escaneado progresivo a 50 o 60 cuadros por segundo y se llama 720p. La figura de 59,94 sólo se usa para la compatibilidad con el sistema de televisión analógica de Estados Unidos.

Como puede apreciarse, la combinación de una pantalla panorámica con una imagen el cuádruple de la resolución estándar, es muy atractiva, sobre todo si lo combinamos con sonido envolvente.

4.1.3 CINEMATOGRAFIA

Producción de películas

Las técnicas de televisión en alta definición, que fueron usadas inicialmente en cine para efectos especiales y luego en postproducción, son ahora de uso común sobre todo por la facilidad y calidad proporcionando efectos más atractivos y llamativos. Estas técnicas usadas en cine no son las mismas que para televisión. Hay dos estándares referidos como 2K y como 4K y se refieren al número horizontal de pixels.

Algunos productores o directores de cine, están empezando a usar cámaras electrónicas (video) de alta definición para sus películas. Los fabricantes de estas cámaras están respondiendo a las nuevas necesidades innovándolas constantemente en calidad y funciones para facilitar la libertad creativa que se tiene en cine, incluyendo nuevas ópticas, filtros y todo tipo de accesorios y ajustes. Pero el paso a la cinematografía digital a gran escala es lento y se necesitará tiempo para establecerse y demostrarse realmente los beneficios.

La industria cinematográfica protege históricamente la alta calidad de los negativos de las películas y, actualmente, está estudiando el almacenamiento electrónico de esas películas a resoluciones extremadamente elevadas, con más del doble del número de pixels usados hoy en la televisión de alta definición.

Distribución de películas.

Las películas se distribuyen, actualmente, de forma física entre las salas de cine. Los sistemas de proyección son mecánicos y, salvo el excelente trabajo de algunos operadores, la calidad de las imágenes se va degradando con el paso de las películas. Las distribuidoras

se aseguran que esta degradación no vaya en aumento y proporcionan nuevas copias cuando se necesiten; pero con un coste de dinero.

Una ventaja de la proyección electrónica digital es que no hay casi degradación de la calidad de imagen cuando se reproduce y, si se desea, puede distribuirse por redes de comunicación digital y por satélite. Hoy en día ya existen proyectores electrónicos capaces de mostrar imágenes de alta definición sobre pantallas grandes. Esta increíble flexibilidad asociada con señales electrónicas están empezando a cambiar las formas de mostrar las películas en algunas salas de cine, e incluso abriendo el abanico de posibilidades para esas salas, ya que, independientemente de mostrar películas, pueden mostrar acontecimientos deportivos, conciertos de música rock o clásica, distribuidas originalmente en video de alta definición.

El 1 de enero de 2004 se hizo la primera emisión de televisión en alta definición digital para Europa, por parte del canal Euro1080, usando en el formato 1080i (1920 x 1080 a 25 fps en entrelazado). Fue el concierto de año nuevo y se emitió a través del satélite Astra a varias salas de cine situadas en varios países de Europa. Las personas que no pudieron estar en Viena, tuvieron la oportunidad de ver el concierto en una sala cinematográfica en alta definición y con sonido digital multicanal.

Estos cambios, como se ha dicho, tendrán un profundo efecto en las distribuidoras de películas y en los usuarios de las salas de cine. El coste de equipar una sala con un proyector digital y una red de comunicaciones es alto, aunque algunos propietarios de multisalas, ya han equipado, al menos, una de ellas para usarla con este fin.

Muchas ciudades y pequeñas comunidades no tienen salas de cine, pero sí que disponen de salones de actos en los ayuntamientos o colegios. Se pueden adquirir pequeños proyectores electrónicos y mostrar esas películas o eventos educativos. [15]

4.1.4 ALTA DEFINICION. LOS ESTANDARES.

HDTV es Televisión en Alta Definición. HD es Alta Definición. Frame Cuadro. Field Campo

La alta definición ha tenido varios formatos durante mucho tiempo, y se han propuesto varios estándares. La industria del cine está empezando a usar los formatos HDTV con el propósito de obtener altas resoluciones para mostrar las imágenes con la mejor calidad en las grandes pantallas de cine. Así, han consolidado el estándar de alta definición más común.

Widescreen (pantalla panorámica)

Primero, y lo más importante, todos los formatos de alta definición adoptan la misma relación de aspecto de pantalla panorámica 16:9.

Square Pixels (pixels cuadrados)

Segundo, en todos los estándares de alta definición, los pixels son cuadrados. Esto incluye a la industria informática, permitiendo integrar de forma más simple los gráficos generados por ordenador en las imágenes de alta definición.

Colorimetría.

Todos los estándares HDTV de las dos familias usan la colorimetría definida en la ITU-R BT.709. Esta **NO** es la misma colorimetría que se usa en los sistemas de televisión estándar PAL o NTSC.

Dos “familias” de estándares.

Los estándares HDTV han reconocido la convergencia entre la electrónica, cinematografía e industria informática, siendo una parte importante para la reproducción en modernas televisiones y producciones de cine.

Existen dos “familias” de formatos de televisión en alta definición (HDTV) que se distinguen por el número de pixels y líneas. Una de las familias tiene 1.080 líneas activas de imagen mientras que la otra, tiene 720 líneas. Cada familia define varias frecuencias de visualización o imágenes por segundo.

Una de las elecciones más importante de la alta definición, ha sido el escaneado entrelazado y progresivo. La HDTV admite ambos, reconociendo las ventajas de cada uno de ellos. La forma más común para referirse a un estándar de alta definición, es usar el número de líneas y la frecuencia visualización. Por ejemplo, 1080/50i y 720/60p se pueden usar para definir el estándar, donde el primer número indica siempre el número de líneas, el segundo número indica la frecuencia de visualización, y la “i” o la “p” indica si el escaneado es entrelazado (**i**) o progresivo (**p**).

Alta definición 1920 x 1080 “Common Image Format” (HD-CIF)

Esta familia está definida internacionalmente por la SMPTE 274M y la subdivisión ITU-R BT.709-5. El estándar BT.709 define un formato de imagen y frecuencia de visualización, y todas sus variantes tienen 1920 pixels horizontales y 1080 líneas activas de imagen.

Con una relación de aspecto 16:9 y siendo el pixel cuadrado, $(1080 \times 16/9 = 1920)$ encaja en el mundo informático.

El formato HD-CIF de 1920 x 1080 contiene 2,07 millones de pixels en una sola imagen de televisión (comparado con los cerca de 400.000 pixels de una imagen PAL o NTSC). Así, el aumento potencial de resolución es de un factor de casi 5 veces.

Las variantes se refieren a las diferentes frecuencias de visualización, y la forma en que las imágenes son capturadas; de forma progresiva o entrelazada.

La SMPTE define once (sí, 11) formatos de escaneado de HDTV 1920 x 1080, ocho de ellos progresivos y tres entrelazados.

La ITU, ahora en su quinta revisión, define diez sistemas de escaneado – ocho progresivos y dos entrelazados. Estos incluyen 25fps para Europa, 30fps para Estados Unidos y Japón y 24fps para la industria cinematográfica.

El formato común de imagen (CIF) facilita el intercambio de programas entre diferentes entornos y hace posible que cualquier equipamiento pueda trabajar en cualquier entorno. Como tal, este es un gran paso adelante comparado con los sistemas actuales totalmente incompatibles. Hay que recalcar que el actual documento ITU BT.709-5, recomienda el uso del formato HD-CIF para la producción de nuevos programas y facilitar así los intercambios internacionales.

Rec. ITU-R BT.709-5

Ahora se encuentra en su quinta revisión, (la alta definición ya lleva muchos años en el mercado y el documento original data del año 1.990). El punto principal está en la segunda parte del documento, ya que la primera parte, que definía los sistemas originales de alta definición, está descatalogado.

En la segunda parte, el formato común de imagen (CIF) está definido “*para tener un parámetro de imagen común, independiente de la frecuencia de la imagen*”. Los parámetros claves son el sistema de escaneado y la colorimetría. Las distintas frecuencias permitidas son las siguientes:

Sistema	Captura	Escaneado
24p, 25p, 30p, 50p, 60p	1920 x 1080, progresivo	Progresivo
24psF, 25psF, 30psF	1920 x 1080 captura progresiva	Cuadro segmentado
50i, 60i	1920 x 1080 entrelazado	Entrelazado

Tabla. 14. Donde i=entrelazado, p=progresivo y psF=progresivo con cuadro segmentado.

El cuadro segmentado (Segmented Frame) es una forma de transportar una imagen progresiva en dos segmentos, así esa señal se “ve” igual que los dos campos de una imagen entrelazada.

En post producción se necesitará trabajar en ambos formatos de señal, tanto en entrelazado como en progresivo, durante un cierto tiempo. Uno de los problemas para monitorizar los nuevos formatos de señal, como el 24p, es el parpadeo (flicker) inducido en los monitores de televisión TRC. El segundo es el procesamiento de las imágenes progresivas (la mayoría de los monitores de televisión TRC muestran las imágenes de forma entrelazada). El formato de cuadro segmentado permite usar los mismos sistemas electrónicos para imágenes progresivas y entrelazadas, y visualizarlas correctamente sobre monitores de TRC. No hay cambios en la característica de la imagen progresiva, y sólo se usa para frecuencias de hasta 30fps. Tampoco hay problemas para monitorizar la señal con los nuevos visualizadores planos, tanto de LCD o de plasma. El interface digital de una señal entrelazada o psF es el mismo, aunque el contenido de esa señal es diferente.

Alta definición 1280 x 720 Progressive Image Sample Structure

Definido internacionalmente por la SMPTE 296M, aunque no por la ITU, es una familia que incluye ocho sistemas de escaneado - todos en formato progresivo - teniendo todos una resolución de 1280 pixels horizontales y 720 líneas activas. Proporciona 921.600 pixels en una imagen, pero al estar definido como un formato de imagen sólo progresivo, acarrea algunas implicaciones.

Las frecuencias de visualización son 23,98p, 24p, 25p, 29,97p, 30p, 48p, 50p, 59,95p y 60p.

Sistema	Captura	Escaneado
24p, 25p, 30p, 50p, 60p	1280 x 720 progresivo	Progresivo
23,98p, 29,97p, 59,94p	1280 x 720 progresivo	Progresivo compatible NTSC

Tabla 15. La colorimetría cumple la norma ITU-R BT.709

Frames/Fields. Cuadros/campos

La resultante de que haya muchas variantes por cada familia de Alta Definición, es la multitud de frecuencias de cuadro o campos.

Históricamente, en Europa y otras partes del mundo el sistema de televisión estándar tiene una velocidad de 25 imágenes por segundo. En Estados Unidos y Japón se usan 30 imágenes por segundo, mientras que la industria cinematográfica usa 24 imágenes por segundo.

Entrelazado o Progresivo.

Las emisiones de televisión estándar han sido siempre entrelazadas. Esto significa que cada imagen se transmite en dos mitades o campos: el primer campo contiene las líneas impares; el segundo campo contiene las líneas pares. Así, las imágenes, consistentes en dos campos entrelazados, se transmiten al doble de la frecuencia de cuadro – 50Hz o 60Hz. Es ahora el momento de empezar a llamar a esas velocidades de campos por su número de cuadros – **25i y 30i**.

Debido a las complicaciones del sistema NTSC, existe un pequeño número de frecuencias de cuadros “compatible NTSC”, cuya frecuencia nominal se divide entre 1.001. Así, la versión NTSC de 60Hz es de 59,94 y la de 30Hz de 29,97.

Finalmente, la industria cinematográfica ha usado siempre la velocidad de 24 imágenes por segundos (las cámaras graban siempre cuadros). Esto produce una especie de parpadeo característico debido a la baja frecuencia de muestreo temporal, pero resulta artísticamente atractivo por muchos productores y directores, sobre todo para la producción de dramáticos. La Alta Definición se está usando también en producciones cinematográficas, por eso se han definido velocidades de 24 imágenes por segundo en Alta Definición para cine.

Transmisión y producción.

Las grandes cadenas de televisión americanas emiten actualmente en ambos formatos de alta definición, 1080i y 720p usando la misma frecuencia de cuadro/campo que la señal de televisión estándar.

En Europa, la primera cadena de televisión de alta definición – Euro1080 -, que empezó a emitir en enero de 2004, está usando el formato 1080i a 25i (50Hz)

En alta definición, es normal producir imágenes usando un estándar y emitirlo en otro estándar. Por ejemplo, se puede grabar a 25 imágenes o 50 campos por segundo y emitirlos a 60 campos por segundo.

Debido a la gran cantidad de datos necesarios para las variantes de escaneo progresivo a 1920 x 1080, las cámaras actuales no disponen de frecuencias mayores de 30Hz, salvo cambios de última hora. Las imágenes entrelazadas pueden ser capturadas fácilmente a 60 campos por segundo.

La frecuencia de datos a 720p es más manejable, y hay cámaras que graban imágenes a 60 cuadros por segundo, proporcionando una frecuencia de reproducción variable. Esto es particularmente interesante en documentales de naturaleza y acontecimientos deportivos.

Existen conversores de alta definición en el mercado, que pueden convertir entre frecuencias de cuadro y entre familias de 1080 y 720, aunque todavía no compensan la variación de la conversión entre 25i y 30i.

Para producción, es posible masterizar material a 24p. El material se maneja de la misma forma que las imágenes de cine – en Estados Unidos usando el 3-2 pulldown: en Europa aumentando la velocidad hasta 25 imágenes por segundo. La baja frecuencia de muestreo temporal de 24 imágenes por segundo, no es apropiado para todo tipo de material – en

particular para deportes de mucha acción- donde el parpadeo puede molestar la visión por parte del espectador.

Los equipamientos de post producción en Alta Definición son, en general, muy flexibles y pueden trabajar con imágenes adquiridas usando todos los estándares de ambas familias. La elección del formato de adquisición (frecuencia de cuadro y tipo de escaneado) dependerá del estilo creativo y del contenido.

La única limitación es el requerimiento basado en 50Hz o 60Hz para la emisión, y esto se puede considerar en la pre producción. Se pueden elegir las variantes de 24p, 25p y 50i para Europa y afines, o convertir 24/30p o 60i para USA y Japón aunque, como se dijo antes, la conversión de frecuencia es posible pero sin compensación (de momento). [16]

4.1.5 AUDIO PARA ALTA DEFINICIÓN

4.1.5.1 FORMATOS DE AUDIO PARA ALTA DEFINICION.

En comparación con la grabación y edición de imágenes en alta definición, la creación de la banda sonora para alta definición requiere una pequeña programación de la forma de trabajo, nuevas técnicas y nuevo equipamiento. La grabación, edición y mezcla de audio puede utilizar estas técnicas aunque los procesos son muy parecidos a la televisión estándar y cinematografía. La banda sonora para programas en alta definición puede hacerse de varias formas, desde monofónica hasta sonido envolvente en 5.1 canales. Sin embargo, se asume que en la mayoría de los programas, se usa, al menos, la mezcla estéreo y surround en su forma básica.

Existen dos tipos de sonido envolvente. La matriz envolvente original Dolby Surround Prologic, proporciona cuatro canales de sonido, con altavoces frontales izquierdo, central y derecho y un canal envolvente monofónico. El altavoz central posiciona los diálogos en pantalla y el canal envolvente que introduce al espectador dentro de la escena. Esta matriz envolvente codifica los cuatro canales en dos, llamados Lt y Rt, grabándose en cualquier

medio estéreo y transmitiéndose sobre cualquier red estándar. Actualmente la usan algunos operadores de TV en sus programas o películas en los sistemas de televisión analógicas.

La televisión digital y el DVD ofrecen un segundo sistema de sonido envolvente llamado 5.1, usado en la mayoría de películas cinematográficas. El sonido 5.1 es transportado como un flujo de datos digitales codificados de gran calidad, eficacia y realismo, proporcionando al espectador una verdadera experiencia audiovisual. Las tecnologías de codificación incluye Dolby Digital , DTS, MPEG y AAC – el Dolby Digital es el estándar universal para DVD, junto al MPEG en Europa, adicionalmente se incluye el DTS y, para aplicaciones de bajos flujos de datos, el AAC. Estas tecnologías tienen sus variantes para emisiones de televisión – ver tabla 16.

	Flujo para 5.1 canales	Requiere versión estéreo	Calidad de sonido broadcast	Incluido en equipos domésticos	Usos
Dolby Digital AC-3	384 kbs	No	Si	Todos	DVD, DVB y ATSC
AAC	320 kbs	Si. 128 kbs adicional	Si	Sólo Japón	ISBD
DTS	1,5 Mbps Propuesto 384 kbs	Si. 192 kbs adicional	Si	En muchos	DVD. Propuesto DVB

Tabla 16. Sistemas de codificación y de transmisión.

El 5.1 usa altavoces frontales izquierdo, central y derecho, dos altavoces envolventes posteriores izquierdo y derecho, y un canal de efectos de baja frecuencia LFE “0.1” todos posicionados discretamente para que el sonido se pueda localizar de una forma más exacta y auténtica. Los dos canales posteriores son de ancho de banda completo y permite al productor crear un ambiente envolvente más realista. El canal LFE proporciona sonidos

extras muy graves, añadiendo un fuerte impacto real para efectos especiales, como explosiones.

Las transmisiones digitales están diseñadas para soportar sonido envolvente 5.1 (otra cosa es que se use), pero también pueden usarse la matriz envolvente, estéreo o mono, aunque lo ideal sería usar la emisión en 5.1 para preservar posibles copias a DVD posteriores.

4.1.5.2 Sonido envolvente. Producción y postproducción.

Los principios de grabación y mezcla de pistas en 5.1 para dramáticos están ya establecidos aunque, como dije antes, otra cosa es que se usen. Sin embargo, la creación de contenidos en 5.1 en directo es algo que está empezando ahora a desarrollarse.

La postproducción de documentales y dramáticos en 5.1 no necesita más complicación que para hacerlos en estéreo. En la producción, hay que grabar los máximos elementos por separado, así, flexibilizaremos la posición de las voces y efectos en la mezcla final, aunque los diálogos los posicionaremos en el canal central, desplazándolos ligeramente a un lado u otro, dependiendo de la posición en pantalla. Hay que tener en cuenta la mezcla en estéreo, para que no se nos desvíe más de lo normal y rompa la estructura audiovisual. Se pueden incluir diferentes atmósferas estereofónicas, tanto en los altavoces frontales, como posteriores, para introducir al espectador en la acción.

Para grabar eventos en directo, como conciertos y deportes, necesitaremos varios micrófonos adicionales para conseguir que el espectador se encuentre inmerso en el espectáculo. Se colocarán dos micrófonos frontales para crear la perspectiva frontal, con los comentarios aplicados al canal central. Por encima de la audiencia, añadiremos varios micrófonos a la izquierda y derecha de la posición de la cámara principal, que podrán ser asignados a los canales laterales o posteriores, dependiendo de la posición de la (las) cámara.

Para grabar un concierto de música clásica, se usarán varios micros frontales para capturar una imagen sonora frontal de la orquesta, incluyendo el canal central y se podrán usar micros adicionales para capturar la reverberación del recinto (incluidos los asistentes) que serán reproducidos por los altavoces posteriores para añadir realismo a la banda sonora.

4.1.5.3 Metadatos

Cuando se crea una mezcla en 5.1 es importante incluir una información, llamada metadatos. El propósito principal de estos metadatos es optimizar la reproducción de la banda sonora a cada receptor o decodificador. Es esencial que estos metadatos se ajusten cuidadosamente en cada mezcla para que los usuarios no tengan que preocuparse en el momento de realizar la escucha. En algunos países, no se hace la transmisión completa del audio en la emisión del contenido HD, sino que se deriva desde el flujo 5.1 hacia una mezcla estéreo, siendo el decodificador, gracias a los metadatos, el que se encarga de hacer el ajuste necesario.

Emisión 5.1 + metadatos →	Decodificador aplica metadatos →	5.1 surround →	Home cinema completo
		5.1 dinámica reducida →	Home cinema modo nocturno
		Mezcla estéreo →	TV estéreo, VCR, Auriculares
		Mezcla mono →	TV monofónica

Tabla 17. Flujo de audio con sus metadatos.

El control de los metadatos puede permitir desde la mezcla inicial en 5,1 hasta la mezcla monofónica, pasando por una reducción de la gama dinámica para escuchar a bajo volumen por la noche.

4.1.5.4 Gestionando la mezcla.

Los magnetoscopios de alta definición, pueden grabar cuatro o más canales de audio (hasta 8). Las mezclas en 5.1 para emisiones se suelen distribuir utilizando codificadores Dolby E, que permite codificar todos los canales de audio y los metadatos, usando sólo dos pistas. Esto simplifica la distribución usando la estructura estándar estéreo.

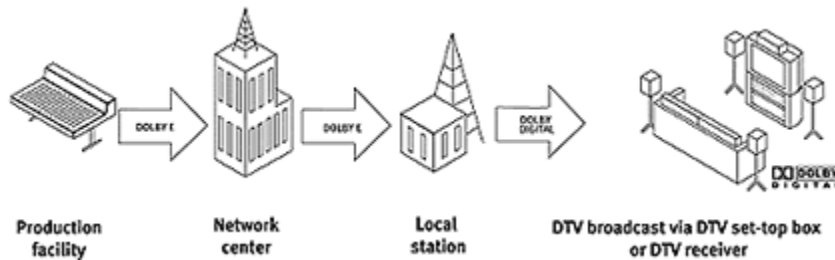


Fig. 73. la grabacion, edicion y mezcla de audio HD.

Técnicamente, las dos primeras pistas del magnetoscopio se usan, convencionalmente, para la mezcla estéreo o para la matriz surround. Las otras dos pistas se usan para poner la mezcla 5.1 codificada en Dolby E con los metadatos incluidos. Como el rango dinámico de la mezcla 5.1 es muy elevado, se comprime la información acorde a un flujo de datos estandarizado y que guarda una relación entre el número de canales y la tasa de flujo, generalmente entre 384 y 448kbs. Las pistas adicionales se usan para referencias de músicas, efectos, voces dobladas en otros idiomas, etc.

4.1.5.5 Reutilizacion de la mezcla.

La mezcla 5.1 con sus metadatos creados para la emisión, pueden usarse directamente para le creación de la banda sonora de un DVD. También puede usarse para una presentación en un cine electrónico aunque, hay que recordar, que la mezcla para equipos domésticos difiere de la mezcla para grandes salas cinematográficas y pueden verse alteradas por varias causas: el volumen de la sala, el nivel acústico, la calidad y número de altavoces, la potencia suministrada, etc. Por estas razones, la mezcla de las bandas sonoras para

películas, suelen hacerse en estudios de grabación preparados con condiciones similares a salas cinematográficas.

El uso de un formato codificado, como el Dolby E, sobre una cinta HD hace que dispongamos de la mezcla en 5.1 con sus metadatos para emisión, o varias mezclas estéreo o 5.1 en versiones internacionales para DVD junto a sus idiomas y referencias de audio.

4.1.5.6 Emisión.

En el momento de la emisión, el audio multicanal, acompañado de sus metadatos, son descodificados desde su formato de origen y vueltos a codificar usando el esquema de emisión. Este flujo ya codificado se multiplexa usando el estándar correspondiente de emisión.

Actualmente existen dos formatos de transmisión de audio 5,1 en alta definición. El sistema ATSC y el DVB usan la especificación AC-3, conocida como Dolby Digital, mientras que el sistema japonés ISDB usa la codificación AAC basada en mpeg audio. Se han realizado pruebas satisfactorias con el sistema de codificación DTS y se propondrá como opcional para el DVB. Los tres sistemas manejan audio en 5.1, pero varían en ancho de banda, estandarización, degradación de metadatos y base instalada de descodificadores. Sin embargo, sólo una mezcla en 5.1 es necesaria en la producción, independientemente del sistema de transmisión utilizada. (Ver tabla 16)

4.1.6 Normas para HDTV

Recomendación para los formatos HDTV de 1125/60 y 1250/50 líneas/campos, definiendo valores y una estructura de muestreo de 4:2:2 o 4:4:4 que corresponda a 5,5 veces a la de la norma ITU-R 131601. Las cadencias de luminancia (Y) son definidas como 74,25 MHz y 37,125 MHz para las diferencias de color Cb y Cr, a 8 bits o 10 bits. Esta norma define una forma expandida de la norma 601 y usa píxeles no cuadrados. El documento original de la norma apenas se refería a 1035 y 1152 líneas activas, correspondiendo al número total de

líneas de barrido de 1125 y 1250. En el año 2000, la revisión de la norma ITU-R BT.709-4 incorporó los patrones de 1080 líneas activas, recomendando su uso en todas las nuevas producciones. Esta revisión definió también el patrón de 1080 líneas con píxeles cuadrados como formatos CIF (Common Image Formats) para intercambio internacional.

4.1.7 Formatos e intervalos

El área de una imagen de televisión que transporta la información de imagen está definida como un área activa. Fuera de esta área existen líneas e intervalos de extinción de campos (blanking) que corresponden, aunque no exactamente a las áreas definidas para los sistemas analógicos originales de 525 y 625 líneas. En las versiones digitales, las áreas activas y de extinción de intervalos son las definidas en las normas ITU-R BT601, SMPTE RP125 y EBU-E. Para normas HDTV de 1125 líneas (1080 líneas activas) que pueden tener cadencias de 60, 30, 25 o 24 Hz, las líneas activas son siempre del mismo tipo (1920 muestreos de pixel a 74,25 MHz y 25,86 microsegundos) conforme se define en los documentos SMPTE 274M e ITU-R.BT 709-4. Sólo las extinciones de línea difieren de forma que la parte activa pueda mapearse píxel por píxel entre estos formatos. Las normas de televisión digital tienen que hacer referencia apenas a los contenidos en términos de área activa de imagen tal como 1920x1080, 1280x720, 720x576. Para los formatos de 625 y 525 líneas, la referencia es de 720 muestras de luminancia a 13,5 MHz = 53,3 microsegundos. En vídeo digital, no existen medias líneas como existen en analógico. El cuadro anexo muestra el intervalo de extinción para las normas SD y HD más comunes.

Formato	625/50	525/60	1125/60i	1125/50i	1125/24p
Líneas activas	576	487	1080	1080	1080
Intervalo de extinción	24	19	22	22	45/cuadro
Línea Campo 1					
Línea Campo 2	25	19	23	23	
Extinción de líneas	12m	10,5ms	3,8ms	9,7ms	11,5ms

Tabla 18. Formato Común de Imagen (Common Image Format, CIF)

La ITU definió los formatos de imagen CIF desde una señal de definición estándar para ordenadores de 352 x 240 píxeles hasta los formatos HDTV. Para producción en alta definición, el formato preferido es el que está definido como HD-CIF en la norma ITU- R BT.709-4, descrito como 1920x1080, con relación de aspecto 16:9 y cadencias de imagen de barrido progresivo de 24, 25 y 30Hz y formatos de campos entrelazados de 50 y 60 Hz. Esta definición ayuda a asegurar que el formato 1920x1080 sea la base actual para intercambio internacional de programas.

Técnica.

Partiendo de que el vídeo de SD está formado a partir de la descomposición de la imagen en 483 líneas con 720 elementos de imagen (pixels) por línea en un sistema de barrido de pantalla de 625 líneas resulta lógico que para aumentar la definición se aumente el número de líneas que barren la pantalla y para mantener equilibradas las resoluciones horizontales y verticales, aumentar proporcionalmente el número de pixels por línea. Para que pueda mantenerse una relación entre la definición estándar y la alta, el número de líneas se situó en el doble (1250) y teniendo en cuenta que la relación de aspecto se definió en 16:9 en

lugar de los 4:3 tradicionales, el número de píxeles por línea pasó a 2376. Esto es lo que define la norma ITU-R BT, 709-4 para la señal de vídeo digital de alta definición.

Esta norma, actualizada en 2000, define que el número de líneas activas sea de 1080 y que el número de píxeles activos se establezca en 1920. En lo que respecta a la imagen, con excepción de la normalización de la colorimetría, esta norma es idéntica a la norma SMPTE 295M y, con las adaptaciones relativas al número de cuadros y la inclusión de algunos formatos con barrido progresivo, a las normas SMPTE 260M y 274M. De hecho, todos los formatos descritos en esas normas comparten un formato común de imagen CIF (Common Image Format) constituido por 1080 líneas activas con 1920 píxeles activos por línea, manteniendo la imagen una relación de aspecto de 16:9.

Digitalización

Ya que las imágenes aún son captadas mayoritariamente por dispositivos analógicos y constituidas por señales de colores primarios (R, G y B), componentes de luminancia (Y) y diferencias de color (Cr y Cb), toda la información debe pasarse a digital a través del muestreo, cuantificación y codificación. Para asegurarse que los niveles recuperados correspondan a los digitalizados, la digitalización debe hacerse con la introducción de márgenes de protección. Así, el nivel de blanco (700mV) debe corresponder a una digitalización igual a 940 de los 1024 niveles de resolución correspondientes a los 10 bits de la cuantificación, el pico de blanco a 1019, el negro a 64, el pico de negro a 48, las crominancias (+350 y -350mV) a 960 y 64 con el nivel de extinción en 512 y los picos (+396,1 y 396,9 mV) iguales a 1019 y 4.

Siendo el número de píxeles 2376, las líneas 1250, el número de cuadros por segundo 25 y las palabras de 10 bits, la frecuencia de muestreo tiene que ser igual a $2376 \times 1250 \times 25 \times 10 = 74,25$ MHz para la componente de luminancia (Y), y la mitad de ese valor, 37,135 MHz, para los componentes de color Cr y Cb. De esta forma, se obtienen flujos de tres palabras de 10 bits en paralelo, no existiendo informaciones en los intervalos correspondientes a las extinciones horizontales ($2376-1920=456$ píxeles) y verticales ($1250-1080=170$ líneas). Estos flujos de datos presentan cadencias de 74,25 Mega palabras

por segundo (MWps) para la luminancia (Y), y de 37,125 MWps para los colores Cr y Cb. La primera operación que se hace con estos flujos de datos es la multiplexación en el tiempo (TDM-Time Division Multiplexing) de los flujos correspondientes a las dos componentes de crominancia.

La segunda operación es la de inserción al final de cada línea de cuatro palabras indicativas del final de línea activa (EAV - End of Active Line), siendo la primera siempre 001111111111, la segunda y la tercera siempre 000000000000 y la cuarta corresponde a la cuantificación de informaciones relativas a las situaciones vertical (V), horizontal (H) y al cuadro (F) de la línea que termina (información XYZ), seguidas de dos palabras donde se indica el número de la línea (como la cuantificación del número de líneas, 1250, excede una palabra -sólo puede cuantificar 1024- hay que recurrir a dos palabras) y dos palabras que transportan información relativa a los códigos de detección de errores (CRC, Cyclic Redundance Codes). Pasados 444 píxeles, en los cuales puede estar insertada información auxiliar o complementaria (código temporal, audio, teletexto....) vuelven a insertarse cuatro palabras, relativas ahora al arranque de línea activa siguiente (SAV, Start of Active Video) idénticas a las de EAV.

La tercera operación es otra multiplexación en el tiempo, ahora entre la luminancia (Y) y los colores Cr/Cb, cuyo resultado es un flujo de datos relativos a YCrCb, con una cadencia de 148,5 MWps y una señal cronométrica de 148,5 MHz que será multiplicada por diez. La cuarta operación es la de serialización (transposición de las palabras en paralelo para bits en serie) pasando el flujo de datos a 1.485 Mbps, sobre el control de reloj de 1.485 MHz.

Las restantes operaciones se destinan a mejorar la resistencia a la corrupción de la señal. Se trata de operaciones de scrambling (donde son tratadas aleatoriamente a amplias secuencias de 0's y 1's) para lograr que el haz no retorne a cero (NRZ, el haz dejará de variar entre 0 y 1, sea 1 el nivel que sea para variar entre los dos niveles, uno positivo y otro negativo) e interrupciones de NRZ, NRZI donde se introducen transiciones de nivel en las secuencias largas de 1's: todo temporizado a 1.485 MHz.

Características de la señal

La salida de este procesamiento es un haz de datos con una cadencia de 1485 Mbps y una amplitud pico a pico de 800 milivoltios en 75 Ohms, que debe tener pérdidas por devolución inferiores a 15 dB, tiempos de subida y de caída inferior a 270 picosegundos (ps) y jitter inferior a 673 ps entre 10 y 100 Hz y 134,6 ps por encima de los 100 Hz. La relación señal /ruido cuantificada es igual a 6 X número de bits más 6, o sea, para 10 bits, 66 dB. Para manutención de estas especificaciones se recomienda que los cables coaxiales no excedan los 85 metros, además debe utilizarse fibra óptica.

La señal recibida por el equipamiento terminal (receptor) llega siempre con algún grado de corrupción debido al ruido térmico. Si la relación señal /ruido fuera mucho más baja, podría perjudicarse la reconstrucción de la señal original. Pero como es el circuito de entrada del receptor el que contribuye con la mayor parte del ruido, este circuito debe tener la mejor calidad posible. En el receptor se lleva a cabo la corrección de las degradaciones de conexión a la fuente, la recuperación del reloj, el paso de NRZI a NRZ, la reordenación, la conversión serie /paralelo, las dos desmultiplexaciones (la primera para Y y Cr/Cb y la segunda para Cr y Cb), y las conversiones digital /analógicas (DAC) que recuperan los componentes analógicos Y, Cr y Cb, pasando a partir de aquí el procesamiento a ser de dos señales analógicas.

4.1.8 El HDV (Vídeo en alta definición a nivel doméstico)

NAB 2004 se recordará como el año en que el formato HDV (vídeo digital de alta definición) ganó legitimidad entre los productores, personal independiente y editores, así como una gran cantidad de profesionales, que demostraron un interés considerable, sobre todo en el periodismo electrónico (newsgathering).

Ediciones técnicas a un lado, cuando hablamos de adquisición de video a 720p y 1080i en HDTV y reproducción a precios de DV (por debajo de los 4.000€ para una cámara de video), la comunidad creativa se interesa de forma muy rápida.

Debido a que el nuevo formato emplea la misma cassette, velocidad de la cinta y anchura de pista que el formato DV, es relativamente fácil que los fabricantes desarrollen productos que sean compatibles con las soluciones DV existentes.

Pero como suele ocurrir, hay detractores que dicen que el formato es limitado y confunde a la gente diciendo que el HDV no será similar en pantalla a las imágenes de HD capturadas con un equipo 100Mb/s. Dicen que el HDV está utilizando la vieja tecnología de la compresión (MPEG-2) en vez de los nuevos codecs de DVCPRO-HD o de HDCAM que hacen un trabajo mucho mejor al preservar la información original del color y de mantener los artefactos tales como las "bandas" al mínimo.

JVC, estaba sólo el año pasado cuando anunció una cámara HDV de un sólo CCD, la JY-HD10U, a 720p. Esta cámara captura imágenes de HD a 25Mb/s en cassettes DV estándar o miniDV y utiliza compresión MPEG-2. En la convención de este año del NAB 2004, JVC anunció que ha vendido más de 2.000 unidades de su cámara HDV en los ESTADOS UNIDOS solamente en los últimos ocho meses.

Bajo el eslogan "HD para todos", JVC mostró un prototipo de tres CCD, el camcorder conmutable HD/SD ENG/EFP que satisface mejor (en términos de capacidad de grabación y tamaño) a los locutores y a los profesionales de la producción. El nuevo camcorder HD de JVC utiliza tres sensores de 2/3 de pulgada de tecnología CMOS, con una resolución nativa de 1920x1080 pixels cada uno, e incluye un codificador MPEG-2 capaz de grabar en SD y HD, incluyendo HD a 24 frames en formato progresivo (24p). La cámara utilizará el formato de grabación HDV, que proporciona 276 minutos de grabación en HD sobre cintas del tamaño normal de DV y 60 minutos de grabación HD en cintas miniDV. La compañía también demostró la transmisión sin hilos en tiempo real (WiFi), de vídeo comprimido HDV con buenos resultados.

El formato HDV fue definido en 2003 por un consorcio de compañías de electrónica de consumo que incluye Canon, JVC, Sharp y Sony. El grupo puso las especificaciones de HDV abiertas y a disposición cualquier grupo con la esperanza de que muchos crearan

productos complementarios de apoyo. Las cámaras HDV utilizan el interfaz ITU-R1394 para transferir las imágenes de la cinta a un sistema de edición por ordenador.

Se dice a menudo que ningún formato nuevo puede sobrevivir sin la ayuda de los fabricantes de la industria pero, debido a su especificación abierta, el HDV goza ahora de una extensa compatibilidad de sistemas y de tecnologías que cubre toda la cadena, desde la producción, hasta la postproducción.

Comparable al formato DV de muchas maneras, el HDV ofrece menos información de color (4 bits contra el proceso de 8 bits del DV). Según el consorcio HDV, la tecnología de procesado proporciona una calidad perceptiblemente mejor de la imagen que los sistemas DV. Los miembros de HDV afirman que cambiando el método de la corrección de error de la anchura de una pista (según lo especificado en el formato DV) a la corrección de error entre múltiples pistas, el formato HDV ofrece una capacidad mejor de corrección de error y una mayor resistencia a la pérdida de datos causados por drop-outs.

Editando con un grupo de imágenes (GOP) largo del formato de HDV, las imágenes comprimidas en MPEG-2 no están en el límite. El MPEG-2 con compresión intraframe del formato HDV permite un vídeo de buena calidad con flujos binarios más bajos, permitiendo más contenido por cinta, pero la compresión intraframe incrementa la dificultad de editar el contenido porque todos los frames individuales de un vídeo no están disponibles para el editor en la línea de tiempo.

Los fabricantes han trabajado duro para conseguir una plataforma flexible para trabajar. Pinnaclesys, que ofrece varios sistemas de edición que apoyan HDV, entre ellos el nuevo Liquid Edition 6, ha dicho, que editar los clips estándar del MPEG intraframe (o DV) es fácil y directo. Si se hace un corte básico entre dos clips durante una edición, el archivo que resulta se recompone juntando los dos clips de nuevo. Si un editor elige una transición entre los frames, sólo los frames de la transición serán recodificados y la nueva pieza se sumará al flujo de salida. Poco o nada de pérdida de calidad existirá en el punto de la transición.

Cuando editamos los frames codificados en el formato IBP de Pinnaclesys, no es fácil realizar un corte o una transición exacta entre frames. La realización de un corte toma una cierta metodología de trabajo debido al hecho de que los frames situados entre los seleccionados puede que no estén disponibles para el editor. Pinnaclesys ha perfeccionado su tecnología de edición IBP así que el usuario no tiene que preocuparse. El Liquid Edition permitirá editar el contenido de HDV (codificado IBP) tan fácilmente como editar material de DV.



Etiqueta estándar HD Ready

La “European Information & Communications Technology Industry Association” ha anunciado una etiqueta estándar para todos los dispositivos de visualización que sean compatibles con la Alta Definición (HD Ready) tales como TV, pantallas de plasma, LCD, proyectores, retroproyectores y monitores.

La etiqueta “HD ready” está soportada por todos los miembros de la EBU (European Broadcasting Union) y proporciona al consumidor la garantía que el dispositivo visualizador que compre a determinado fabricante, sea compatible con las emisiones actuales y las futuras emisiones en HDTV.

4.1.9 HD. Recepción y visualización.

La siguiente figura ilustra la cadena de recepción y visualización en HD. Consiste en un sintonizador HD (satélite, cable o terrestre) llamado set top box (STB), y un visualizador HD, o de una televisión integrada HDTV, que combina alguna o todas las partes del dispositivo. La etiqueta “HD ready” sólo se refiere al visualizador HD.

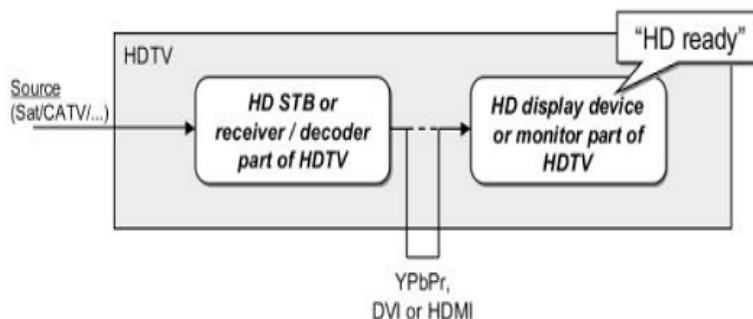


Fig 73. HD ready. Recepción y visualización.

Requerimientos para la etiqueta “HD ready”

Un visualizador ha de cumplir los siguientes requerimientos para poder mostrar la etiqueta “HD ready”:

1. Visualizador.

La resolución nativa mínima del dispositivo visualizador (LCD, Plasma, etc) será de 720 líneas físicas en formato 16:9.

2. Interfaces de Vídeo.

El visualizador ha de aceptar la entrada en HD a través de:

- Conexión analógica por componentes YPbPr (*), y
- Conexión digital por DVI o HDMI.

Las entradas HD aceptarán los siguientes formatos de video HD:

- 1280x720 @ 50 y 60Hz con exploración progresiva (“720p”), y
- 1920x1080 @ 50 y 60Hz con exploración entrelazada (“1080i”)

La entrada DVI o HDMI ha de soportar la protección de contenidos (HDCP)

Nota sobre los PDP (Panel de plasma)

Mucho se habla de la resolución de los paneles de plasma, que suele tener, en la mayoría de los casos, 848 x 480 pixels. Según esto, este visualizador **NO** cumpliría con la etiqueta “HD Ready” pero, la mayoría de estos visualizadores, suelen llevar un circuito electrónico, de mayor o menor calidad, que hace un reescalado de la señal, permitiendo que una señal de alta definición se muestre en el visualizador sin problemas. **No** veremos la señal nativa de alta definición pero, **SÍ** podremos ver imágenes de alta definición en ese visualizador, ya que, de no llevar ese circuito, no veríamos nada, de ahí la etiqueta “HD Ready”

(*) El visualizador “HD ready” ha de soportar, como mínimo, la entrada HD en formato de componentes analógicos de video, más conocido por las siglas Y, Pb, Pr, para asegurarse la total compatibilidad con las fuentes HD existentes hoy en el mercado. La señal por componentes YPbPr se realiza mediante conexiones estándar con tres jacks RCA ó BNC con adaptador. [17]

4.2 BALANCEO INTELIGENTE DE VIDEO

Introducción

Si su trabajo es el de crear soluciones para las presentaciones, existen altas probabilidades de que haya encontrado el gran desafío para presentar videos estándar para televisión en los proyectores LCD y DLP actuales de alta resolución. A pesar de que muchos de estos dispositivos ofrecen una entrada de video, posiblemente usted haya descubierto que el uso previo de un duplicador o cuadruplicador de líneas para convertir un formato de video para televisión a uno de computadora, proporciona un resultado final mejor. Y si alguna vez intentó mostrar un video para televisión en una de las nuevas pantallas planas TFT, habrá notado que estos dispositivos no tiene una entrada directa para videos de televisión, lo que hace que los duplicadores o cuadruplicadores de líneas resulten indispensables.

No obstante, existe un nuevo y revolucionario concepto acerca de la conversión de videos para televisión a videos de computadoras. El balanceo Inteligente de video hace que los

duplicadores y cuadruplicadores de líneas sean, sino obsoletos, al menos una segunda opción muy distante.

La siguiente guía analizará a continuación de qué se trata exactamente el balaceo inteligente de video, cómo difiere de los duplicadores y cuadruplicadores de líneas tradicionales, y por qué puede ofrecerle videos más claros, definidos y brillantes para ser exhibidos en pantallas de proyectores LCD y DLP, o pantallas de plasma.

4.2.1 Conversión de videos para televisión a videos de computadora

¿Qué implica esto?

Al margen del tipo de enfoque que se utilice para la conversión de una señal entrelazada de televisión en una señal de video no entrelazada para computadora, el hecho sigue siendo que los dos formatos de señal son intrínsecamente diferentes y cualquier dispositivo de conversión deberá adaptarse a estas diferencias.

4.2.2 Resolución

La señal de televisión tiene un número predeterminado de líneas horizontales. Para NTSC ese número es de 525 líneas, mientras que para PAL y SECAM es de 625. Sin embargo, no todas estas líneas son visibles. De hecho, sólo 483 líneas de NTSC son visibles, así como en PAL y SECAM sólo es posible visualizar 576 líneas en la pantalla de televisión. Las líneas restantes se llaman líneas vacías, no contienen información sobre la imagen y se esconden en los bordes superior e inferior de la pantalla.

En contraposición, el número de líneas horizontales de una pantalla de computadora puede variar bruscamente, desde resoluciones muy bajas de 480 líneas horizontales visibles, hasta resoluciones altísimas de 1280 líneas o más. Las computadoras más nuevas vienen provistas de tarjetas de video que le permiten al usuario elegir diferentes tipos de resoluciones de pantalla. Obviamente, a resoluciones más altas, mayor es la definición y claridad de los detalles y los textos. Por ejemplo, una pantalla de computadora compuesta por 768 líneas horizontales es capaz de contener y mostrar más detalles que una imagen de computadora compuesta sólo por 480 líneas, o una imagen de televisión de 483. El número

relativamente bajo de líneas horizontales de una imagen de video para televisión, limita la capacidad de mostrar textos muy pequeños u otros detalles visuales complejos.

4.2.3 Señales Entrelazadas versus No entrelazadas

Una señal de televisión "entrelazada" significa que cada "pantalla" completa de información está compuesta por dos "campos" independientes, las líneas pares y las impares. Las líneas impares se dibujan primero en la pantalla. Luego se dibujan las pares, entre las impares, antes que éstas se hayan esfumado completamente. Esto sucede tan rápido, que es muy difícil que el ojo humano lo pueda percibir, particularmente cuando se está visualizando una imagen con movimiento. No obstante, el formato entrelazado es inadecuado para la exhibición correcta y clara de ciertos objetos estáticos, especialmente aquellos con líneas horizontales muy finas. La aparición y desaparición de los campos pares e impares de la imagen causa un parpadeo visible de las líneas finas horizontales como los que se pueden apreciar en las hojas de cálculo.

Por el contrario, la señal de la computadora fue específicamente diseñada para mostrar imágenes detalladas sin movimiento, tales como los textos pequeños y las hojas de cálculo. Esto se logra mediante una señal no entrelazada, también llamada de "barrido progresivo". Este barrido progresivo, dibuja las líneas horizontales en la pantalla de la computadora progresivamente, de izquierda a derecha y de arriba a abajo, en una sola pasada. Por ende, el problema del parpadeo que ocurre en la señal entrelazada del video para televisión, no existe en el video de computadora.

4.2.4 Formato de la señal

Los videos para televisión están definidos por los estándares NTSC, PAL o SECAM, que pautan la cantidad de líneas de una imagen, el modo en que se define la información de colores y la velocidad a la cuál las líneas son dibujadas en pantalla, de arriba a abajo (velocidad de actualización).

No obstante, dentro de NTSC, PAL y SECAM, existen diversos formatos de señales que cumplen con estos estándares. El formato más frecuentemente usado es el video

combinado. En el video combinado, toda la información - para el rojo, verde, azul (RGB) y el sync - está combinada en una señal única. El video-S, que ofrece una calidad de imagen superior, separa la información cromática (color) de la información de luminosidad y del sync. Otras variantes de NTSC y PAL incluyen RGB a 15 kHz, como el video componente y el video-D (digital). Cabe recordar que aún cuando todos estos formatos difieren en la forma en que la información del video es combinada en una única señal, presentan ciertas características en común. Todas son señales entrelazadas, todas tienen la misma cantidad de líneas visibles, ya sea 483 (NTSC) o 576 (PAL y SECAM), y tienen una velocidad de actualización que es invariable y predeterminada. En el caso de NTSC, dos campos entrelazados que comprenden un mismo marco, son dibujados en la pantalla 30 veces por segundo (es decir, una velocidad de 30 Hz), y en el caso de PAL y SECAM, esto sucede 25 veces por segundo (25 Hz).

Contrariamente a lo que ocurre con los videos para televisión, no existe un estándar que los videos de computadora deban cumplir. Como mencionamos anteriormente, existe un amplio rango de resoluciones de pantalla que son comúnmente usadas. Asimismo, existe un amplio rango de velocidades de actualización, la mayoría comprendidas entre los 60 y 85 Hz. Y aún cuando casi todas las pantallas de las computadoras son de tipo no entrelazadas, algunas tarjetas de video ofrecen la opción entrelazada para la exhibición en pantalla.

No obstante, lo que sí tienen en común todas las señales de video es el modo en que describen la información cromática y luminosa al monitor. Todos los formatos para video VGA, SVGA y Mac, transmiten la información del rojo, verde y azul (RGB) como señales independientes. (Existen diferencias acerca de cómo las computadoras transmiten la información sync en combinación con las señales de color). El hecho de que las computadoras mantengan las señales rojo, verde y azul separadas, permite que los monitores puedan mostrar una vasta variedad de colores con una mínima distorsión.

4.2.5 Duplicadores y Cuadruplicadores de Líneas

¿Cómo funcionan?

¿Cómo realizan los duplicadores y cuadruplicadores de líneas, la conversión de videos de televisión a videos de computadora?

Los duplicadores y cuadruplicadores de líneas generan una señal de video para computadora que es "una función de" o está "relacionada con" la señal original del video de televisión. Ejecutando determinados procesos sobre la señal de televisión entrelazada, son capaces de modificarla de modo tal que resulta admisible por los equipos que reciben solamente señales no entrelazadas de videos para computadora . No obstante, el tipo de señal original que ingresa en el duplicador o cuadruplicador determina el tipo de señal de salida que se puede obtener finalmente.

Cómo analizamos en la sección anterior, existen 483 líneas visibles en una señal NTSC, y 576 en PAL y SECAM. Un duplicador de líneas toma estas cantidades de líneas entrelazadas y las convierte en un formato no entrelazado. Sin embargo, el número total de líneas es el mismo - 483 y 576, respectivamente - sólo que se producen dos veces. Los cuadruplicadores difieren de los duplicadores porque crean una línea adicional para cada línea original, convirtiendo una señal NTSC en una señal no entrelaza de 966 líneas visibles, y una señal PAL o SECAM en 1152 líneas visibles.

La velocidad de actualización de la señal convertida, generada por los duplicadores y cuadruplicadores de líneas, también es una función del formato de la señal de entrada. La señal NTSC tiene una velocidad de actualización de 30 Hz. Esto quiere decir que una "pantalla" completa de información es dibujada 30 veces por segundo. Los duplicadores y cuadruplicadores de líneas, simplemente duplican esta velocidad de actualización de 30 HZ, creando una señal de salida con una velocidad de actualización de 60 Hz. Cuando la señal original de entrada es una señal PAL o SECAM con una velocidad de actualización vertical de 25 Hz, la señal de salida convertida por los duplicadores y cuadruplicadores tiene una velocidad de actualización de 50 Hz.

¿Qué encierra un nombre?

Muchas personas creen erróneamente que los duplicadores de líneas emiten una salida con dos veces la cantidad original de líneas de la imagen de video, y los cuadruplicadores, con cuatro veces esta cantidad. Esto no es exactamente cierto. Los duplicadores de líneas duplican el número de líneas dibujadas en pantalla, en cada pasada individual de arriba hacia abajo, mientras que los cuadruplicadores las cuadruplican. Sin embargo, cuando se utiliza un duplicador, el número total de líneas que forma la imagen sigue siendo el mismo que el original, pero se duplica cuando se utiliza un cuadruplicador. (¡Sí, es muy confuso!)

"Duplicar" y "cuadruplicar" pueden referirse también, al cambio de la velocidad de actualización horizontal de la salida del video, con respecto a la salida original del video para televisión. En otras palabras, cada línea horizontal individual de salida convertida por un duplicador, se "dibuja" en la pantalla, de izquierda a derecha, dos veces más rápido que cada línea individual del video de televisión original. Utilizando un cuadruplicador, cada línea individual se genera a una velocidad cuatro veces mayor.

4.2.6 Un comentario importante sobre la compensación del movimiento

Los diferentes duplicadores y cuadruplicadores de líneas, y balanceadores de video (aún en estudio), ofrecen distintos grupos de características. No es el propósito de esta guía evaluar estas variadas características y sus funciones. No obstante, la compensación del movimiento es una de las características que merece ser analizada dentro del contexto de esta guía. A fin de ayudarlo a comprender de qué se trata, resulta más fácil explicar los problemas que pueden ocurrir.

Como hemos discutido anteriormente, una de las funciones de los duplicadores y cuadruplicadores de líneas, y de los balanceadores de video, es la de convertir una señal entrelazada en un formato no entrelazado. Un modo en que los productos logran esto es almacenando el primer campo de una imagen entrelazada entrante y "manteniéndola suspendida" por una fracción de segundo, mientras que el segundo campo, que contiene el resto de las líneas de la imagen, es recibido. Los campos pares e impares son combinados

en un marco único que es capaz de salir bajo un formato no entrelazado, al doble de la velocidad original del marco.

Ahora bien, imagine que el video para televisión que usted está convirtiendo a video de computadora, es un juego de fútbol. En muchos de los marcos, habrá un juego de fútbol moviéndose a lo largo de la pantalla. En el video NTSC original para televisión, los campos impares de la imagen se producen en una 60a. de segundo, aparte de los campos pares. El juego de fútbol, al moverse rápidamente, se verá levemente distinto en los campos pares que en los impares. Luego, cuando los campos sean combinados en un único marco no entrelazado, el juego de fútbol aparecerá con bordes supuestamente dentados o mellados, debido a que la ubicación de la imagen en las líneas pares diferirá levemente de la que tendrá en las impares. El término informal y no tan técnico de este problema es "dientes o mellas" (El nombre técnico es "distorsión temporal").

Existen muchas técnicas usadas por los duplicadores y cuadruplicadores de líneas, y balanceadores de video para eliminar, o al menos reducir, la aparición de "dientes o mellas". La eficacia, así como también el costo asociado a estas técnicas varían mucho y no todos los duplicadores y cuadruplicadores de líneas, ni todos los balanceadores de video ofrecen esta característica.

4.2.7 Limitaciones de los Duplicadores y Cuadruplicadores de líneas

Las limitaciones de los duplicadores y cuadruplicadores de líneas se deben a la naturaleza de su propio funcionamiento. Dado que el formato de la señal de entrada afecta aspectos claves de la señal de salida convertida, esto es el número de líneas y la velocidad de actualización vertical, los duplicadores y cuadruplicadores tienen limitaciones en cuanto al tipo de señales de salida que pueden ofrecer.

Como mencionamos anteriormente, cuando se comienza con una señal NTSC, los duplicadores siempre producen una señal con 483 líneas visibles no entrelazadas y los cuadruplicadores generan una imagen con 966 líneas. Cuando se utiliza una señal PAL o SECAM, los duplicadores crean una imagen con 576 líneas visibles no entrelazadas mientras que los cuadruplicadores crean una de 1152.

Si usted está familiarizado con las resoluciones estándar de las computadoras, sabrá que no existen resoluciones estándar que tengan este número de líneas. Por ello, las incompatibilidades aparecen cuando se trata de recibir una imagen con un número de líneas distinto de aquél que el diseño del proyector o monitor puede admitir. Algunas de las siguientes dificultades pueden presentarse: 1) la imagen aparece recortada; 2) la imagen se ve estirada o distorsionada; o 3) el proyector o la pantalla ejecutan un procesamiento adicional de la imagen entrante para hacerla compatible con su formato de exhibición. Este reprocesado adicional prácticamente degrada la calidad de la imagen.

Otra limitación es la velocidad de actualización vertical predeterminada producida por los duplicadores y cuadruplicadores.

La señal de salida convertida de un duplicador o cuadruplicador siempre tiene una velocidad de actualización vertical de 60 Hz o de 50 Hz, según cuál sea el tipo de señal original: NTSC o PAL/SECAM. La mayoría de los monitores y proyectores norteamericanos son capaces de recibir imágenes a una velocidad de actualización de 60Hz, y aquellos países que utilizan estándares PAL o SECAM, tienen equipos están diseñados para tomar señales a 50 Hz. No obstante, muchos proyectores tienen la capacidad de recibir y mostrar rangos de actualización mayores, ofreciendo así una imagen mejor. Cuando usted usa un duplicador o cuadruplicador, está limitado a usar un rango más bajo, de 60Hz o 50 Hz.

Por último, muchos duplicadores y cuadruplicadores de líneas sólo pueden generar imágenes con una proporción dimensional de 4:3. Esto les impide ser capaces de exhibir videos con formato de buzón en las nuevas pantallas de plasma de 16:9, o monitores CRT (Tubo de Rayos Catódicos). Y en el futuro, a medida que surjan más y más aparatos nuevos, como las pantallas LCD que ofrecen exhibiciones de 16:9, los duplicadores y cuadruplicadores de líneas no proporcionarán un medio de aprovechamiento para estas capacidades.

Habiendo estudiado la tecnología de los duplicadores y cuadruplicadores con sus consiguientes limitaciones, analizaremos ahora una nueva alternativa a estas técnicas: El balanceo inteligente de video.

4.2.8 Balanceo Inteligente de Video

¿De qué se trata?

El balanceo inteligente de video es una nueva concepción acerca de la conversión del video para televisión al formato de video de computadora. Más que generar una salida que depende del formato de entrada, un balanceador de video es capaz de producir una imagen convertida en un amplio rango de resoluciones y rangos de actualización, completamente independientes del formato original de entrada del video de televisión.

¿Cómo logra hacer esto?. Al igual que los duplicadores y cuadruplicadores de líneas, el balanceador de video combina la información de los campos pares e impares de una señal de video entrante pasándola a una imagen combinada no entrelazada. Luego, aprovechando los últimos avances de la tecnología electrónica, el balanceador de video utiliza sofisticados algoritmos de procesamiento para trabajar la imagen, cambiándole su resolución, su velocidad de actualización, y hasta su proporción dimensional, a fin de corresponderse exactamente con las especificaciones de salidas requeridas. En lugar de generar resoluciones "anticuadas" de 483 o 966 líneas, los balanceadores de video ofrecen salidas convertidas a resoluciones estándar de 640 x 480, 800 x 600, y aún mayores, como las 1280 x 1024, todas en una gran variedad de rangos de actualización.

Ventajas del Balanceo de Video

Las ventajas del balanceo de video son obvias. Este elimina las muchas limitaciones impuestas por la tecnología de duplicación y cuadruplicación de líneas.

- Cuando se utiliza un balanceador de video, el proyector o monitor, no se ve forzado a recortar, distorsionar o reprocesar posteriormente la imagen para mostrarla en pantalla.
- Los balanceadores de video pueden proporcionar salidas a múltiples rangos de actualización. Un ejemplo de cómo esta característica puede resultar beneficiosa es la capacidad de exhibir videos PAL convertidos, en proyectores que no soportan rangos de actualización de 50 Hz. Otro ejemplo sería la exhibición de los videos convertidos

mostrándolos a 75 Hz, eliminando todos los parpadeos visibles, mejorando la salida de luz del monitor o proyector y relajando la tensión visual.

- Los balanceadores de video no están limitados a una proporción dimensional de 4:3. Esto abre nuevas oportunidades para la exhibición (16:9) de videos con formato de buzón.

Además de todas estas ventajas, hay una más que requiere una pequeña explicación. Los balanceadores de video permiten generar una salida que equipara exactamente a la resolución original de su pantalla LCD, proyector DLP o pantalla de plasma - algo imposible cuando se utiliza un duplicador o cuadruplicador de líneas.

¿Qué es la "Resolución Original" o "Mancha Ideal"?

Los proyectores LCD y DLP, así como las pantallas de plasma, son el tipo de productos más frecuentemente utilizados en conjunto con los duplicadores y cuadruplicadores de líneas, y los balanceadores de video. Y, al revés de lo que ocurre con los monitores CRT (monitores estándar para computadoras), este tipo de dispositivos tiene un número predeterminado de píxeles que forman su pantalla. Comúnmente nos referimos a este número de píxeles como la "resolución original" o "mancha ideal", y aún cuando la mayoría de las pantallas y proyectores tiene la capacidad de mostrar resoluciones adicionales más allá de su "resolución original", la mejor calidad de imagen siempre se consigue utilizando esa resolución original. Por ejemplo, imagine que está utilizando un proyector cuya resolución original es 800 x 600, pero que también puede mostrar resoluciones de 640 x 480. Es fácil comprender por qué una imagen proyectada, que actualmente tiene una información de identificación de 800 x 600 píxeles, se verá mejor que una imagen de 640 x 480 que se despliega tratando de cubrir el espacio de los 800 x 600 píxeles reales de la pantalla. (Vea la ilustración).

Dado que los balanceadores de video pueden producir una señal de salida que concuerde exactamente con la resolución original de la mayoría de los proyectores LCD y DLP y pantallas de plasma, estos dispositivos tiene la capacidad de trabajar en su máximo potencial, mostrando una imagen lo más nítido y claro posible.

Otra ventaja muy importante

¿Nos creería si le decimos que es el precio? El balanceo de video no necesita ser un proceso costoso. Gracias a los últimos avances de la tecnología en el mundo de la electrónica, los nuevos balanceadores de video son capaces de realizar conversiones de videos para televisión a videos de computadora a un precio increíblemente competitivo. Claro que la combinación de las características adicionales, como la compensación del movimiento, los controles de procesamiento de video, controles de posición, formatos de entrada y tipos de conectores influyen en el costo de los balanceadores de video del mismo modo que afectan a los duplicadores y cuadruplicadores de líneas tradicionales. Pero si comparamos el funcionamiento básico del balanceador de video con el de un duplicador o cuadruplicador de líneas, la comparación sencillamente no existe. Los balanceadores de video pueden utilizarse exactamente para las mismas aplicaciones que los duplicadores y cuadruplicadores de líneas, pero haciendo un trabajo mucho mejor y al precio de un simple duplicador de líneas y a una quinta parte del costo de muchos cuadruplicadores. [18]

4.3 VIDEO SOBRE REDES.

Introducción

La transmisión de vídeo sobre redes de telecomunicaciones está llegando al punto de convertirse en un sistema habitual de comunicación debido al crecimiento masivo que ha supuesto Internet en estos últimos años. Lo estamos utilizando para ver películas o comunicarnos con conocidos, pero también se usa para dar clases remotas, para hacer diagnósticos en medicina, videoconferencia, distribución de TV, vídeo bajo demanda, para distribuir multimedia en Internet...Debido a la necesidad de su uso que se plantea en el presente y futuro, se han proporcionado distintas soluciones y sucesivos formatos para mejorar su transmisión. Pero hoy, ya hemos oído hablar negativamente de los sistemas actuales de distribución de vídeo debido a su dudosa calidad en redes como Internet. Estas aplicaciones normalmente demandan un elevado ancho de banda y a menudo crean

cuellos de botella en las redes. Este es el gran problema al que esta sometida la transmisión de vídeo. ¿Por qué es el vídeo tan problemático?

¿Qué es el vídeo?

El vídeo no es nada más que la reproducción en forma secuencial de imágenes, que al verse con una determinada velocidad y continuidad dan la sensación al ojo humano de apreciar el movimiento natural. Junto con la imagen, el otro componente es el sonido.

4.3.1 Transmisión de vídeo: vídeo digital.

La transmisión digital y la distribución de información audiovisual permiten la comunicación multimedia sobre las redes que soportan la comunicación de datos, brindando la posibilidad de enviar imágenes en movimiento a lugares remotos. Pero no es todo tan bonito a la hora de transmitirlo por red, debido a que nos encontramos con sucesos como lentitud entre la reproducción de imágenes, errores de transmisión, o pérdidas de datos...

Existen dos formas de transmisión de datos, analógica y digital. Una de las características del vídeo es que está compuesto por señales analógicas, con lo que se pueden dar las dos formas de transmisión. En los últimos años la transmisión de datos se ha volcado hacia el mundo digital ya que supone una serie de ventajas frente a la transmisión analógica. Al verse la información reducida a un flujo de bits, se consigue una mayor protección contra posibles fallos ya que se pueden introducir mecanismos de detección de errores, se elimina el problema de las interferencias, podemos disminuir el efecto del ruido en los canales de comunicación, conseguir codificaciones más óptimas y encriptado, mezclar con otros tipos de información a través de un mismo canal, y poder manipular los datos con ordenadores para comprimirlos, por ejemplo.

Además si queremos difundir el vídeo por vías digitales tendremos que digitalizarlo, con lo que debe ser capturado en su formato analógico y almacenado digitalmente logrando así que sea menos propenso a degradarse durante la transmisión. Existen dos tipos de redes de comunicación, de conmutación de circuitos y de conmutación de paquetes. En la conmutación de circuitos, donde la comunicación está permanentemente establecida durante toda la sesión, un determinado ancho de banda es asignado para la

conexión, y el tiempo de descarga del vídeo puede predecirse, pero tienen la desventaja de que las sesiones son punto a punto y limitan la capacidad de usuarios. En la conmutación de paquetes pueden acomodarse más fácilmente las conferencias multipunto. Aquí el ancho de banda está compartido pero es variable, lo que supone una importante mejora puesto que, si el bit rate (o número de bits por segundo) es fijo la calidad de la imagen variará dependiendo del contenido de los fotogramas. Debe cumplirse que el ancho de banda, la resolución, y la compresión de audio sean idénticos para cada cliente que recibe el vídeo, lo que dificulta la configuración del sistema. El vídeo es muy sensible al retardo de la red, ya que puede provocar cortes en las secuencias. La pérdida de alguna información en el vídeo sin comprimir no es muy relevante, ya que al perderse un fotograma, el siguiente fotograma proporciona la suficiente información para poder interpretar la secuencia. En cambio el vídeo comprimido es mucho más sensible a errores de transmisión, ya que las técnicas de compresión que se valen de la redundancia espacial y temporal pueden perder la información de esta redundancia y los efectos de la falta de datos pueden propagarse en los próximos fotogramas. Es por eso que actualmente la comunicación con vídeo vía Internet no prometen una elevada fiabilidad de transmisión.

Algunas técnicas de compresión compensan esta sensibilidad a la pérdida de datos enviando la información completa sobre un fotograma cada cierto tiempo, incluso si los datos del fotograma no han cambiado. Esta técnica también es útil para los sistemas de múltiples clientes, para que los usuarios que acaban de conectarse, reciban las imágenes completas.

Nos podemos preguntar cuál es la tecnología de red adecuada para las aplicaciones de vídeo, pero siempre dependeremos del entorno en el que trabajemos. Por ejemplo si disponemos de un alto ancho de banda el tipo de red adecuada sería ATM; para un entorno de red de área local podríamos usar Fast Ethernet, y actualmente para que el usuario de Internet, ADSL.

Pero la solución para resolver el cuello de botella del ancho de banda del vídeo no está en un solo tipo de red, sino en una infraestructura de red flexible que pueda manejar e integrar diferentes redes y que deje paso también a futuras redes sin cambiar el hardware. También debe ser capaz de negociar las variables de ancho de banda, resolución, número de

fotogramas por segundo y algoritmo de compresión de audio. Así que se necesita un nodo que permita la interconectividad entre todas las redes. Es el MCU (unidad de control multipunto). Cada red- RDSI, IP, ATM- usa protocolos específicos que definen la naturaleza de las ráfagas de vídeo. Las combinaciones de protocolos y estándares son muchas: para vídeo H.261 o H.263, CIF o QCIF, de 7.5 fps a 30 fps; y para audio G.711, G.728, G.722 o G.723. Por ejemplo en una conferencia múltiple el número de posibles combinaciones de estándares y protocolos es muy elevado y puede saturar el MCU. Muchos MCU no son capaces de negociar todas estas variables, forzando a los terminales de los clientes a reducir sus protocolos al más bajo común denominador de todos los participantes, bajando así la calidad del vídeo.

4.3.2 Digitalización

La información a digitalizar será la de las imágenes. Cada cuadro de la imagen es muestreado en unidades de píxeles, con lo que los datos a almacenar serán los correspondientes al color de cada pixel.

Tres componentes son necesarias y suficientes para representar el color y para ser interpretado por el ojo humano. El sistema de codificación de color usado es el RGB (Red, Green, Blue).

Para digitalizar una señal de vídeo analógico es necesario muestrear todas las líneas de vídeo activo. La información de brillo y color son tratadas de forma diferente por el sistema visual humano, ya que es más sensible al brillo que al color. Con lo que se usa un componente especial para representar la información del brillo, la luminancia, una para el color y la saturación, la crominancia. Cada muestra de color se codifica en señal Y-U-V (Y- luminancia, U y V crominancia) partiendo de los valores del sistema RGB. Con este sistema las diferencias de color pueden ser muestreadas sin resultados visibles, lo que permite que la misma información sea codificada con menos ancho de banda. Un ejemplo de conversión de señal analógica de televisión en color a una señal en vídeo digital sería:

Sistema PAL: 576 líneas activas, 25 fotogramas por segundo, para obtener 720 pixels y 8 bit por muestra a 13,5Mhz:

- Luminancia(Y): $720 \times 576 \times 25 \times 8 = 82.944.000$ bits por segundo
- Crominancia(U): $360 \times 576 \times 25 \times 8 = 41.472.000$ bits por segundo
- Crominancia(V): $360 \times 576 \times 25 \times 8 = 41.472.000$ bits por segundo

Número total de bits: 165.888.000 bits por segundo (aprox. 166Mbits/sg). Ninguno de los sistemas comunes de transmisión de vídeo proporcionan transferencias suficientes para este caudal de información. Las imágenes de vídeo están compuestas de información en el dominio del espacio y el tiempo. La información en el dominio del espacio es provista por los pixels, y la información en el dominio del tiempo es provista por imágenes que cambian en el tiempo. Puesto que los cambios entre cuadros colindantes son diminutos, los objetos aparentan moverse suavemente.

El valor de luminancia de cada pixel es cuantificado con ocho bits para el caso de imágenes blanco y negro. En el caso de imágenes de color, cada pixel mantiene la información de color asociada; una imagen completa es una composición de tres fotogramas, uno para cada componente de color, así los tres elementos de la información de luminancia designados como rojo, verde y azul, son cuantificados a ocho bits. Pero la transmisión digital de vídeo tiene también alguna desventaja respecto a la analógica, por ejemplo, en una videoconferencia, cuando distintos usuarios envían sonido al mismo tiempo, si el proceso fuera analógico las distintas ondas se sumarían y podríamos escuchar el conjunto de todas ellas. Al ser digital, los datos llegan en paquetes entremezclados, lo que dificulta la compresión.

Tipos comprimido/descomprimido

Como hemos dicho para cada punto de la imagen se le asigna un determinado número de bits que representarán el color de dicho punto. Si la imagen es en blanco y negro, bastará un bit para representarlo, mientras que para 256 colores serán necesarios 8 bits. De esta forma tendremos la imagen digitalizada, pero almacenar esta información dependerá del número de pixels que utilicemos por imagen. Por ejemplo una imagen de 640 x 480 puntos con 256

colores ocupa 300 Kb, y si tenemos una secuencia de vídeo a 25 fotogramas por segundo significaría que un solo segundo ocuparía 7.500 Kb. Y todo esto sin contar el audio. La información de vídeo compuesta de esta manera posee una cantidad tremenda de información; por lo que, para transmisión o almacenamiento, se requiere de la compresión de la imagen.

La compresión del vídeo generalmente implica una pérdida de información y una consecuente disminución de calidad. Pero esto es aceptable porque los algoritmos de codificación están diseñados para descartar la información redundante o que no es perceptible por el ojo humano. Aunque sabemos que la calidad del vídeo es inversamente proporcional al factor de compresión.

La compresión es un arma de doble filo, ya que el vídeo comprimido es más sensible a los errores. Un error en vídeo comprimido puede hacer ilegible la imagen, con lo que se añade redundancia para recuperar esa información.

El vídeo comprimido en general debe transmitir información por un canal más pequeño del que necesitaría para ser transmitido y poder ser visualizado en tiempo real. Así la información de audio y vídeo deben ser procesadas por los codecs antes de ser transmitidos. Los codecs derivan de las palabras compresor y descompresor, y son los módulos de software que permiten la compresión y descompresión de los ficheros de audio y vídeo para que puedan ser transmitidos por redes de baja velocidad.

La digitalización y la compresión pueden darse conjuntamente y en tiempo real para facilitar la comunicación y la interacción. Los codecs más utilizados son los siguientes: Microsoft Video1, Microsoft RLE, Intel Indeo R2, Intel Indeo R3, Intel YUV9, CinePak, Captain Crinch, Creative Compressor.

Las señales recibidas deben ser decodificadas antes de poder ser visualizadas por el usuario. Durante este proceso se puede producir:

- lo que se llama "vídeo fantasma" o suavización de imagen, que es la forma con la que los codecs compensan los elevados flujos de información. Cuando ocurre esto, el codec

comprime la información reduciendo el "frame rate" (número de imágenes por segundo), el cual puede hacer que los movimientos rápidos parezcan borrosos. El codec también modifica la resolución para comprimir la información lo cual puede hacer que la imagen se vea desplazada. Entonces, para reducir estos efectos, se disminuye el flujo de información visual.

También puede darse un retardo de audio. En la red de Internet por ejemplo la mayoría de los usuarios están conectados a velocidades de 56.6 kilobits por segundo (Kbps), 33.6 kbps o 28.8 kbps, y el vídeo descomprimido para ser enviado en calidad broadcast requiere un ancho de banda de red de 160 megabits por segundo (Mbps), en calidad CD requiere aproximadamente 2.8 Mbps, y con los modems actuales sería imposible conseguir las velocidades requeridas para su transmisión. Aquí es donde juegan un papel importante los codecs.

Los codecs se optimizan para conseguir la mayor calidad posible en bajos índices de transferencia. Son usados para codificar el vídeo en tiempo real o pregrabado y ser mandado por la red para que el usuario final solamente con una aplicación que lo descomprima puede al instante visionar en su terminal.

Compresión

La técnica de compresión de vídeo consiste de tres pasos fundamentalmente, primero el preprocesamiento de la fuente de vídeo de entrada, paso en el cual se realiza el filtrado de la señal de entrada para remover componentes no útiles y el ruido que pudiera haber en esta. El segundo paso es la conversión de la señal a un formato intermedio común (CIF), y por último el paso de la compresión. Las imágenes comprimidas son transmitidas a través de la línea de transmisión digital y se hacen llegar al receptor donde son reconvertidas al formato común CIF y son desplegadas después de haber pasado por la etapa de post-procesamiento. Mediante la compresión de la imagen se elimina información redundante. Se ayuda de la redundancia espacial y temporal. La redundancia temporal es reducida primero usando similitudes entre sucesivas imágenes, usando información de las imágenes ya enviadas. Cuando se usa esta técnica, sólo es necesario enviar la diferencia entre las imágenes, es decir las zonas de la imagen que han variado entre dos fotogramas consecutivos, lo que elimina la necesidad de transmitir la imagen completa. La compresión espacial se vale de

las similitudes entre píxeles adyacentes en zonas de la imagen lisas, y de las frecuencias espaciales dominantes en zonas de color muy variado.

El método para eliminar las redundancias en el dominio del tiempo pueden ser eliminadas mediante el método de codificación de intercuadros, que también incluye los métodos de compensación/estimación del movimiento, el cual compensa el movimiento a través de la estimación del mismo.

En el otro extremo, las redundancias en el dominio espacio es llamado codificación intracuadros, la cual puede ser dividida en codificación por predicción y codificación de la transformada usando la transformada del coseno.

La transformada del coseno o DCT es una implementación específica de la transformada de Fourier donde la imagen es transformada de su representación espacial a su frecuencial equivalente. Cada elemento de la imagen se representa por ciertos coeficientes de frecuencia. Las zonas con colores similares se representan con coeficientes de baja frecuencia y las imágenes con mucho detalle con coeficientes de alta frecuencia. La información resultante son 64 coeficientes DCT. El DCT reordena toda la información y la prepara para la cuantización.

El proceso de cuantización es la parte del algoritmo que causa pérdidas. La cuantización asigna un número de bits específico a cada coeficiente de frecuencias y entonces comprime los datos asignando unos cuantos bits a los coeficientes de alta frecuencia. sin que lo note el observador. Los parámetros de la cuantización son optimizados, pero el proceso aún deteriora la calidad del vídeo. Generalmente se acepta que un factor de compresión de 2:1 (aproximadamente 10Mb/seg), se pueden apreciar visualmente algunas pérdidas en la integridad del vídeo.

El proceso de decodificación es básicamente el inverso del proceso de codificación. La compresión del audio está descrita por tres parámetros: ratio de muestreo (numero de muestras por segundo), bits por muestra (numero de bits para representar cada valor), y número de canales (mono o estéreo).

Los estándares de vídeo digital más conocidos son: MPEG, Quicktime, AVI, MOV, real vídeo, ASF...Y para vídeo analógico: NTSC, PAL, SECAM

4.3.3 Soluciones de vídeo

El vídeo puede servirse como un fichero, o en tiempo real. A esta última forma de enviar el vídeo se le conoce como streaming.

Streaming video.

Streaming video, o vídeo en tiempo real, es la tecnología que permite la transmisión y recepción de imágenes y sonidos de manera continua a través de una red. A diferencia de otros formatos de audio y vídeo, en los que es necesario esperar que el archivo sea cargado en el equipo para su visualización, esta tecnología permite apreciar el contenido conforme se va teniendo acceso a la información del archivo.

EL servidor de streaming permite visionar el vídeo de forma continua porque hace uso de un buffer, donde van cargándose algunos segundos de la secuencia antes de que sean mostrados. Entonces cuando se detecta un periodo de congestión de red, se visualizarán los datos que tenemos ya almacenados en el buffer. De esta forma el cliente obtiene los datos tan rápido como el servidor y la red lo permitan. Hay pocos formatos hoy en día que soporten este tipo de visualización progresiva, probablemente en el futuro próximo, el estandar para el streaming vídeo será en Advanced streaming format (ASF). El streaming puede decirse que funciona de forma inteligente ya que asegura al usuario que recibirá la más alta calidad posible dependiendo de la velocidad de conexión o de los problemas de conexión de la red. Tradicionalmente la congestión de la red forzaba al usuario a detener la visualización del vídeo almacenando en un buffer la información para posteriormente continuar mostrando la secuencia. Con los nuevos formatos de streaming como el MPEG-4, el cliente y el servidor pueden degradar la calidad de forma inteligente para asegurar una reproducción continua del vídeo.

¿Cómo funciona?

Si se dan problemas de congestión de red, primeramente el servidor de vídeo disminuye el número de fotogramas que está enviando para mantener la calidad del audio e ir llenando el búffer mínimamente. Si las condiciones empeoran, el servidor deja de mandar frames de vídeo, pero mantiene la calidad del audio. Finalmente, si la calidad del audio empieza a degradarse, el cliente reconstruye de manera inteligente las secuencias que tiene almacenadas para no perder calidad.

Video en Internet.

Existen dos métodos para la distribución de contenido con audio y vídeo sobre la Web. El primer método usa un Servidor Web estándar para repartir los datos a un medio visualizador. El segundo método usa un servidor de streaming. ¿Cómo funciona un servidor web para distribución de vídeo? Una vez disponemos del vídeo digitalizado el archivo será codificado o comprimido a un fichero para ser distribuido sobre una red con un específico ancho de banda como un modem de 56.6 kbps. Entonces el fichero se almacena en un servidor web. Ahora sólo hemos de crear una página web en un servidor con un enlace al fichero, el cual cuando sea pulsado por un cliente permitirá la descarga automática.

El protocolo usado es el HTTP (Hypertext Transport Protocol), que opera en la parte alta del TCP el cual maneja la transferencia de datos. Este protocolo no está optimizado para aplicaciones en tiempo real, con lo que el objetivo es maximizar el ratio de transferencia, para lograrlo usa un algoritmo llamado "comienzo lento", TCP manda primero datos con un bajo ratio, y gradualmente va incrementando el ratio hasta que el destinatario comunica una pérdida de paquetes. Entonces el TCP asume que ha llegado al límite del ancho de banda y vuelve a enviar datos a baja velocidad, y a volverá a incrementar la velocidad repitiendo el proceso anterior. TCP se asegura una fiable transmisión de datos con la retransmisión de los paquetes perdidos. Sin embargo lo que no puede asegurar es que todos los paquetes recientes llegarán al cliente para ser visualizados a tiempo, con lo que podremos experimentar pérdida de imágenes en las secuencias de vídeo.

Una solución sobre Internet.

Internet no puede considerarse un medio adecuado para la transmisión de vídeo en tiempo real. La calidad de los vídeos transmitidos en tiempo real es bastante pobre con lo que debes elegir poca velocidad y mucha calidad o ganar velocidad y perder calidad en imagen. Pese a todo esto existen soluciones desarrolladas que con las mejoras de las técnicas y de la velocidad de los modems han ido evolucionando.

Veamos como podríamos montar un servidor de vídeo como podría ser el que hasta ahora ha sido el estándar de transmisión de vídeo sobre Internet, RealVideo. Para poder distribuir vídeo sobre Internet, y especialmente un sistema de stream vídeo, los componentes que necesitamos son un codificador para digitalizar el vídeo y comprimirlo, un software de servidor web, que puede ser una máquina distinta o la misma que el codificador, y una conexión a la red con suficiente ancho de banda, dependiendo del número de usuarios a los que queremos dar servicio. El usuario final necesitará solamente el programa cliente para descargar y visualizar los flujos de vídeo.

Videoconferencia.

Como dijimos, el problema básico del vídeo es el gran ancho de banda que requiere. En videoconferencia suele trabajarse con ventanas de 300x200 pixeles. El vídeo estándar utiliza 30 imágenes (frames) por segundo, por tanto, 30 imágenes de 60 KB dan la friolera de 1,8 millones de bytes por segundo, la mayoría de los fabricantes se orienta hacia la adopción de la RDSI.

Otra solución al cuello de botella del ancho de banda en videoconferencia es la compresión de las imágenes. Complejos algoritmos basados en la transformada discreta del coseno (DTC) suelen ser los más utilizados. Lo más significativo de la compresión es la ratio que se puede alcanzar.

El formato JPEG permite definir la calidad de las imágenes, es decir, cuanta pérdida queremos aceptar en la compresión.

Los sistemas de videoconferencia establecen la compresión mediante un dispositivo de hardware y/o software llamado códec (codificador/decodificador), encargado además de muestrear las señales de audio y vídeo así como de transmitir las.

4.3.3.1 APLICACIONES

El gigantesco desarrollo actual de la MultiMedia y la proyección que se le puede suponer indica una invasión de productos tanto en la empresa como en el domicilio. Estos productos van a ser utilizados en cada caso para diversas aplicaciones, entre las que podemos diferenciar dos grandes grupos:

- Aplicaciones profesionales. En este apartado podemos citar:

- Aulas de formación.
- Edición de prensa.
- Edición de CD-ROM's y CDI.
- Aplicaciones médicas.
- Aplicaciones en la Red Global.
- Visualización de telemantenimientos.
- Teletrabajo.
- Aplicaciones publicitarias y promocionales.
- ...

- Aplicaciones domésticas:

- De aprendizaje.
- Lúdicas.
- De consumo.
- Integración con otros medios del hogar.
- ...

4.3.4 CONECTIVIDAD

Conceptos básicos

Después del repaso general a lo que globalmente podemos definir como terminal audiovisual, vemos que el siguiente paso es la conectividad entre diversos terminales, tanto

para formar redes especializadas para la ejecución de complejos proyectos de infografía, como para diseñar ramas multimedia en redes corporativas.

Además, para una conexión a Internet, dando por supuesto que tanto el servidor como el medio proporcionan las características adecuadas, es fundamental disponer de un potente terminal.

Los cuatro puntos de apoyo en que debe basarse una red con proyección de futuro para tener un tiempo de vida razonable son, independientemente de la tecnología:

Cableado y electrónica de red

Una cadena es tan buena como el más débil de sus eslabones. Por eso, para plantear una red de un modo profesional, es fundamental que la base física de la conectividad cumpla con las especificaciones de calidad necesarias para un funcionamiento óptimo. Es decir, que el primer punto a tener en cuenta es definir el tipo de cable en función del tráfico y el tipo de instalación, de modo que pueda certificarse posteriormente, eliminando de este modo una fuente habitual de fallos, consistente en no dar importancia a este primer y básico paso.

Servidor y terminales

La elección de máquinas con garantías, buen soporte técnico y con baja tasa de fallos es el segundo paso. De ello hablaremos posteriormente.

Hardware y software multimedia

Tanto en el servidor como en los terminales, debemos disponer de los mejores complementos, aplicados los criterios de selección que anteriormente hemos citado, tanto para sistemas infográficos como multimedia.

Software de red y aplicaciones

En este apartado debemos definir el software de red a utilizar, mediante criterios de fiabilidad, potencia, actualización y ampliación, compatibilidad de las aplicaciones, etc..., siendo un valor añadido muy importante los programas de prevención de fallos y las soluciones para telemantenimiento.

4.3.5 CONECTIVIDAD EN INFOGRAFÍA: SERVIDORES DE VIDEO Y AUDIO.

Nadie puede pasar por alto los cambios en los procesos de grabación que están teniendo lugar en la industria. Los discos están aquí y las cintas están en retirada. Bueno, esto es lo que a algunos nos gustaría creer. Por lo que respecta a todo lo que no sea grabación de

material, distribución de programas y archivo esta puede ser pronto la situación. Incluso algunos procesos de grabación pueden ir en breve directamente a disco. El cambio de soporte para imágenes y sonido siempre ha tenido efectos profundos. Las consecuencias de cambiar de la pantalla grande a la pequeña y de la película al video han sido enormes. De la misma manera, cambiar de cinta a disco, no tendrá un impacto menor, como ya está empezando a verse claramente.

Los sistemas digitales de video han sido tradicionalmente monopuesto, por consideraciones básicamente técnicas.

El paso siguiente ha consistido en la integración de diversas instalaciones mediante servidores de video. Es esencial que las características de éstos se adecúen a las tareas a desarrollar, ya que los servidores van a desempeñar un papel central en cualquier instalación. También es importante comprender algunos principios de la grabación de video en discos.

Los servidores de video se van a ver implicados de un modo creciente en los procesos de grabación de video y hoy en día pueden identificarse cuatro grandes áreas de aplicación para ellos:

- VOD (Video bajo demanda)
- Informativos
- Reproducción para teledifusión
- Postproducción

Los servidores para transmisión y VOD requieren un acceso relativamente lento, puesto que únicamente van a reproducir programas largos, con duraciones que van desde 5 segundos hasta 2 horas o más según los casos. Las necesidades de acceso son facilitadas además por el hecho de que estos equipos van a ser programados en su respuesta, por ejemplo, desde un sistema automatizado de noticias. Incluso los equipos de video bajo demanda se supone que pueden responder en un margen de segundos e incluso minutos.

Un servidor para transmisión, al que se le pide que detenga una repetición y comience una distinta, puede tardar varios segundos en responder, siendo entonces totalmente inservible para la operación manual o para utilización en directo.

Las áreas de postproducción e informativos son diferentes, puesto que puede ser necesario acceder en cada momento a un solo fotograma, exigiendo mucho más en cada instante al servidor.

El fundamento de un servidor es su sistema de almacenamiento, por lo que es necesario definirlo óptimamente, según una serie de parámetros:

- Acceso aleatorio real
- Capacidad de manejar video digital 601, sin compresión, o al menos 1:4
- Array de discos redundante, para seguir funcionando si falla un disco
- Escalabilidad en función de las necesidades
- Arquitectura abierta al crecimiento

El ancho de banda determina cuantos canales pueden emplearse y con que grado de compresión. Evidentemente, cuanto mayor es la compresión, mayor es el número de canales que pueden utilizarse. En el caso del video bajo demanda, se pone un mayor énfasis en el número de canales disponibles que en la calidad. En postproducción las prioridades se invierten, lo cual significa utilizar una ligera compresión, por lo que el número de canales disponibles resulta crucial.

La idea detrás de un servidor es proporcionar acceso rápido al material almacenado, pero con múltiples servidores siempre va a existir la posibilidad de conflicto; el acceso de cada usuario a los datos es un tema claramente importante. ¿Que ocurre cuando varios canales desean acceder al mismo material al mismo tiempo?. ¿Puede el servidor dar respuesta a todos o, por contra, algunos tendrán que esperar?. Una solución sería hacer suficientes copias de los datos para asegurarse la disponibilidad deseada. Sin embargo, esto incrementa las necesidades de gestión y emplea mayor cantidad del escaso bien que es el espacio de almacenamiento. Por otro lado, tal proceso de copiado debe planearse con antelación, no puede hacerse sobre la marcha, limitando así el espacio para cambios de última hora y también la flexibilidad.

Aunque una situación de acceso simultáneo por parte de varios usuarios puede parecer, en principio, poco creíble, su solución proporciona grandes beneficios y compensa el esfuerzo técnico y económico.

Esta solución significa que todos los usuarios pueden acceder a cualquier material en cualquier momento sin tener que planificarlo con antelación y sin necesidad de una copia múltiple previa. De igual modo, todos los puertos deben poder grabar al mismo tiempo.

Por supuesto, aplicaciones menos complejas requieren soluciones menos complejas.

Una gestión efectiva del sistema es esencial para operar de modo práctico. Si pensamos en una instalación completa, en algún momento el material en el servidor deberá finalmente ser almacenado como piezas completas o acabadas, tarea para el gestor del sistema. El usuario simplemente utiliza el equipo, sin tener que preocuparse por las complejidades inherentes al sistema de almacenamiento.

4.3.5.1 APLICACIONES

Prensa electrónica.

Basándonos en nuestro sistema servidor de video como punto de referencia, podemos utilizarlo para aplicaciones concretas, tales como la edición de un periódico electrónico que incluya, además de los tradicionales elementos de texto y grafismo, ficheros audiovisuales equivalentes a los editados en los noticiarios de televisión, que ¿por qué no? pueden ser los mismos que los de la cadena de TV perteneciente al mismo grupo.

La tecnología va a hacer cada vez más difusa la diferenciación entre medios, y el "lector" cada vez se aproximará más a lo que podríamos definir como un "televidente interactivo", con la revolución que va a suponer para la información, formación, entretenimiento y publicidades.

No obstante, no debemos olvidar que cambia el continente, pero es fundamental la calidad e interés de los contenidos.

Redes ATM.

La tecnología ATM emplea el concepto de conmutación de celdas, el cual combina los beneficios de la conmutación de paquetes tradicionalmente utilizada en las redes de datos y la conmutación de circuitos utilizada en redes de voz.

ATM se basa en el concepto de conmutación de paquete rápido, donde se supone una fiabilidad muy alta a la tecnología de transmisión digital, típicamente sobre fibra óptica y

por lo tanto la no necesidad de recuperación de errores en cada nodo. Ya que no hay recuperación de errores, no son necesarios los contadores de número de secuencia de las redes de datos tradicionales y tampoco se utilizan las direcciones de red, usando en su lugar el concepto de identificador de circuito o conexión virtual (VCI).

ATM ha sido definido para soportar de forma flexible, la conmutación y transmisión de tráfico multimedia, comprendiendo datos, voz, imágenes y video. En este sentido, soporta servicios en modo circuito y en modo paquete.

Sin embargo, a diferencia de la conmutación de circuitos, ATM no reserva slots para la conexión. En su lugar, una conexión obtiene slots o celdas solo cuando está transmitiendo información. Cuando una conexión está en silencio no utiliza slots, estando estos disponibles para otras conexiones. Con esta idea de partida, se ha decidido que la unidad de conmutación y transmisión sea de tamaño fijo y longitud pequeña. Esta unidad es conocida como celda y tiene una longitud de 53 bytes divididos en 5 de cabecera y 48 de información o carga útil.

Las celdas pequeñas y de longitud constante son ventajosas para tráfico con tasa de bit constante (voz, video) y son muy útiles en general, ya que permiten un tiempo de latencia muy bajo, constante y predecible, así como una conmutación por hardware a unas velocidades muy elevadas. También, en el caso de pérdidas de celdas por congestión o corrupción, la pérdida no es muy grande, siendo en muchos casos remediable o recuperable. La tecnología ATM está orientada a conexión. Esto significa que antes de que el usuario pueda enviar celdas a la red, es necesario realizar una llamada y que ésta sea aceptada para establecer una conexión virtual a través de la red.

Un conmutador ATM podría describirse como una caja que mantiene en su interior una gran cantidad de ancho de banda, siendo cedido este recurso según aumentan las necesidades. Se dice que ATM proporciona ancho de banda bajo demanda.

Dentro de las diversas aplicaciones donde ATM puede ser una buena solución, total o parcialmente, podemos citar:

- Redes de empresa homogéneas
- Grupos de trabajo virtuales
- Desarrollos en colaboración
- Computación distribuída con uso intensivo de ancho de banda

- Videoconferencia de sobremesa multiventana
- Soporte y formación remota

4.3.6 CONECTIVIDAD EN MULTIMEDIA

Lo anteriormente expuesto, en aplicaciones más diversas y menos exigentes en general, nos lleva a la conectividad en multimedia, donde debemos contemplar dos opciones:

Rama multimedia de una red

En este caso se trata de incrementar las prestaciones de una red existente, de modo que pueda responder a aplicaciones que cada vez demanda más el mercado, tanto a nivel de presentación en los terminales como de videocomunicaciones. Los departamentos de autoedición, oficinas técnicas, etc disponen de la posibilidad de mejorar los resultados de su trabajo, aportando una nueva imagen y un nuevo contenido a los proyectos. Podemos hablar de proyectos virtuales como un modo de ver mediante realidad virtual el resultado final de un diseño de arquitectura, obras públicas, y cualquier otro sector profesional, con objeto de poder realizar modificaciones previas e interactivas. Uno de los sectores que más están utilizando estos recursos es el de restauración medioambiental.

Red multimedia

Es el siguiente paso dentro del planteamiento anterior, con incorporación de capacidades multimedia a todos los terminales de la red, y posibilidad de conexión con otras redes o a través de Internet, siempre con una presentación audiovisual cada vez de más calidad.

4.3.7 LA RED GLOBAL: INTERNET

La informática del futuro próximo podrá estar basada en los super Pcs con gran facilidad de uso y potentes capacidades de cálculo y de representación gráfica y multimedia, o bien, en las denominadas autopistas de la información, mediante sencillos terminales conectados a la red. El debate suscitado sobre una u otra opción está enfrentando estrategias de grandes compañías del mundo informático, si bien hoy por hoy, una mezcla de ambas filosofías o tecnologías de la información parece lo más prudente, ya que tomar partido por cualquiera de las dos alternativas expuestas nos alejaría de las ventajas de la no elegida.

Podemos plantear los puntos básicos sobre los que se apoya el desarrollo de Internet:

- La Red Global es un medio de comunicación
- Distinción entre marketing en la red y marketing de la red
- Relaciones extraempresariales (Clientes, proveedores..)
- Comercio electrónico
- Relaciones intraempresariales
- Filosofía "Intranet"

En base a estos puntos, nos encontramos con un fenómeno descontrolado y en crecimiento exponencial, que finalmente tendrá un cauce mucho más definido, cuando terminal, medio y servidor proporcionen servicios de verdadero interés para el periférico más importante del PC: la persona que lo utiliza.

4.3.8 APLICACIONES

Sistemas de videoconferencia.

Los sistemas de videoconferencia se presentan como una ventana abierta al mundo, tanto para comunicación interna entre sedes de una compañía como para reuniones genéricas, incidiendo en el ahorro de gastos y tiempo que permite, con la consecuente mejora de la productividad y la comunicación.

No obstante, no debemos olvidar que la comunicación cara a cara, las reuniones fuera del entorno de trabajo y el placer de viajar son propios de la naturaleza humana, y que la técnica debe utilizarse para complementar y no para sustituir, con el objetivo de establecer una mejora en las comunicaciones.

Vivimos unos tiempos en que la negociación es un elemento imprescindible, y sin duda un sistema de videoconferencia aporta su granito de arena para diseñar un plan de conversaciones sucesivas, pero... ¿cómo brindar con cava una vez cerrado el acuerdo?

Dos tipos de solución

La videoconferencia proporciona comunicaciones inmediatas, interactivas y de gran impacto visual. Con un sistema de videoconferencia se puede obtener acceso inmediato a otras personas, intercambiar información independientemente de la localización donde se

encuentren nuestros interlocutores y conseguir una comunicación donde las expresiones faciales y el lenguaje corporal puedan tener mayor significación que las propias palabras. Las posibilidades del sistema no se limitan exclusivamente a los temas de oficina, pudiendo utilizarse para aplicaciones tan diversas como:

- Reuniones con clientes para la confirmación de un diseño publicitario, arquitectónico o mecánico.
- Intercambio de opiniones sobre cuestiones técnicas o de servicio entre las oficinas centrales y las delegaciones.
- Informes de la situación de ventas a las oficinas principales.
- Reuniones frecuentes y breves entre el personal de la compañía.
- Cambios de impresiones entre investigadores universitarios.
- Reuniones de planificación sobre nuevos productos entre departamentos comerciales y de diseño.
- Formación para el personal de la empresa.
- Discusión de contratos y documentos.
- Eliminación de viajes innecesarios y potenciación de viajes más interesantes.
- Dinamización de la actividad negociadora.
- Incorporación doméstica para aplicaciones de teletrabajo.

Y así podríamos seguir añadiendo aplicaciones en cualquier campo donde se precise una mejora de las comunicaciones.

Podemos definir dos soluciones diferentes a problemas diferentes.

- Videoconferencia entre directivos, normalmente ubicada en salas de reuniones, donde el componente informático es de menor importancia que el audiovisual.

En este caso, el manejo suele ser mediante mando a distancia y menús en pantalla del televisor, incluyendo control remoto de cámaras, etc.

Normalmente el sistema va instalado en un mueble, la comunicación es vía RDSI, y existen una serie de periféricos opcionales, tales como fax/escaner, cámara de documentos, entradas y salidas de señal de video, etc siendo un parámetro fundamental es la sencillez de uso, desde mando a distancia o pantalla táctil, de modo que cualquier operación se realiza

de modo intuitivo (conexión de los equipos, control de las cámaras, envío de documentos, control de audio, etc)

Normalmente existen diversos accesorios tales como cámaras para gráficos, escaner para documentos, sistema de videopresentación, tabla de dibujo, placas gráficas, unidades adicionales de audio y monitorización, supresión de ecos, sincronizador automático de labios, etc

Una de las fórmulas de amortización es el alquiler a terceros, como complemento a una completa programación de actividades.

Estos sistemas permiten reunión de grupos de unas seis personas por localización y enlace multipunto, normalmente mediante un controlador que gestiona hasta 30 puntos.

Existe una variante -que puede presentarse como un tercer tipo de solución- conocida como "roll about", basada en equipamiento susceptible de ser trasladado de una sala a otra, lo que permite mayor flexibilidad de reuniones.

- Videoconferencia basada en PC. Está orientada a la comunicación entre técnicos o profesionales, con posibilidad de actuación interactiva sobre el mismo documento (entre dos usuarios únicamente), intercambio de documentos, etc, todo gestionado mediante un software estándar, en entorno windows, donde la diferencia entre muchas de las soluciones está en el interface de usuario, y en la disponibilidad de módulos opcionales para prestaciones complementarias.

La calidad de imagen, además de estar limitada por condicionantes técnicos según el hardware, el medio de transmisión, la velocidad, los algoritmos de compresión y los estándares, va a depender del hardware adicional (tarjeta compresora/descompresora de video). Este hardware también influirá en la conectividad exterior a sistemas audiovisuales ya instalados.

Es preferible partir de un software general y compatible, en entorno windows, personalizarlo con programación a partir de APIs, y utilizar un buen hardware de uso extendido, que permita diversos algoritmos de compresión. Si el sistema se va a utilizar únicamente para comunicación interna de la compañía, podemos usar MJPEG en vez de P64 para comprimir el video, pero si vamos a necesitar conexiones con otras compañías, deberemos ajustarnos al estándar H320, H221 o H261, tanto para transmisión como para codificación.

Otras aplicaciones

Los medios de comunicación, las revistas especializadas, las publicaciones de los profesionales, etc, no son sino meros testigos de la cantidad inmensa de proyectos innovadores que día a día surgen y cuyo destino final es la red.

Podemos citar campos tan diversos y cada vez más relacionados, como:

- Realidad virtual
- Enseñanza
- Medicina
- Hogar
- Teletrabajo y autoempleo

Donde la red cada vez tiene más presencia. Es muy sencillo visitar el museo de arte moderno de Nueva York, aprender los últimos descubrimientos de la teoría atómica o jugar con un nuevo juego cada día, del mismo modo que cada vez va a ser más sencillo consultar desde la casa de salud rural con los médicos especialistas de cualquier hospital remoto para confirmar o establecer el diagnóstico adecuado. Pero no debemos olvidar que para ver de verdad una pintura de El Bosco o Velazquez hay que ir al Museo del Prado. [19]

Aspectos a tener en cuenta

El objetivo de esta exposición ha pretendido ser el de proporcionar un poco de claridad, o al menos dirigir hacia la búsqueda de la claridad, con objeto de que el punto de partida de cualquier proyecto de conectividad, y más si incorpora infografía y multimedia, sea una definición clara de las necesidades y limitaciones. Esta es la forma correcta de poder obtener la mejor solución técnico-económica, siendo fundamental un correcto dimensionado o definición de los siguientes aspectos:

- Económicos
- Sociales
- Técnicos
- Legales
- Humanos

4.4 DVD (Disco Digital o Disco Versátil Digital).

DVD (que puede significar video Disco Digital o Disco Versátil Digital) es la siguiente generación en la tecnología de almacenamiento en disco óptico. Es esencialmente un CD más rápido y con más capacidad que puede almacenar tanto video de calidad de cine y audio de mayor calidad que un CD, así como datos de computador. El DVD intenta abarcar entretenimiento doméstico, computadoras e información de negocios en un único formato digital, reemplazando eventualmente al CD de audio, a la cinta de video, el laserdisc, CD-ROM, los cartuchos de video juegos.

El DVD se ha extendido ampliamente gracias al soporte de todos los grandes fabricantes de electrónica, de todos los grandes fabricantes de hardware de computador y todos los grandes estudios de cine y música. Con este soporte sin precedentes, el DVD ha sido el producto electrónico de más éxito de todos los tiempos a menos de tres años de su introducción. Es importante entender la diferencia entre formatos físicos (tales como DVD-Video, DVD-ROM) y formatos de aplicación (tales como DVD-Video o DVD-Audio). DVD-ROM es el formato básico que almacena datos.

DVD-Video (a menudo llamado DVD a secas) define como están los programas de video almacenados en un disco y son leídos en un lector DVD-Video o un computador con DVD. La diferencia es similar entre la que hay entre CD audio y CD-ROM. El DVD-ROM incluso incluye variaciones grabables (DVD-R, DVD-RAM, DVD-RW, DVD+RW). Los formatos de aplicación incluyen DVD-Video, grabación DVD-Video, DVD-Audio, grabación DVD-Audio Grabación chorros DVD, y SACD. Hay incluso formatos de aplicaciones especiales para consolas de videojuegos tales como la Sony PlayStation 2.

4.4.1 ¿Cuales son las características del DVD-Video?

1. Unas 2 horas de video digital de alta calidad (un disco de doble cara y doble capa puede almacenar 8 horas de video de alta calidad, o 30 horas de calidad de video VHS).
2. Soporte para películas en formato ancho en televisores estándar o de formato panorámico (formatos 4:3 y 16:9)

3. Hasta 8 pistas de audio digital (para varios idiomas, DVS, etc), cada una de las cuales con hasta 8 canales.
4. Hasta 32 pistas de subtítulos/karaoke Salto automático "sin brusquedades" de video (para múltiples argumentos o clasificaciones en un disco)
5. Hasta 9 ángulos de cámara (diferentes puntos de vista que se pueden seleccionar durante la lectura)
6. Menús y características interactivas sencillas (para juegos, preguntas, etc.
7. Textos identificativos en varios lenguajes para nombre de la película, nombre del álbum, título de la canción, actores, equipo, etc.
8. Rebobinado y avance rápido "instantáneo" (no hay pegatinas de "se bueno, rebobina la cinta" en los discos de arriendo).
9. Búsqueda instantánea por título, capítulo, pista y código de tiempo.
10. Duración (no hay desgaste con la lectura, únicamente daños físicos)
11. No es susceptible a campos magnéticos, resistente al calor.
12. Tamaño compacto (fácil de manejar, almacenar y enviar, los lectores pueden ser portátiles, la creación de copias es más barato que las cintas y laserdisc)

Nota: La mayoría de los discos no contienen todas estas características (múltiples pistas de audio, salto sin costuras, control paterno, etc) puesto que cada característica debe ser compuesta especialmente. Algunos discos podrían no permitir la búsqueda o el salto.

La mayoría de los lectores soportan como estándar las características siguientes:

1. Elección de idioma (para la selección automática de escenas de video, pistas de audio, pistas de subtítulos y menús)
2. Efectos especiales en la lectura.
3. Congelado, paso a paso, lento, rápido y escaneo (no hay reproducción marcha atrás)
4. Control paterno (para negar la lectura de discos o escenas con material obsceno)
5. Programación (lectura o selección de secciones en una secuencia deseada)
6. Lectura Aleatoria y lectura repetida.
7. Salida audio digital (PCM estéreo y Dolby Digital)
8. Compatibilidad con CDs de audio

4.4.2 ¿Como es la calidad del DVD-Video?

El DVD tiene la capacidad de producir video con calidad casi de estudio y calidad de audio mejor que la del CD. El DVD es exageradamente mejor que la cinta de video doméstico y generalmente mejor que el laserdisc. Sin embargo la calidad depende de muchos factores de producción. Mientras la experiencia en la compresión y la tecnología mejora veremos incrementarse la calidad, pero mientras los costos de producción decrecen veremos también discos pobremente producidos. Unos pocos DVDs baratos incluso usan codificación MPEG-1 (que no es mejor que VHS) en vez de el MPEG-2 de mayor calidad.

El DVD video es generalmente codificado desde las cintas master de estudio digital a formato MPEG-2. El proceso de compresión usa una pobre compresión que elimina información redundante (tales como áreas de la imagen que no cambian) e información que no es fácilmente perceptible por el ojo humano. El video resultante, especialmente cuando es complejo o de imagen cambiante rápidamente, podría a veces contener "artefactos" tales como falta de trozos de imagen, imagen borrosa y ruido de video dependiendo de la calidad del procesado y de la cantidad de compresión. A velocidades medias de 3.5 a 5 Mbps (millones de bits/segundo) estos artefactos podrían ser ocasionalmente reconocibles. Velocidades mayores pueden dar mayor calidad sin casi diferencia perceptible con el master original a velocidades sobre 6Mbps.

Mientras la tecnología de compresión MPEG avanza, se consigue mejor calidad a menores velocidades. El video proveniente del DVD a veces contiene "artefactos" visibles tales como cintas de color, borrones, débiles resplandores, detalles perdidos e incluso efectos como superficies que "flotan" con retraso detrás de imágenes en movimiento. Hay que entender que el termino "artefacto" se refiere a cualquier cosa que no estaba originalmente presente en la película. Los artefactos vienen a veces causados a veces por una pobre codificación MPEG como también es causado más a menudo por TV mal ajustadas, malos cables, interferencias eléctricas, una poco gradual reducción del ruido digital antes de la codificación o mejoras de la imagen, pobre transferencia de película a video, grano en la película, fallos en el lector, errores de lectura en discos, etc.

La mayoría de los DVDs muestran pocos artefactos visibles de compresión MPEG en un sistema bien configurado. Si piensas de otra forma estás mal interpretando lo que ves. Algunas de las primeras demos de DVD no fueron muy buenas, pero esto es simplemente una indicación de como de mala puede ser la calidad del DVD si no se procesa y reproduce correctamente. Muchos discos de demostración se abalanzaron al procedimiento de codificación con la intención de ser distribuidas tan rápido como fuese posible. Contrariamente al sentido común y aunque parezca tonto, estas demos NO están preparadas para mostrar lo mejor del DVD.

Las demos en las tiendas deben ser vistas con un poco de paciencia, puesto que la mayoría de los vendedores son incapaces de ajustar correctamente un aparato de televisión. La mayoría de los televisores tienen la nitidez muy alta para la claridad del DVD. Esto exagera el video de alta-frecuencia y causa distorsión, tal como el control de agudos de un CD fijado alto que hace que el sonido sea chillón. Muchos lectores de DVD sacan video con un nivel de negro de 0 IRE (estándar japonés) en vez de 7.5 IRE (estándar USA). En televisores que no están propiamente ajustadas esto puede causar alguna ausencia de trozos en escenas oscuras.

El DVD video tiene una fidelidad de color excepcional, de modo que colores turbios o lavados son casi siempre un problema de visualización, no del lector DVD o del disco. La calidad de audio del DVD es excelente. El DVD incluye la opción del audio digital PCM (pulse code modulation) con tamaños de muestra y velocidades mayores que el CD audio. Alternativamente, el audio para la mayoría de las películas se almacena como sonido surround multi-canal discreto usando compresión de audio Dolby Digital o DTS similar al formato de sonido surround usado en los cines.

Al igual que en el video, la calidad de audio depende de lo bien que se haga el procesado y la codificación. Aun a pesar de la compresión, la calidad del Dolby Digital y DTS puede ser próxima o mejor que la del CD. La evaluación final de la calidad del DVD esta en la mano del consumidor. La mayoría de los espectadores lo valoran mejor que el laserdisc, pero nadie puede garantizar la calidad del DVD, del mismo modo que nadie puede descartarlo

basándose en rumores o demos. Al final es una cuestión de percepción del individuo y el nivel de calidad entregado por el sistema de reproducción.

4.4.3 ¿Cuales son las desventajas del DVD?

Pasaran años hasta que las películas, shows de televisión, otros tipos de video, y software de computador estén ampliamente disponibles. Indefiniciones en las especificaciones y pruebas inadecuadas de discos y lectores han dado lugar a incompatibilidades.

1. Algunas películas no funcionan totalmente (o para nada) en algunos lectores.
2. No se puede grabar (aún).
3. Tiene protección anticopia incorporada y bloqueo regional.
4. Usa compresión digital. El audio o video mal comprimido podría ser borroso, chillón, vago, sin trozos.
5. El proceso de audio compatible hacia atrás para estereo/Dolby Surround puede reducir el margen dinámico.
6. No soporta totalmente HDTV.
7. Algunos lectores DVD y drives podrían no ser capaces de leer CD-Rs.
8. Los actuales lectores DVD y drives no pueden leer discos DVD-RAM.
9. Muy pocos lectores pueden leer a velocidad normal hacia atrás.
10. Las variaciones y opciones tales como DVD-Audio, DVD-VR, pistas de audio DTS no son soportadas por todos los lectores.

4.4.4 ¿Que títulos están disponibles en DVD?

Al igual que con el hardware, las previsiones optimistas de cientos de títulos de películas para navidades de 1996 también fallaron. Únicamente un puñado de títulos DVD principalmente videos musicales estaban disponibles en Japón en noviembre de 1996 para el lanzamiento del DVD. Las películas actuales empezaron a salir en diciembre en Japón. En abril había sobre 150. Las primeras películas que aparecieron en USA el 19 de Marzo de 1997 por Lumivision, fueron Africa: The Serengeti, Antartica, The Rainforest, y Animation Greats.

El lanzamiento de la Warner Bros del 24 de marzo U.S. se limitó a 7 ciudades. Casi 19.000 discos se compraron en las primeras dos semanas del lanzamiento en USA, más de lo esperado. InfoTech predijo sobre 600 títulos a finales de 1997 y más de 8.000 para el 2000. En diciembre del 1997, sobre un millón de películas DVD habían ya sido vendidas.

A finales de 1999, se habían vendido unos 100 millones de discos. A finales del 2000 había sobre 10.000 títulos en USA y unos 15.000 en todo el mundo. Comparado con otros lanzamientos (CD, LD, etc) es un número inmenso en muy poco tiempo. La disponibilidad del hardware y software DVD en Europa llegó sobre un año a 18 meses por detrás de USA. Se anunciaron un número de lanzamientos con poco seguimiento, pero el DVD empezó a estar establecido alrededor de finales de 1998. Hay buenas bases de datos con posibilidad de búsqueda en DVD File, Express.com y DVD Planet. Para una lista de películas en formato panorámico en DVD visita WidescreenReview.

Concorde video lanzó una película en formato PAL "12 Monkeys" en Alemania al final de marzo. Fueron amenazados por Philips con un pleito por no incluir una pista multicanal MPEG, pero el asunto ya se resolvió. Muchos títulos inicialmente de DVD-ROM están disponibles solo como packs de hardware o software hasta que el mercado crezca más.

4.4.5 ¿Que son los "códigos regionales" "códigos de país"?

Los estudios cinematográficos quieren controlar los lanzamientos locales en diferentes países porque los lanzamientos en el cine no son simultáneos (una película podría salir en video en USA cuando todavía es un éxito en los cines en Europa). Incluso, los estudios venden los derechos de distribución a diferentes distribuidores extranjeros y les gustaría garantizar un mercado exclusivo. Por tanto ellos han requerido que el estándar DVD incluya códigos que se puedan usar para prevenir la reproducción de ciertos discos en ciertas zonas geográficas.

Cada lector tiene un código de la zona en la que es vendido. El lector se negará a leer discos que no son permitidos en esa región. Esto quiere decir que los discos comprados en un país podrían no leerse en lectores comprados en otro país. Alguna gente cree que los códigos regionales podrían ser considerados una restricción ilegal de comercio, pero no ha habido

casos legales que confirmen esto. Los códigos regionales son totalmente opcionales. Los discos sin código se leerán en cualquier lector en cualquier país. No es un sistema de cifrado, es un byte de información en el disco que el lector chequea.

Algunos estudios originalmente anunciaron que solamente sus nuevos lanzamientos tendrían códigos regionales, pero hasta ahora casi todos los lanzamientos se leen en una única región. Los códigos regionales son parte permanente del disco, no se desbloquearán después de un periodo de tiempo.

Hay 8 regiones (también llamadas "zonas"). Los lectores y los discos se identifican con el número de la región sobre impresionado en un plano del mundo. Si un disco se lee en más de una región tendrá más de un número en el mapa.

1. USA, Canadá, territorios USA
2. Japón, Europa, Sur África, Medio Este (incluyendo Egipto)
3. Sureste de Asia y Este de Asia (incluye Hong Kong)
4. Australia, Nueva Zelanda, Islas de Pacífico, América Central, Sudamérica, Caribe.
5. Europa del Este (Unión Soviética), subcontinente Indio, África, Corea del Norte, Mongolia
6. China
7. Reservado
8. Especial para usos internacionales (aviones, cruceros, etc...)

Técnicamente no hay nada que se llame disco de región 0 o lector de zona 0.

Hay discos que son multi-región. Hay lectores para todas las regiones. Es decir, algunos lectores, pueden ser modificados para leer discos independientemente de su código regional. Esto generalmente invalida la garantía, pero posiblemente no es ilegal. (La única cosa que obliga a los fabricantes de lectores para hacer lectores codificados por zonas es la licencia CSS). Algunos discos, como los de FOX, Buena Vista/Touchstone/Miramax, MGM/Universal, Polygram y Columbia TriStar incluyen código de programa que chequean

la región correcta (There's Something About Mary, Algopasa con Mary y Psycho son algunos ejemplos).

A finales del 2000, Warner Bros empezó a utilizar el mismo chequeo inteligente de región que otros estudios habían estado utilizando durante un año. Ellos le llaman "región code enhancement" (RCE, también conocido como REA), y recibió mucha publicidad. El RCE se añadió a algunos discos como The Patriot. Los discos RCE podrían chequear la configuración de país o idioma del lector además del código regional.

Un serio efecto de esto es que algunos lectores legales fallan la prueba. Estos "astutos discos" que tienen un chequeo activo de la región no se leerán en lectores "multizona" que están configurados para todas las zonas (FFh) pero pueden ser leídos en lectores de "código cambiable" que permiten que cambies la región usando el control remoto. Podrían incluso no funcionar en lectores con "conmutación-automática" que reconocen y se adaptan a la región del disco. (Depende de la configuración de región por defecto del lector. Un disco puede tener fijadas todas sus marcas de región de tal forma que el lector no sabe a cual configurarse, entonces el puede preguntar al lector la región y dar error si es la errónea. Una configuración por defecto de la región 1 confundirá a los discos astutos de la región 1). Un efecto secundario es que algunos discos RCE fallan en lectores legales, como el Fisher DVDS-1000.

Los códigos regionales también se aplican a sistemas DVD-ROM, pero usan solo con discos DVD-Video, no con discos DVD-ROM que contengan software de computador. Los sistemas de lectura de los computadores chequearan por los códigos de la región antes de leer películas de un DVD-Video. Los nuevos lectores "RPC2" DVD-ROM te permiten cambiar el código de la región varias veces (RPC significa Regional Protection Control). Una vez que el lector ha alcanzado el limite (generalmente 5 cambios) no puede ser cambiado de nuevo a menos que el vendedor o fabricante resetee el lector. La utilidad Drive Info puede decirte si tienes un lector RPC2 (dirá, este lector tiene protección regional). Drive Info e informacion para evitar las restricciones de las regiones DVD-ROM están disponibles en Internet en páginas como Visual Domain y DVD Infomatrix. Despues

de diciembre de 1999, solo se fabrican lectores RPC Phase II. Los códigos regionales no se aplican al DVD-Audio.

4.4.6 ¿Que hay acerca del DVD-Audio o Music DVD?

Cuando el DVD fue lanzado en 1996 no había formato de DVD-Audio, a pesar de que la capacidad del DVD-Video sobrepasa en mucho al CD. El DVD Forum buscó información adicional de la industria de la música antes de definir el formato de DVD-Audio. Un borrador del estándar fue lanzado por el DVD Forum's Working Group 4 (WG4) en enero 1998 y la versión 0.9 fue lanzada en Julio. La especificación formal del DVD-Audio 1.0 (menos la protección contra copia) fue aprobada en Febrero de 1999 y lanzada en Marzo, pero los productos fueron retrasados en parte por el lento proceso de seleccionar características de protección contra copia (cifrado y marcas de agua) con complicaciones introducidas por el Secure Digital Music Initiative (SDMI).

El lanzamiento de lectores programado en Octubre 1999 fue retrasado posteriormente hasta mediados del 2000, presumiblemente debido a preocupaciones causadas por el crack del CSS, pero incluso porque el hardware no estaba listo, las herramientas de producción no estaban depuradas y por el pobre soporte de las casas de música. Pioneer lanzo lectores de DVD-Audio en Japón a finales de 1999, pero no leerán discos protegidos contra copia. Matsushita espera tener lectores universales DVD-Video/DVD-Audio de las marcas Panasonic y Technics disponibles en julio del 2000 por 700U\$ a 1200U\$. Pioneer, JVC, Yamaha y otros podrían incluso lanzar lectores DVD-Audio al mismo tiempo. Sin embargo, parece no haber discos de DVD-Audio para ser leídos en los nuevos lectores. Entre medias, el estándar DVD-Video incluye sonido envolvente y calidad mejor que la del CD (mira en 3.6.2). El formato DVD-Audio es un formato separado del de DVD-Video.

Los discos de DVD-Audio pueden ser diseñados para trabajar en lectores de DVD-Video, pero es posible hacer un disco de DVD-Audio que no se lea en ningún lector de DVD-Video, puesto que la especificación de DVD-Audio incluye nuevos formatos y características, con contenido almacenado en una separada "Zona DVD-Audio" en el disco (el directorio AUDIO-TS) que los DVD-Video nunca miran. Se necesitan nuevos lectores de DVD-Audio, o nuevos "lectores universales" que puedan leer tanto DVD-Video como

discos DVD-Audio. Los lectores universales también se les llaman VCAPs (video-capable audio players).

Petición a los productores: Los lectores universales no estarán disponibles por algún tiempo, pero tú puedes hacer discos universales hoy en día. Con un pequeño esfuerzo, todos los discos DVD-Audio pueden hacerse para que trabajen en todos los lectores DVD incluyendo una versión en Dolby Digital del audio en la zona DVD-Video.

Petición a los desarrolladores de sistemas de edición de DVD-Audio: Haz que tu software haga eso por defecto o recomienda encarecidamente esta opción durante la edición.

El DVD-Audio (y los lectores universales) trabajaran con los receptores actuales. Ellos sacan salida PCM y Dolby Digital, y algunos soportarán los formatos adicionales de DTS y DSD. Sin embargo, la mayoría de los receptores actuales no pueden decodificar el audio multicanal PCM de alta-definición, e incluso si pudiesen, no pueden ser transportadas en los conectores estándares digitales de audio. Los lectores de DVD-Audio con (DACs) conversores digitales a analógicos high-end pueden únicamente ser unidos a receptores con entradas para dos canales o para seis, pero se perderá alguna calidad si el receptor convierte de nuevo a digital para procesado. Los futuros receptores con conexiones digitales mejoradas tales como IEEE 1394 (FireWire) serán necesarias para poder usar toda la resolución digital del DVD-Audio.

4.4.7 ¿Que estudios están apoyando el DVD?

Todos los grandes estudios de cine, la mayoría de los estudios de música. Cuando los lectores DVD estuvieron disponibles a principios de 1997, Warner y Polygram fueron los únicos grandes estudios en lanzar títulos. Títulos adicionales aparecieron de pequeños desarrolladores. Los otros estudios gradualmente se fueron uniendo al campo del DVD. Dreamworks fue el último estudio importante en anunciar total soporte a DVD. Paramount, Fox y Dreamworks inicialmente soportaban solo Divx, pero en verano de 1998 anunciaron que iban a soportar el DVD.

4.4.8 ¿Puedo grabar en DVD desde VCR, TV, etc?

Si. Cuando se introdujo el DVD en el año 1997 solo podía leer. Los grabadores de DVD aparecieron en Japón a finales de 1999, y en el resto del mundo a finales del 2000. Las primeras unidades fueron caras: 2500\$ a 4000\$. Los grabadores de DVD son todavía muy caros (entre los 500\$ a 2000\$), pero eventualmente podrían ser tan baratos como un video VHS. Los grabadores de DVD están ya siendo añadidos a receptores de satélite y cable, grabadores en disco duro, y cajas similares.

Un grabador de DVD es como un VCR, tiene un sintonizador y entradas AV, y puede ser programado para grabar programas. Una diferencia importante es que no tienes que rebobinar ni avanzar nunca más, las grabaciones en un disco son accesibles instantáneamente, generalmente desde un menú en pantalla. Ten en cuenta que los grabadores de DVD no pueden copiar la mayoría de los discos DVD de películas, pues están protegidos.

Desafortunadamente hay más de un formato de DVD grabable, y no todos se leen entre ellos perfectamente. No tiene nada que ver con la vieja batalla VHS con Betamax como muchos en la prensa podrían hacerte creer, pero es bastante confuso.

No te confundas también con los lectores grabadores de DVD para computadores. Estos grabadores pueden almacenar datos, pero crear DVD-Video con todas sus prestaciones, requiere software adicional para hacer la codificación (MPEG), la codificación de audio (Dolby Digital, MPEG o PCM), la navegación y la generación de datos de control.

4.4.9 ¿Es diferente la carpeta de la del CD?

Los fabricantes están preocupados por que los consumidores crean que los DVDs se pueden leer en sus lectores de CD, de tal modo que les gustaría que la carpeta sea diferente. Hay un numero de cajas de DVD que son tan anchas como una caja de CD y tan altas como una cinta de VHS como recomienda la VSDA (Video Software Dealers association). Sin embargo nadie esta siendo forzado a usar un tamaño de carpeta más grande. Algunas compañías usan el estuche normal o carpetas de papel o vinilo. Los discos Divx venían en

cajas de plástico y cartulina Q-pack del mismo tamaño que una caja de CD. La mayoría de las películas están empaquetadas con el Amaray "keep case," una caja de plástico de tipo concha (se abre como los libros) con bolsillos transparentes de vinilo para portadas, este es el más popular entre los consumidores. Hay incluso una "súper caja", versión expandida de una caja de CD que es común en Europa.

4.4.10 ¿Reemplazara el DVD al lector de cintas de video?

No lo hará pronto. Los DVD grabables lo son únicamente para datos de computador, no para video o televisión. Serán necesario un tiempo antes de que el tamaño del mercado de lectores desciendan al coste de los VCRs. Sin embargo, el DVD tiene muchas ventajas sobre los VCRs, incluyendo fundamentalmente disminuir los costos tecnológicos para la producción de hardware y de los discos (lo cual es atractivo para los fabricantes), de tal modo que si el DVD es un éxito comercial podría reemplazar muchos VCRs en 15 a 20 años.

4.4.11 ¿Reemplazara el DVD al CD-ROM?

Si, algunos fabricantes de lectores CD-ROM tienen pensado dejar la producción de lectores CD-ROM después de unos pocos años en favor de lectores DVD-ROM. Debido a que los lectores DVD-ROM pueden leer CD-ROMs hay una migración hacia adelante.

4.4.12 ¿Pueden los grabadores de CD-R crear DVDs?

NO, el DVD usa un láser de longitud de onda más pequeña para permitir unos pits más pequeños en pistas que están más próximas. El láser del DVD debe incluso enfocar más finamente y a diferente nivel.

En efecto, un disco grabado en los actuales grabadores de CD-R podría no ser legible por un lector DVD-ROM. No es probable que existan "actualizaciones" para convertir lectores CD-R a DVD-R puesto que esto posiblemente saldría más caro que comprar un nuevo lector DVD-R.

4.4.13 ¿Cuanto duran los discos DVD?

Los discos impresos para gran publico duraran mucho más que tu, desde 50 a 300 años. Los discos DVD-R se espera que duren entre 40 a 250 años, casi tanto como los discos CD-R. Los formatos grabables (DVD-RAM, DVD-RW, and DVD+RW) se espera que duren de 25 a 100.

4.4.14 ¿Que es un disco de doble capa? ¿Funcionara en todos los lectores?

Un disco de doble capa tiene dos capas de datos, una de ellas es semitransparente de tal modo que el laser puede enfocar a través de ella y leer la segunda capa. Puesto que ambas capas se leen desde el mismo lado, un disco de doble capa puede almacenar casi dos veces lo que un disco de una sola cara, típicamente unas 4 horas de video.

Muchos discos usan dos capas. En principio solo unas pocas plantas de duplicado podían hacer discos de doble capa pero hoy en día casi todas las plantas tienen esta capacidad. La segunda capa puede usar tanto un etiquetado "PTP" (parallel track path) donde ambas pistas vayan en paralelo (para datos independientes o efectos de conmutación especiales) o un etiquetado "OTP" (opposite track path) donde la segunda pista va en espiral opuesta, es decir, así la cabeza de lectura lee hacia afuera desde el centro en la primera pista y desde afuera hacia adentro en la segunda pista. El etiquetado OTP se hace así para proporcionar video continuo entre ambas capas. El cambio de capa puede ocurrir en cualquier momento en el video; no tiene que ser en un capítulo determinado. No se garantiza que la conmutación entre capas será imperceptible.

Los cambios de capa son invisibles en algunos lectores, pero otros lectores pueden causar que el video se congele por menos de medio segundo o hasta 4 segundos. El que no se note dependerá del modo en que el disco es hecho así como en el diseño del lector. El OTP también se llama RSDL (Reverse-Spiral Dual Layer).

La ventaja de las dos capas es que las películas largas pueden usar mayores velocidades de datos para mejor calidad que con una simple capa. Mira en 1.26 para ver detalles de cambio de capa. Hay varios modos de reconocer los discos de doble capa: 1) el color dorado, 2) un

menú en el disco para seleccionar la versión panorámica o la panorámica con barras, 3) dos números de serie en un lado.

La especificación DVD requiere que los lectores de salón y de computador puedan leer discos de doble capa. Hay muy pocas unidades que tienen problemas con discos de doble capa, eso es un error de diseño y debería ser corregido gratis por el fabricante. Algunos discos están diseñados con un "cambio de capa imperceptible" que técnicamente va por delante de lo que permite la especificación del DVD, esto causa problemas en unos pocos lectores viejos. Todos los lectores incluso leen discos doble cara si les cambias la cara manualmente. Ningún fabricante ha anunciado un modelo que lea ambas caras. Los costos añadidos no son justificables puesto que los discos pueden almacenar 4 horas de video en una cara usando dos capas (Los primeros discos usaron dos caras debido a que la producción de doble-capa no estaba ampliamente soportada, esto no es problema hoy día). Los lectores LD/DVD de Pioneer pueden leer las dos caras de un LD, pero no las de un DVD.

4.4.15 ¿Es un estándar mundial el DVD-Video? ¿Trabaja con NTSC, PAL y SECAM?

El MPEG video del DVD esta almacenado en formato digital, pero esta formateado para uno o dos mutuamente incompatibles sistemas de televisión: 525/60 (NTSC) y 626/50 (PAL/SECAM).

Hay tres diferencias entre discos creados para leerse en diferentes sistemas: tamaño de la imagen y formato de aspecto del píxel (720x480 sobre 720x576), velocidad de los cuadros (29.97 contra 25), y sonido envolvente (Dolby Digital contra MPEG). Por tanto hay dos clases de DVDs: DVDs NTSC y DVDs PAL/SECAM.

El video proveniente de las películas se codifica generalmente a 24 cuadros/seg pero se preformatea para una de las dos velocidades. Las películas formateadas para verse en PAL son generalmente aceleradas hasta un 4%, de tal modo que el audio debe ser ajustado de acuerdo a esto antes de ser codificado.

Todos los lectores DVD PAL pueden leer pistas de audio Dolby Digital, pero no todos los lectores NTSC pueden leer pistas de audio MPEG. PAL y SECAM comparten el mismo formato de escaneado, de tal modo que los discos son los mismos para ambos sistemas. La única diferencia es que los lectores SECAM sacan la señal de color en el formato requerido por las televisiones SECAM.

Algunos lectores solamente leerán discos NTSC, algunos lectores solo leerán discos PAL y otros leerán ambos. Todos los lectores DVD vendidos en países PAL leen los dos. Estos lectores multiestándar parcialmente convierten NTSC a una señal PAL de 60 Hz (4.43 NTSC). En este caso, el lector utiliza el formato de codificación de subportadora de color PAL 4.43 pero mantiene la velocidad de escaneo 525/60 NTSC.

La mayoría de televisores modernos PAL puede manejar esta clase de señal "pseudo-PAL" de 60Hz. Unos pocos lectores PAL multiestandar sacan NTSC 3.38 verdadero de un disco 525/60 NTSC, lo cual requiere una TV NTSC o una TV multiestándar. Algunos lectores tienen un conmutador para elegir la salida entre PAL 60 Hz y NTSC cuando leen discos NTSC. Hay unos pocos lectores standards-converting (convierte-estándar) que convierten desde discos NTSC 525/60 a salida estándar PAL.

Una conversión de estándares correcta requiere hardware caro para manejar escalamiento, conversión temporal y análisis de movimiento de objetos. Debido a que la calidad de la conversión es pobre en lectores DVD, usar una salida pseudo-PAL con una TV compatible proporciona mejor imagen. (El sonido no se ve afectado por la conversión de video). La mayoría de los lectores NTSC no puede leer discos PAL. Un muy pequeño número de lectores NTSC (como el Apex y SMC) pueden convertir PAL 625/50 a NTSC. Hay conversores externos, como el Emerson EVC1595 (U\$350). También hay conversores de alta calidad como el TenLab. Un productor puede elegir poner video 525/60 en un lado del disco y 625/50 en el otro. La mayoría de los estudios incluyen pistas de audio en Dolby Digital en sus discos PAL.

Hay actualmente tres tipos de lectores DVD si cuentas los computador es. La mayoría del hardware y software DVD pueden leer ambos video NTSC y PAL y audio Dolby Digital y MPEG. Algunos PCs pueden solamente mostrar el video convertido en el monitor del

computador, pero otros pueden dar salida de señal de video para una TV. Resumen final: Los discos NTSC (con audio Dolby Digital) se leen en cerca del 95% de las instalaciones DVD en todo el mundo. Los discos PAL se leen en muy pocos lectores fuera de los países PAL. (Esto es independiente de las regiones).

4.4.16 ¿Que es el firmware y porque debo actualizarlo?

Los lectores de DVD son simples computadores. Cada uno tiene un programa software que controla como reproduce los discos. Puesto que el software esta almacenado en un chip, se llama firmware. Algunos lectores tienen fallos en su programación que hace que tengan problemas leyendo ciertos DVDs. Para corregir estos fallos, el lector tiene que ser actualizado con un cambio del chip de firmware. Esto se hace generalmente en un centro de asistencia técnica, aunque algunos lectores pueden ser actualizados simplemente insertando un CD.

4.4.17 ¿Que pasa con las animaciones en el DVD? ¿Se comprimen mal?

Alguna gente dice que las animaciones, especialmente animaciones hechas a mano tales como dibujos animados, no se comprimen bien con MPEG2 o incluso acaban de mayor tamaño que el original. Otra gente dice que las animaciones se comprimen mejor. Ninguna de las dos afirmaciones está en lo cierto. Supuestamente el "jitter" (nerviosismo) entre cuadros causados por diferencias en los dibujos o en su alineamiento causa problemas. Un experto en animación en Disney cito que eso no sucedía con técnicas de animación modernas. E incluso si sucediese, la característica de estimación de movimiento del MPEG2 podría compensarla. Debido al modo en que el MPEG2 divide una imagen en bloques y los transforma en información frecuente, puede tener un problema con los bordes agudos que son comunes en las animaciones. Esta pérdida de información de alta frecuencia puede mostrarse como un "anillado" o puntos borrosos en los bordes (llamado efecto Gibbs) Sin embargo, a las velocidades de datos usadas generalmente para el DVD este problema no sucede.

4.4.18 ¿Por que algunos discos requieren cambio de cara? ¿No pueden los DVDs almacenar cuatro horas por cara?

Incluso sabiendo que la tecnología de doble capa permite sobre cuatro horas de reproducción continua en una única cara, algunas películas están partidas en los dos lados del disco requiriendo que se cambien de cara durante la reproducción. Esto es generalmente debido a que los productores fueron muy vagos para optimizar la compresión o para hacer un disco de doble capa. El tener una mejor calidad de imagen es una pobre excusa para incrementar la velocidad de transferencia; en muchos casos el video se ve mejor si se codifica bien a una velocidad menor. La carencia de capacidad de producción para doble capa es otra mala excusa, al principio muy pocas plantas podían hacer discos de doble capa, pero ya no es este el caso. Ningún lector puede cambiar automáticamente de cara, pero no es necesario puesto que la mayoría de las películas duran menos de 4 horas y pueden coger fácilmente en un lado de doble capa (RSDL).

4.4.19 ¿Porque esta la imagen apretada haciendo que las cosas parezcan muy flacas?

Respuesta: RTFM Estas viendo una imagen anamórfica creada para ser mostrada únicamente en una televisión panorámica. Necesitas ir al menú de configuración de tu lector y decirle que tienes una televisión 4:3 convencional y no una televisión panorámica 16:9. Entonces automáticamente le pondrá barras a la imagen de manera que puedas ver el ancho completo en las proporciones correctas. En algunos casos, puedes cambiar el formato mientras el disco se esta reproduciendo. En la mayoría de los lectores tienes que parar el disco antes de cambiar el formato. Algunos discos tienen etiquetas indicando panorámico en una cara y estándar en la otra. Para ver la versión panorámica debes dar la vuelta al disco.

4.4.20 ¿Usan todos los lectores de video el Dolby Digital (AC3)? ¿Todos tienen 5.1 canales?

La mayoría de los discos DVD-Video contienen pistas sonoras en Dolby Digital. Sin embargo, no se requiere. Algunos discos, especialmente aquellos que solo contienen audio,

tienen pistas PCM. Es incluso posible que un disco PAL 625/50 contenga audio MPEG, pero hasta ahora el audio MPEG no esta muy extendido.

No asumas que la etiqueta de "Dolby Digital" es una garantia de 5.1 canales. Una pista de sonido en Dolby Digital puede ser mono, dual mono, estéreo, Dolby Surround estéreo, etc. Por ejemplo *Blazing Saddles* y *Caddyshack* son películas en mono, de tal modo que la pista de sonido Dolby Digital en estos DVDs tiene un único canal. Algunos envases de DVD tienen pequeñas letras o iconos por debajo del logotipo de Dolby Surround que indican la configuración de canales. En algunos casos, hay mas de una pista digital: una pista en 5.1 canales y una pista especialmente mezclada para Dolby Surround. Es perfectamente normal para tu lector DVD indicar una pista de audio Dolby Digital mientras tu receptor indica Dolby Surround: quiere decir que el disco contiene una señal de dos canales Dolby Surround codificada en formato Dolby Digital.

4.4.21 ¿Cual es la diferencia entre el DVD de primera, segunda y tercera generación?

No hay una respuesta clarificadora a esta pregunta, ya que obtendrás una respuesta diferente de cada persona a la que preguntes. El término segunda generación y tercera generación se usa tanto para referirse a lectores de DVD-Video como a lectores DVD-ROM. En general, únicamente significan versiones más modernas de los sistema de reproducción DVD. Los términos no han sido usados (aún) para sistemas DVD que pueden grabar, leer video juegos y similares. De acuerdo con algunas personas, los lectores DVD de segunda generación salieron en otoño de 1997 y los lectores de tercera generación son aquellos que salieron al principio de 1998. Según otras personas, el DVD de segunda generación serán lectores de "alta definición" que no vendrán hasta el 2003 o así. Hay muchas confusas variaciones entre estos extremos, incluyendo los puntos de vista de los lectores compatibles DTS o lectores Divx o lectores de escaneo progresivo o lectores con 10 bits de video o lectores que pueden leer Matrix que constituyen la segunda, tercera o cuarta generación.

Las cosas están mas claras en el campo del PC, donde la segunda generación (DVD II) generalmente significa lectores DVD-ROM 2x que pueden leer CD-Rs, y tercera generación (DVD III) generalmente significa lectores DVD-ROM 5x (o a veces 2x o 4,8x

o 6x), unos pocos de los cuales pueden leer DVD-RAMs y algunos de los cuales son de formato RPC2. Alguna gente se refiere a los lectores RPC2 o 10X como cuarta generación.

4.4.22 ¿Que es lo que pasa entre el DTS y el DVD?

Digital Theater Systems Digital Surround es un formato de codificación de audio similar al Dolby Digital. Requiere un decodificador, o bien en el lector o bien en un receptor externo. Alguna gente dice que por su menor compresión el DTS suena mejor que el Dolby Digital. Otros dicen que no hay diferencias perceptibles significativas, especialmente a la velocidad de transferencia de 768 kbps que es el 60% más que el Dolby Digital.

Debido a las muchas diferencias en producción, mezcla, decodificación y niveles de referencia, es casi imposible comparar exactamente los dos formatos (DTS generalmente produce un nivel sonoro más elevado, lo cual hace que suene mejor). Los discos DTS eran codificados inicialmente por ellos, pero en Octubre del 1999 se pusieron a la venta codificadores DTS. Los títulos DTS son considerados generalmente como piezas especiales orientados a los entusiastas del audio. La mayoría de los discos en DTS están también disponibles en versión solo Dolby Digital. El DTS es un formato opcional en el DVD.

4.4.23 ¿Es compatible el CD-ROM con el DVD-ROM?

Si. Todos los lectores DVD-ROM leerán CD-ROMs (Yellow Book). El software en un CD-ROM ira bien en un sistema DVD-ROM. Sin embargo. los DVD-ROMs no son leíbles por los lectores de CD-ROM.

4.4.24 ¿Es compatible el CD-R con el DVD-ROM?

A veces, el problema es que los CD-Rs (Libro Naranja, parte II) son "invisibles" a la longitud de onda del láser DVD debido a que el color usado en los CD-Rs no refleja el haz. Algunos lectores de DVD-ROMs de primera generación y muchos lectores DVD-Video de primera generación no pueden leer discos CD-R's. La formula del tinte usado por diferentes fabricantes de CD-R también afecta a la capacidad de lectura. Es decir, algunas marcas de CD-R tienen mejor reflectividad a la longitud de onda del laser DVD, pero incluso esto no funciona con fiabilidad en todos los lectores. La solución común para el

lector de DVD o DVD-ROM es usar dos lasers a diferentes longitudes de onda: uno para leer DVDs y el otro para leer CDs y CD-Rs.

4.4.25 ¿Es compatible el video CD con el DVD?

A veces, no se requiere en las especificaciones del DVD pero es trivial soportar el Video CD (estandar del White Book) puesto que cualquier decodificador MPEG2 puede decodificar MPEG1 desde un video CD. Sobre dos tercios de los lectores DVD pueden leer Video CD's. Los modelos de Panasonic, RCA, Samsung y Sony leen video CD. Los modelos japoneses de Pioneer leen los videos CDs pero los americanos que sean más viejos que el DVL-909 no los leen.

Los lectores Toshiba más antiguos que los modelos 2100, 3107 y 3108 no leen los video CD's. La resolución del VCD es de 352x288 para PAL y 352x240 para NTSC. El modo en que la mayoría de los lectores de DVD y video CD tienen en cuenta la diferencia es recortando las líneas extra o añadiendo líneas en blanco. Cuando se leen VCDs PAL, los lectores Panasonic y RCA NTSC aparentemente cortan 48 líneas (17%) de la parte de abajo. Los lectores Sony NTSC ajustan todas las 288 líneas para que cojan. Debido a que los VCDs PAL están codificados para leer a 25cps una película de 24cps, hay generalmente una aceleración del 4%. El tiempo de lectura se acorta, y el audio es desplazado en velocidad a menos que haya sido procesado digitalmente antes de la codificación para poner el audio a su velocidad normal. Esto también pasa con DVDs PAL.

Todos los computadores con DVD-ROM pueden leer video CDs (con el software correcto). Los lectores estandar VCD no pueden leer DVDs.

4.4.26 ¿Es compatible el Super Video CD con el DVD?

Generalmente no. El Super Video CD (SVCD) es una mejora del Video CD que fue desarrollado por un comité de fabricantes e investigadores respaldados por el gobierno chino, en parte para decrementar un poco los royalties de la tecnología del DVD y en parte para crear presión para decrementar los precios de lectores y discos DVD en China.

En terminos de calidad de audio y video, SVCD esta entre Video CD y DVD, usando un lector CD 2x para soportar 2.2 Mbps VBR MPEG-2 video (a una resolucion de 480x480 (NSTC) o 480x576 (PAL)) y dos canales audio MPEG-2 Layer II. Al igual que el DVD, puede sobreimprimir graficos por subtítulos. Es técnicamente fácil hacer un lector DVD-Video compatible con SVCD, pero se esta haciendo principalmente en modelos de lectores DVD asiaticos.

4.4.27 ¿Es el Enhanced CD compatible con el DVD?

Si. Los lectores DVD leerán música de Enhanced Music CDs (Blue Book, CD Plus, CD Extra), y los lectores DVD-ROM leerán música y datos de Enhanced CDs. Los formatos viejos ECD tales como modo mezclado y pista cero (pregap, hidden track) deberían también ser compatibles, hay un problema con Microsoft y otros los lectores CD/DVD-ROM que saltan la pista cero.

4.4.28 ¿Es compatible el MP3 con el DVD?

Generalmente NO. El MP3 es compresión de audio MPEG-1 Layer 3 (MP3 no es MPEG-3, que no existe). La especificación DVD-Video permite Layer 2 únicamente para audio MPEG. El MP3 puede leerse en un computador con un lector DVD-ROM. Unos pocos lectores DVD-Video (Apex/REC/VDDV, DiViDo, Euro Asia/Smart, I-Jam, Lasonic, Nintaus, Raite [AV Phile/ Hoyo/ KiSS/ Monica/ Monyka/ Tokai/ Yamakawa], Revoy, Sampo, Shinco) pueden leer CDs MP3.

4.4.29 ¿Es compatible el laserdisc con el DVD?

No, los lectores DVD estándar no leerán laserdisc, y no puedes leer un disco DVD en cualquier lector laserdisc estándar. (El laserdisc usa video analógico, el DVD usa video digital, son formatos muy diferentes) Sin embargo, Pioneer produce lectores combinados que leen laserdisc y DVDs (y también CDVs y CDs de audio). Se rumorea que Denon y Samsung están trabajando también en lectores LD/DVD.

4.4.30 ¿Que efecto tendra el FMD en el DVD?

Quien sabe. Hasta ahora, el FMD (fluorescent, multilayer disc) de Constellation 3D's no ha salido del laboratorio. Puedes estar seguro que los informes causando la súbita muerte del DVD son exagerados salvajemente y no estan basados en la realidad. La tecnología de multicapas fluorescentes, que puede ser usada en tarjetas o discos, usa un laser en un material grabador fluorescente causando que emita luz. Puesto que no depende de luz laser reflejada, es posible crear muchas capas de datos (C3D ha hecho un prototipo con 50 capas en su laboratorio). Puede usar el mismo laser de 650 nm del DVD, de tal modo que los lectores FMD podrian hacerse para que leyesen DVDs. El 20 de junio C3D anunció un programa para hacer FMDs con 25 gigas por lado que podria ser leible por lectores DVD con unas "modificaciones pequeñas y baratas).

El FMD es una nueva tecnología, que no usa una pista grabada, soportada por una pequeña compañía. El DVD está basado en decadas de desarrollo de tecnología de almacenamiento óptico por docenas de compañías. Es posible que el FMD pueda llegar a establecerse en unos pocos años, pero el DVD esta tan atrincherado que los lectores FMD tendrán que leer DVD para que tengan exito. Por tanto no hay que preocuparse de que el DVD se quede obsoleto en algun momento. Quizas el FMD sea la tercera generación del formato DVD, siguiendo el desarrollo de la actual version de laser azul de alta densidad.

4.4.31 ¿Como afecta el MPEG4 al DVD?

El MPEG-4 en un formato estandar de codificacion de video diseñado principalmente para transmision de video a baja velocidad, aunque es actualmente mas eficiente que el MPEG-2 a velocidades de DVD y HDTV. El MPEG-4 incluso proporciona multimedia avanzado con sus media objects, pero la mayoría de las implementaciones solo soportan video sencillo (Simple Visual Profile). El DVD usa codificación de video MPEG-2. Los lectores de DVD estandar no reconocen el formato MPEG-4. Los ficheros MPEG-4 pueden ser almacenados en CDROM para su uso en computadores. Por ejemplo Divx;-) usa MPEG-4. Es posible que el MPEG-4 se use en una futura versión de DVD de alta definición. Es incluso posible que un formato similar tal como el H.263 sea usado para la siguiente generacion de DVD.

En cualquier caso, posiblemente no aparecerá antes del 2004 como pronto. Para más información sobre MPEG, mira la página de Tristan's MPEG.org y la página web de MPEG

4.4.32 ¿Que pasa con la grabación en DVD: DVD-R, DVD-RAM, DVD-RW y DVD+RW?

Hay cinco versiones de DVD-ROM grabable: DVD-R para authoring, DVD-R para general, DVD-RAM, DVD-RW y DVD+RW. Todos los lectores grabables pueden leer discos DVD-ROM, pero cada uno usa un tipo diferente para grabación. El DVD-R puede grabar datos solo una vez (solo en secuencial), mientras que DVD-RAM, DVD+RW y DVD-RW pueden ser reescritos miles de veces. El DVD-R fue el primero en estar disponible en otoño de 1997. El DVD-RAM le siguió en verano de 1998. El DVD-RW salió en Japón en diciembre de 1999. El DVD+RW estará disponible a principios del 2001. El DVD grabable estuvo disponible primeramente únicamente para computadoras. Los grabadores de DVD están siendo lanzados en todo el mundo desde finales del 2000.

Los grabadores de video DVD no pueden copiar discos con películas cifradas. Los tres formatos borrables (DVD-RAM, DVD-RW, y DVD+RW) están esencialmente en competición entre ellos. El mercado determinará cuál de ellos tendrá éxito. Por el momento, el DVD-RAM tiene una ventaja de casi dos años. El gran problema es que ninguno de los formatos escribibles es totalmente compatible con los demás o incluso con los lectores actuales. Conforme pasa el tiempo serán más compatibles y más mezclados entre ellos. Por ejemplo, los nuevos grabadores de Pioneer combinan DVD-RW y DVD-R. Los futuros lectores de DVD escribirán en discos DVD-R así como en CD-R/RW. Toshiba, Panasonic y otros lanzaron lectores combinados DVD-ROM/CD-RW a finales de 1999. Hay una buena lista de lectores DVD-R compatibles e incompatibles en HomeMovie.com

4.4.33 ¿Que es el DeCSS y el DivX;-)?

El CSS (Content Scrambling System) es un esquema de cifrado y autenticación pretendido para evitar que las películas DVD sean copiadas digitalmente. Mira 1.11 para más detalles. El DeCSS se refiere al proceso de eliminar CSS, así como al código y

programas DeCSS. El programa de computador para descifrar CSS fue lanzado en Internet en octubre de 1999.

La diferencia entre evitar el cifrado CSS con el DeCSS y interceptar video descifrado y descomprimido con un ripper DVD es que el DeCSS puede ser considerado ilegal bajo los tratados DMCA y WIPO. La información DeCSS puede ser usada para adivinar todas las claves master, de tal modo que un PC estandar puede generar la lista completa de 400 claves, haciendo el proceso secreto sin sentido. En cualquier caso, no hay ningun atractivo en ser capaces de copiar un conjunto de ficheros de películas (a menudo sin menus y otras características especiales del DVD) que podria llevar sobre una semana para descargar en un modem de 56K y podria llenar un disco de 6G o una docena de CD-Rs.

En marzo del 2000, aparecio una redistribución de tecnologia DVD llamada DivX;-) (Si, el smile/emoicon es parte del nombre, cuyos creadores deberian descuartizarlos por la tonta broma que ha confundido a miles de personas.). El DivX ;-)) es simplemente un hack del codec de video MPEG-4 de Microsoft y audio MP3, permitiendo que video sin DeCSS sea descargado y leído en el Windows Media Player.

Los creadores del DivX ;-)) se han aliado con el Project Mayo para desarrollar una nueva version llamada Divx Deux. Hay incluso una variacion de software-libre llamada 3ivx. A pesar de la baja velocidad de transferencia (y por tanto de la baja calidad), el tiempo y esfuerzo que lleva encontrar y descargar los ficheros no merece la pena para la mayoria de los aficionados a las películas. La realidad es que la gente que esta ripeando y descargando DVD lo estan haciendo por el desafio, no para evitar comprar discos.

Los partidarios del DeCSS dicen que fue creado unicamente para permitir que las películas DVD fuesen visualizadas en el sistema operativo Linux, el cual habia sido excluido de la licencia CSS por su naturaleza de codigo abierto. Esto esta especificamente permitido por las leyes DMCA y WIPO. [20]

4.5 CONVERTIDOR ANALOGICO-DIGITAL (ADC, Analogic to Digital Conversion)

¿Qué es ANALOGICO y que es DIGITAL?

El término **ANALÓGICO** en la industria de las telecomunicaciones y el cómputo significa todo aquel proceso entrada/salida cuyos valores son *continuos*. Algo continuo es todo aquello que puede tomar una infinidad de valores dentro de un cierto límite, superior e inferior.

El término **DIGITAL** de la misma manera involucra valores de entrada/salida *discretos*. Algo discreto es algo que puede tomar valores fijos. En el caso de las comunicaciones digitales y el cómputo, esos valores son el CERO (0) o el UNO (1) o Bits (BInary DigiTs).



Fig. 75. representación de una señal analógica y una digital.

Ventajas de la comunicación digital

La transmisión digital es la transmisión de pulsos digitales entre dos puntos, en un sistema de comunicación. La información de la fuente original puede estar ya sea en forma digital o en señales analógicas que deben convertirse en pulsos digitales, antes de su transmisión y convertidas nuevamente a la forma analógica en el lado del receptor.

Algunas de las **VENTAJAS** de la transmisión digital [con respecto a la analógica] son:

1.-La ventaja principal de la transmisión digital es la **inmunidad al ruido**. Las señales analógicas son más susceptibles que los pulsos digitales a la amplitud, frecuencia y variaciones de fase. Esto se debe a que con la transmisión digital, no se necesita evaluar esos parámetros, con tanta precisión, como en la transmisión analógica. En cambio, los pulsos recibidos se evalúan durante un intervalo de muestreo y se hace una sola determinación si el pulso está arriba (1) o abajo de un umbral específico (0).

2.-**Almacenamiento y procesamiento:** Las señales digitales se pueden guardarse y procesarse fácilmente que las señales analógicas.

3.- Los sistemas digitales utilizan la **regeneración de señales**, en vez de la amplificación, por lo tanto producen un sistema más resistente al ruido que su contraparte analógica.

4.- Las señales digitales **son más sencillos de medir y evaluar**. Por lo tanto es más fácil comparar el rendimiento de los sistemas digitales con diferentes capacidades de señalización e información, que con los sistemas analógicos comparables.

5.- Los sistemas digitales están mejor equipados para evaluar un rendimiento de error (por ejemplo, **detección y corrección de errores**), que los analógicos.

6.- Los equipos que procesan digitalmente consumen menos potencia y son más pequeños, y muchas veces con más económicos.

Algunas de las **DESVENTAJAS** de la transmisión digital son las siguientes:

1.- La transmisión de las señales analógicas codificadas de manera digital **requieren de más ancho de banda** para transmitir que la señal analógica.

2.- Las señales analógicas deben convertirse en códigos digitales, antes que su transmisión y convertirse nuevamente a analógicas en el receptor.

3.-La transmisión digital **requiere de sincronización precisa**, de tiempo, entre los relojes del transmisor y receptor.

4.- Los sistemas de transmisión digital son incompatibles con las instalaciones analógicas existentes.

La conversión Analógico-Digital consta de varios procesos:

- Muestreo
- Cuantización
- Codificación

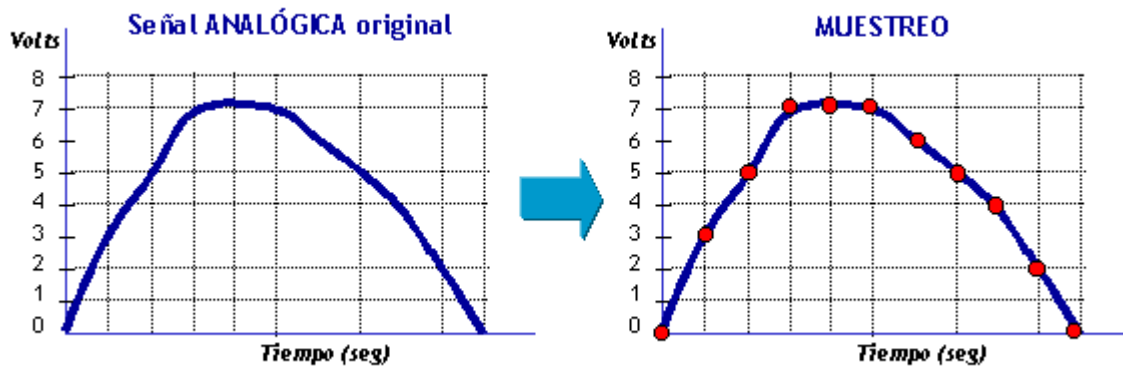


Fig. 76. muestreo de una señal analogica.

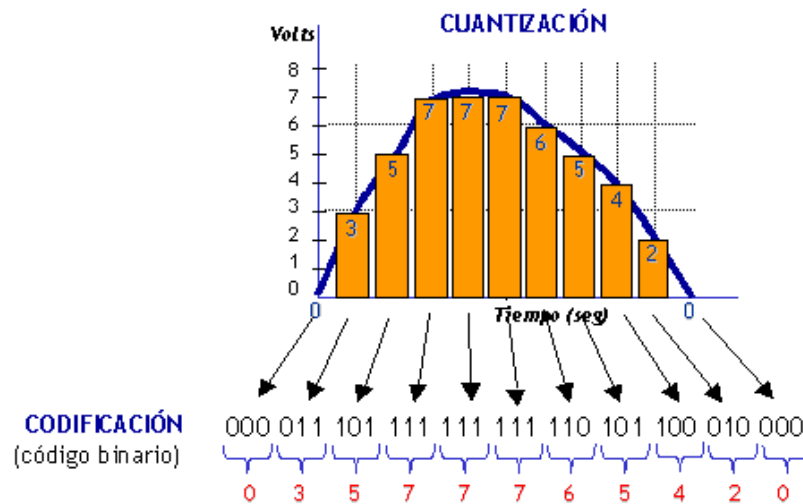


Fig. 77 .cuantificacion y codificacion de la señal.

4.5.1 Muestreo

Toda la tecnología digital (e.g. audio, video) está basado en la técnica de muestreo (sampling en inglés). En música, cuando una grabadora digital toma una muestra, básicamente toma una *fotografía fija* de la *forma de onda* y la convierte en bits, los cuales pueden ser almacenados y procesados. Comparado con la grabación analógica, la cual está basada en registros de voltaje como patrones de magnetización en las partículas de óxido de la cinta magnetica. El muestreo digital convierte el voltaje en números (0s y 1s) los cuales pueden ser fácilmente representados y vueltos nuevamente a su forma original.

Razón de muestreo

La frecuencia de muestreo de una señal en un segundo es conocida como razón de muestreo medida en Hertz (Hz).

$$1 \text{ Hz} = 1/\text{seg}$$

La razón de muestreo determina el rango de frecuencias [ANCHO DE BANDA] de un sistema. A mayores razones de muestreo, habrá más calidad o precisión.

Por ejemplo en audio digital se usan las siguientes razones de muestreo:

24,000 = 24 kHz - 24,000 muestras por segundo. Una muestra cada 1/24,000 de segundo.

30,000 = 30 kHz - 30,000 muestras por segundo. Una muestra cada 1/30,000 de segundo.

44,100 = 44.1 kHz - 44,100 muestras por segundo. Una muestra cada 1/44,000 de segundo.

48,000 = 48 kHz - 48,000 muestras por segundo. Una muestra cada 1/48,000 de segundo.

Una señal de audio muestreada a 48 KHz tiene una mejor calidad [el doble], que una señal muestreada a 24 KHz. Pero, una señal muestreada a 48 KHz, ocuparía el doble del ancho de banda que la de 24 KHz. Por lo que si queremos mayor calidad, lo perdemos en ancho

de banda. Cuando bajan archivos en Internet MP3 por ejemplo, éstos tienen diferentes calidades, un archivo MP3 de mejor calidad, ocupará mayor espacio en disco...

La calidad de un disco compacto [CD] equivale un muestreo de 44.1 KHz a 16 bits, éste es el estándar. Si decimos que los archivos MP3 tienen calidad de CD, es que están muestreados a 44.1 KHz a 16 bits.



Fig. 78. calidad de sonido y bits.

4.5.2 Cuantización:

Es el proceso de convertir valores continuos [e.g voltajes] en series de valores discretos.

Por ejemplo el audio digital [e.g. MP3, WAV, etc] tienen dos características bien importantes, el muestreo (tiempo) y la cuantización (nivel).

Mientras que el muestreo representa el tiempo de captura de una señal, la cuantización es el componente amplitud del muestreo. En otras palabras, mientras que el muestreo mide el tiempo (por instancia 44,100 muestras por segundo), la cuantización es la técnica donde un evento analógico es medido dado un valor numérico.

Para hacer esto, la amplitud de la señal de audio es representada en una serie de pasos discretos. Cada paso está dado entonces por un número en código binario que digitalmente codifica el nivel de la señal. La longitud de la palabra determina la calidad de la representación. Una vez más, una palabra más larga, mejor la calidad de un sistema de audio (comparando una palabra de 8 bits con una de 16 bits o 32 bits) (ver figura).

El bit de resolución de un sistema define el rango dinámico del sistema. 6 dB es ganado por cada bit.

Por ejemplo:

8 bits equivale a **256 estados** = 48 dB (decibeles) **16 bits** equivalen a **65,536 estados** = 96 dB.

Entonces, se debe de tomar muestras a tiempos menores y se debe de cuantizar a mayores niveles (bits), si sucede lo contrario suceden **errores de cuantización**.



Fig. 79. error de cuantizacion.

4.5.3 Codificación

La codificación es la representación numérica de la cuantización utilizando códigos ya establecidos y estándares. el código más utilizado es el código binario, pero también existen otros tipos de códigos que son empleados.

A continuación se presenta una tabla donde se representan los números del 0 al 7 con su respectivo código binario. Como se ve, con 3 bits, podemos representar ocho estados o niveles de cuantización. [21].

$2^{(n)}$ = Niveles o estados de cuantización

donde n es el número de bits.

Num.	Codigo binario	Num.	Codigo binario	Num.	Codigo binario
0	000	3	011	6	110
1	001	4	100	7	111
2	010	5	101		

CONCLUSIONES.

Con el paso del tiempo el video ha evolucionado, quien iba a imaginar que algún día íbamos a poder capturar imágenes y audio, con la calidad con la que podemos tener ahora. Y mucho menos imaginamos que los dispositivos de captura y almacenamiento fueran demasiado pequeños como lo son ahora. En la actualidad podemos almacenar gran cantidad de video en un disco compacto y con una calidad de alta definición.

La evolución en el video ha dado pie a la evolución de los distintos aparatos utilizados para su reproducción. La TV de alta definición, el reproductor de DVD, son algunos ejemplos de esta evolución.

Dado que en estos tiempos ya es digitalizado el video, esto contribuyo al desarrollo de programas que pudieran comprimir la señal digital de video para que pudiera ser enviada, prueba de esto es la creación de estándares de compresión denominados MPEG. El video digital implica grandes ventajas. Como son la calidad y el uso de menos ancho de banda, su codificación.

El video tiene estrecha relación con el audio. Y el incluir audio a las imágenes también implica que el audio sea comprimido para su envío. Este envío se realiza con el MPEG audio. Debido a que el audio también ya ha sido digitalizado y su calidad aumentó al grado de que podemos escuchar desde una canción hasta un video musical con calidad estereofónica o superior y esta puede ser almacenada en un disco compacto al igual que el video.

Hace algunos años podíamos pensar en la comunicación mediante video digital en tiempo real, pero sólo imaginarlo. Hoy en día puede decirse que es una realidad. Pero no está todo logrado. Tenemos la posibilidad de solicitar una película de cine desde casa, y que se nos sea servida inmediatamente vía satélite. Pero no sólo queremos eso, queremos poder disfrutar de la comunicación multimedia y a bajo precio. Ahora mismo esto significa que a través de Internet ya que es la red que nos llega a todos hasta la puerta de nuestras casas.

Finalmente se cumplió con el objetivo de realizar una recopilación bibliográfica con los conceptos más importantes sobre el tema de “VIDEO”, esperando que se le de el uso esperado y contribuya a servir como una fuente de consulta para los alumnos del Instituto de Ciencias Básicas e Ingeniería en especial para los alumnos de la carrera de Ing. Electrónica y Telecomunicaciones.

GLOSARIO

.1 (punto 1): se refiere al canal de efectos de baja frecuencia en formatos de sonido surround de 5.1 y 6.1 canales.

2-pass. (o Multi-pass): En pocas palabras, el "multipass" es el mejor formato de codificación de audio y video conocido hasta ahora. Utilizar esta técnica al codificar un video a otro formato, significa básicamente que el codificador de video (o *video encoder*) analiza el video múltiples veces, desde el comienzo hasta el final, antes del proceso de codificación propiamente tal. Mientras escanea el archivo, el *encoder* escribe información acerca del video original en su propio *logfile* y utiliza esta información para determinar la mejor manera posible de calzar el video dentro de los límites de *bitrate* que el usuario ha configurado para el proceso de codificación. Esto explica porqué la codificación multipass solo es utilizada en codificaciones con VBR (*Variable Bitrate*) (*la codificación vía CBR - Bitrates Constantes- no ofrece ninguna flexibilidad al encoder para determinar el bitrate de cada frame*). La mejor manera de entender porqué se utiliza este método es pensar en una película -- cuando existen tomas que son total y absolutamente negras (cambios de escena por ejemplo) la codificación de 1 pasada (1-pass) CBR utiliza exactamente la misma cantidad de datos para dicha parte, que la que utilizaría para una escena de acción compleja. Utilizando VBR y Multi-pass, el encoder "sabe" qué escenas permanecen Ok con bajos bitrates, y los bitrates que pueden ser usados en escenas más complejas, logrando así una mejor calidad para escenas que requieren un mayor bitrate. SuperVCD y DivX4 son 2 populares formatos que utilizan este método de codificación.

5.1 canales: se refiere a formatos de sonido surround que utilizan 5 canales (designación de "5") de sonidos de alta frecuencia, y un canal (designación de "1") para efectos de baja frecuencia (LFE). Ejemplos de formatos de sonido surround de 5.1 canales son Dolby Digital y DTS.

525/60: Sistema de escaneo de 525 líneas por cuadro (o *frame*), y 60 campos (o *fields*) entrelazados (30 cuadros) por segundo. Es utilizado por la televisión estándar NTSC.

6.1 canales : se refiere al formato de sonido Extended Surround, como el THX Surround EX y el DTS-ES 6.1 Matrix, que tienen un canal de sonido surround posterior, para obtener efectos de 360 grados. Para ser justos, DTS-ES Discrete 6.1 es el único formato de sonido surround de 6.1 canales.

7.1 canales: se refiere a algoritmos cuyo sonido posterior está basado en dos parlantes. (Back surround).

16:9 widescreen/aspect ratio: se refiere a pantallas de televisión con aspecto de 16:9 o DVDs publicados con aspecto de 16:9.

A/D Conversión de Analógico a Digital (Analogue to Digital Conversion). También denominado digitalización o cuantificación. Consiste en la conversión de señales analógicas a digitales, normalmente para su utilización posterior en un equipo digital.

AES/EBU La Sociedad de Ingeniería de Audio (Audio Engineering Society) (AES) junto con la EBU (European Broadcast Union) han definido una norma para audio digital, ahora también adoptado por el ANSI (American National Standards Institute) (Instituto Nacional de Normas Americano). Conocido como 'AES/EBU', esta norma de audio digital permite diversidad de frecuencias de muestreo, por ejemplo, CDs a 44.1 Khz, o VTRs digitales a 48 Khz - 48 Khz es la frecuencia estándar más extendida en el mundo de la postproducción.

AFL After Fader Listening. Cuando en una mesa de mezclas de audio, no hay seleccionado ningún pulsador de preescucha, el sistema se configura de forma que se monitorice la señal de la salida de programa, pudiéndose comprobar el nivel de cada fuente sobre la mezcla total de sonido.

AM Modulación en amplitud.

ANSI (American National Standards Institute). Instituto de estandarización de EE.UU., que ha creado diversos estándares, entre los que podemos citar ASCII.

ASK Amplitude Shift Key. Cuando se utilizan moduladoras digitales y éstas modulan señal senoidal en amplitud.

Audio embebido Audio que se transmite incorporado a una señal de vídeo - se suele referir a su uso en SDI, donde es posible incluir hasta cuatro canales de audio digital junto con el vídeo.

AC-3: Diminutivo de **Audio Encoding 3**, un algoritmo de Dolby Laboratories. Es el nombre original de lo que conocemos por Dolby Digital.

Agudos (Treble) Es la parte alta del rango de frecuencias de audio, aproximadamente es arriba de 10,000Hz.

AIFF: (.aif, .aifc, .aiff) (Audio Interchange File Format) formato de archivos de audio desarrollado por Apple Computer. Se utiliza para almacenar sampleos de alta calidad de audio.

Aliasing: Es una distorsión (artifact) durante la reproducción de video o audio digital, que aparece cuando la frecuencia de la señal es más del doble que la frecuencia de muestreo. La resolución es insuficiente, ya que admite ruido adicional que no está presente en la onda original.

Análogo (Analog) Una señal analógica es una en la cual el voltaje variable es análogo (similar) a la forma de onda acústica, que es continuamente variable.

Artefacto: (Artifact). Efecto no-natural, que no está presente en el video o audio original, y que es producido por una acción o agente externo.

Anamorphic Widescreen: Término referido a Video DVD codificado en un formato widescreen de 16:9 (*aspect ratio*). El video DVD puede ser acomodado a aspect ratios de 4:3 y 16:9

ASF: Advanced Streaming Format: Este formato de archivos almacena información de audio y video, y fué especialmente diseñado para trabajar en redes, como Internet. La información es descargada como un flujo continuo de datos, y por ende, no es necesario esperar la descarga completa del archivo para poder reproducirlo.

Aspect Ratio: El "ratio" es la razón entre el ancho y alto de la pantalla, o "ancho:alto". De esta manera, "4:3" significa que el tamaño horizontal es un tercio mayor que el tamaño vertical. El *aspect ratio* de la televisión estandard es de 4:3 (o 1,33:1).

AVI: Audio Video Interleave: formato de video digital desarrollado por Microsoft. Tanto el video como el audio son comprimidos en forma conjunta, y define cómo están unidos, con el fin de preservar la sincronía de ambos. Un archivo AVI puede utilizar diferentes codecs y formatos.

Baja Frecuencia (Low Frequency) Cualquier frecuencia más baja que 160Hz.

BB Black Burst. Señal que se utiliza para sincronizar elementos que componen un estudio de producción. Esta señal contiene 7 informaciones: frecuencia y fase de líneas, frecuencia y fase de cuadro, frecuencia y fase de subportadora de color y nivel de pedestal negro.

BETACAM Sistema VTR en componentes analógico que utiliza una cinta de ½ pulgada - muy similar al Betamax doméstico. Fue desarrollado por Sony y lo comercializan ellos mismos y otros muchos fabricantes.

Aunque graba en cinta las señales componentes Y, R-Y y B-Y, muchas máquinas funcionan con vídeo codificado (PAL o NTSC) de entrada y de salida. El sistema ha seguido desarrollándose a lo largo de los años para ofrecer modelos para los mercados profesionales así como ancho de banda de luminancia total (Betacam SP), audio PCM y entrada y salida digital serie para el mercado broadcast.

BLU Banda Lateral Única. Su objetivo es mejorar el rendimiento de emisor pero encareciendo el receptor. Se suprime la portadora y una banda lateral.

BLV Banda Lateral Vestigial. En este sistema se suprime parcialmente una banda lateral, de un 5 a un 10% de la original. Permite aumentar el número de canales por cada banda y mejora el rendimiento con respecto a la modulación normal de AM, pero manteniendo la simplicidad en la modulación. Este sistema es el utilizado en la tradicional TV analógica terrestre.

BMP Mapa de bits. Formato de imagen de uso muy extendido en los equipos compatibles con Windows.

BNC El origen de las siglas BNC no está claro, y se le han atribuido muchos nombres, desde «British Naval Connector» a «Bayonet Neill-Councilman». Haremos referencia a esta familia hardware simplemente como BNC, debido a que no hay consenso en el nombre apropiado y a que en la industria de la tecnología las referencias se hacen simplemente como conectores del tipo BNC. Se trata de un conector de tipo coaxial y de bayoneta.

Bitrate: Término bastante utilizado al hablar de calidades de video y audio. Define cuánto (o el promedio) de espacio físico (en *bits*) toma un segundo de audio o video. Por ejemplo, 3 minutos de audio MP3 a 128kBit/sg CBR, toma 2,81 MegaBytes de espacio físico ($1,024 \times 128 / 1,024 / 1,024 \times 180 / 8$).

en donde:

128kbit = 128 x 1024 bits

1 byte = 8 bits

1 kilobyte = 1,024 bytes

1 megabyte = 1,024 kilobytes

3 minutos = 180 segundos.

Algunos ejemplos:

MP3 cerca de 128 kbps (kilobits por segundo)

VCD cerca de 1374 kbps

DVD cerca de 4500 kbps

DV cerca de 25 Mbps (megabits por segundo).

Capacidad de almacenamiento Según la norma de codificación digital ITU-R 601 4:2:2, cada imagen ocupa una gran cantidad de espacio de almacenamiento - sobre todo en lo que se refiere a dispositivos de almacenamiento en ordenadores tales como DRAM y discos. Tanto es así, que las cifras pueden resultar confusas a menos que se recuerden unos cuantos datos de referencia. Afortunadamente, las unidades de mega, giga, y tera hacen que sea fácil expresar las grandísimas cifras implicadas. Las capacidades se pueden obtener directamente a partir de la norma 601.

Para la norma de 625 líneas la imagen activa es: $720(Y) + 360 (Cr) + 360 (Cb) = 1440$ pixels/línea

Con 576 líneas activas/imagen hay $1440 \times 576 = 829.440$ pixels/imagen (es decir, muestreando a 8 bits, una imagen ocupa 830 kbytes) 1 sg. Requiere $830 \times 25 = 20.750$ kbytes, o 21 Mbytes.

Para la norma de 525 líneas la imagen activa es: $720(Y) + 360 (Cr) + 360 (Cb) = 1.440$ pixels/línea.

Con 487 líneas activas por imagen hay $1440 \times 487 = 701.280$ pixels/imagen (es decir, muestreando a 8 bits, una imagen ocupa 701,3 kbytes) 1 sg. Ocupa $701,3 \times 30 = 21.039$ kbytes, o 21 Mbytes.

De este modo, ambos sistemas de 625 y 525 líneas requieren aproximadamente la misma cantidad de almacenamiento para un tiempo determinado.

1 min. Requiere $21 \times 60 = 1.260$ Mbytes, o 1,26 Gbytes

1 hora requiere $1,26 \times 60 = 76$ Gbyte n=ms*

*76 Gbytes será suficiente para ambas normas

Algunos números útiles (referidos a vídeo no comprimido):

1 Gbyte almacena 47 sg.

1 hora ocupa 76 Gbytes

CCD Dispositivo de acoplamiento de carga (Charge Couple Device) (CCD) - constituido por una matriz lineal o por una bidimensional de elementos sensibles a la luz. La luz se convierte en una carga eléctrica proporcional a la luz que incide en cada célula. Las células están acopladas a un sistema de barrido que, después de una conversión de analógico a digital, presenta la imagen como una serie de dígitos binarios. Las primeras matrices CCD eran incapaces de reproducir un rango amplio de luminancia pero ahora ofrecen imágenes con bajo ruido y alta resolución. Mejorando la resolución actual de la televisión, se producen ahora CCDs bidimensionales libres de fallos para su utilización en HDTV.

CCIR Comité Consultor Internacional de Radiocomunicaciones (Comité Consultatif International des Radiocommunications). Ha sido absorbido por el ITU bajo la siglas ITU-R.

CD-ROM Compact Disc - Read Only Memory. Disco compacto que sólo lee memoria. Una modalidad de disco compacto utilizada para almacenar datos: hasta 650 megabytes de imágenes fijas, sonido y gráficos. Su uso está muy difundido por todo el mundo, reemplazó a los diskettes como soporte de almacenamiento y su popularidad solo esta amenazada por las mayores prestaciones del DVD.

Compact Disc Recorder. Grabador de CD-ROM. Aparato que permite grabar un CD-ROM virgen, generalmente en una sola sesión

CIRC Cross Interleave Redundant Code: Código Redundante Entrelazado y Cruzado. Este código lo utilizan los sistemas DAT para recuperar fragmentos de información de audio que se ha perdido o dañado.

CT Código de tiempo. Señal grabada en los sistemas profesionales que definen de una forma más precisa que la señal CTL, la ubicación de las imágenes en la cinta. Estos códigos están formados por señales digitales asociadas a cada imagen, especificando horas, minutos, segundos y cuadros de ese segundo.

CTL Señal de control utilizada en los magnetoscopios. Esta señal se graba en una pista longitudinal y no indica una posición absoluta. Los controladores de edición determinan la posición de la cinta contando estos impulsos.

CBR: Constant Bitrate (Bitrate Constante): Básicamente es un término que describe cómo es codificado el audio y el video, en donde el bitrate no varía a lo largo del clip de audio o video. (*Ver VBR: variable Bitrate*). El VCD standard contiene MPG video y audio de bitrate constante, al igual que el MP3.

CD (Compact disc) Formato de grabación digital inventado por Philips y Sony, cuantificado a 16 bits y 44.1kHz de frecuencia de muestreo.

Closed Caption: señal de video de texto sobrepuesto, que no es visible (en contraposición al open caption, que es parte permanente de la imagen). Normalmente, es la representación de audio hablado. En los EEUU, el standard oficial NTSC de Closed Caption requiere que todo televisor mayor de 13 pulgadas incluya circuitos para decodificar cualquier tipo de información sobrepuesta almacenada en la línea 21 del video digital. El DVD puede proveer de datos closed caption, pero se prefiere el formato de imagen de bitmap sobrepuesta por su versatilidad.

Codec: Acrónimo de "codificación/decodificación", un codec es un algoritmo o programa de computación especializado que codifica o reduce el número de bytes consumidos por archivos y programas grandes. Los archivos codificados con un codec específico requieren el mismo codec para ser decodificados. Algunos codecs conocidos son Divx, MPEG-1, MPEG-2, Xvid, DV type 1 y type 2 para video, y MP3 para audio.

Component Video:(video por componentes) Un sistema de video que contiene tres señales de componente separadas por color; rojo, verde, azul (*red, green blue* RGB) o diferencias de chroma/color (YCbCr, YPbPr, YUV) en forma análoga o digital. El sistema de codificación MPEG-2 utilizado por DVDs, está basado en componente de video digital de diferencias de color. Solo algunos televisores tienen entradas de video por componentes.

D1 Formato de grabación digital en cinta de vídeo según la norma ITU-R 601, 4:2:2 que utiliza una cinta de 19 mm. De ancho, permitiendo grabar hasta 94 minutos en una cassette. Al tratarse de un sistema de grabación por componentes es ideal para trabajos de estudio o de postproducción ya que su gran ancho de banda de crominancia permite realizar llaves de croma de excelente calidad. Además, es posible producir generaciones múltiples con muy poca degradación y los equipos D1 se pueden integrar sin necesidad de transcodificación con la mayoría de los sistemas de efectos digitales, telecines, dispositivos gráficos, grabadores de disco, etc. Al ser por componentes, no hay exigencias de 'color framing'. A pesar de las ventajas, los equipos D1 no se utilizan demasiado para producción de TV, en parte debido a su alto coste.

D2 Modelo de VTR para señales compuestas (codificadas) digitales PAL o NTSC. Utiliza cinta de 19 mm. y graba hasta 208 minutos en una única cassette. Ni las cassettes ni el formato de grabación son compatibles con D1.

El D2 se ha utilizado con frecuencia como un claro sustituto de los VTRs de 1 pulgada. Aunque ofrece buenas posibilidades de funciones especiales y generaciones múltiples sin pérdidas, el tratarse de un sistema codificado implica que las características derivadas de la codificación están presentes. El usuario debe ser consciente del 'cross color', huellas de transcodificación, poco ancho de banda de crominancia y secuencias de 'color framing'. Si se emplea un formato de 8 bits para muestrear la totalidad de la señal codificada se produce una reducción de amplitud en la resolución, haciendo que el D2 sea más susceptible a la aparición de artificios de 'contorneado'.

D3 Modelo de VTR que utiliza cintas de 1/2 pulgada para grabar señales compuestas (codificadas) PAL o NTSC digitalizadas y muestreadas a 8 bits. Se dispone de cassettes desde 50 hasta 245 minutos. Puesto que utiliza una señal compuesta las características son en general las mismas que para el D2 excepto que el tamaño de la cassette de 1/2 pulgada ha permitido la creación de una gama completa de equipos VTR con el mismo formato, incluyendo camascopios.

D5 Formato de VTR que utiliza la misma cassette que el D3 pero graba señales en componentes muestreadas según las recomendaciones ITU-R 601 con 10 bits de resolución.

Mediante una decodificación interna los VTRs D5 pueden reproducir cintas D3 y proporcionar salidas en componentes. Al tratarse de un grabador de vídeo digital por componentes sin compresión, el D5 ofrece las mismas prestaciones que el D1, por lo que resulta adecuado para postproducción y también para su uso general en estudios. Además de su utilidad para los sistemas actuales de TV de 625 y 525 líneas, este formato también permite la grabación HDTV mediante una compresión de alrededor de 5:1.

DAT Digital A Tape. Sistema de almacenamiento digital de audio sobre una cinta de 3,81 mm de ancho, dos canales de sonido digital muestreado a 48 KHz y cuantificado a 16 bits.

dB DECIBELIO (dB) Unidad de medida que expresa relaciones utilizando escalas logarítmicas y que se emplea para expresar magnitudes vinculadas a la percepción humana auditiva o visual. Se pueden asociar muchos atributos diferentes al punto de referencia denominado 0 dB - por ejemplo un nivel estándar de sonido o potencia - y obtener medidas relativas a esa referencia. Muchos niveles de funcionamiento se expresan en dB - por ejemplo la relación señal/ruido (S/N). Las relaciones en dB se definen según la expresión: $20 \log_{10} (\text{Nivel 1}/\text{Nivel 2})$ donde los niveles 1 y 2 pueden ser audio, vídeo o cualquier otro nivel de voltaje apropiado.

DBL Doble Banda Lateral. Sistema de modulación en amplitud donde se suprime la portadora y se transmiten únicamente las bandas laterales, ahorrando energía en la emisión.

DCT 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.

DIN 45500 Das ist Norm. (Esto es norma). Sistema de normalización especialmente concebido para conseguir la unificación de tolerancias, tamaños, calidades, etc., de los elementos empleados y producidos en la industria. Las normas DIN fueron publicadas por vez primera en 1926 en Alemania y han sido adoptadas por gran número de países. Concretamente la norma 45500 se refiere a especificaciones técnicas de equipos de sonido HI-FI.

DNR Dolby Noise Redution. Sistema de reducción de ruido utilizado en las grabaciones sobre cinta magnética. Según su complejidad se diseñaron variantes: Dolby A, Dolby B y Dolby C.

Drop-frame time code El formato de 525/60 líneas/campo utilizado en el sistema NTSC no funciona exactamente a 60 campos por segundo sino a 59.94, o 29.97 cuadros por segundo. El código de tiempos identifica 30 cuadros por segundo. El código de tiempos con salto de cuadro compensa este error perdiendo dos cuadros por minuto excepto en el décimo. Obsérvese que el sistema PAL 625/50 es exacto y no requiere pérdida de cuadros. Ver también: Código de tiempos sin salto de cuadro.

DVB Direct Vídeo Broadcasting. Sistema de difusión de televisión digital basado en el sistema MPEG-2.

DVB-C Sistema DVB diseñado para transmitir por cable que utiliza la modulación QAM.

DVB-MC Utilizando la misma tecnología que la distribución por cable pero con una frecuencia de hasta 10 GHz, para aplicaciones de televisión directa en microondas.

DVB-MS Representa la versión de microondas con los principios básicos del sistema de transmisión DVB-S

DVB-S Sistema DVB diseñado para transmitir por satélite canales de 36 MHz de ancho de banda, con modulación QPSK. Incluye canales de pago. Equivale a la actual transmisión satélite digital.

DVB-T Es la versión digital para transmisiones por tierra, es decir aprovecha las mismas antenas que actualmente funcionan en todos los hogares. Puede usar tanto los sistemas de modulación QPSK , QAM y OFDM

Deinterlace: El proceso de crear un frame único a partir de 2 fields entrelazados de un frame de video. Esta técnica es utilizada principalmente para remover artefactos entrelazados si se requiere un nuevo frame o si el video está siendo utilizado a un rate diferente al que fue creado.

Digital TV: Es el nuevo estandar norteamericano para la televisión digital. Consiste en no menos de 18 formatos distintos de imagen, incluyendo los formatos de definición standard (SDTV) y de alta definición (HDTV).

Demultiplexar: Dividir el audio y el video en archivos separados. También se conoce como "Demux".

Discreto: se refiere a un canal de audio o señal que es independiente de cualquiera de los otros canales. Dolby Digital 5.1 y DTS 5.1 son ejemplos de sistemas de sonido de multi-canales discretos.

Divx: Digital Video Express. Una variación del DVD basado en el sistema pay-per-view, de corta vida. No confundir con DivX;).

DivX;): sistema de codificación (codec) desarrollado por un grupo de hackers, en particular por un tipo llamado gej. Está basado en la tecnología MPEG-4 de Microsoft, conocida como Windows Media Video V3. Basicamente, los encoders de Microsoft no permiten al usuario salvar streams MPEG-4 a la estructura del formato AVI, en vez de ello fuerzan a los usuarios a utilizar ASF. Esto trajo algunas limitaciones, que fueron "corregidas" por DivX;), que también agrega soporte para otras tecnologías de codificación de Windows Media Audio, permitiendo a los usuarios utilizar MP3s en sus películas. El año 2001, los desarrolladores originales de este codec "ilegal" realizaron una versión legal del DivX;), llamada DivX4. DivX4, sin el emoticón ;) , soporta películas con el antiguo formato DivX;) y agrega nuevas capacidades y eventualmente una compresión de mayor calidad. El nombre, DivX;) proviene del desaparecido sistema de DVD por Pay Per View llamado DIVX. Con DivX;) es posible almacenar entre 50 y 120 minutos de *relativa* buena calidad de video en un CD (740MB) . La mayoría de las películas en formato DivX;) pueden almacenarse en un CD (a diferencia de un VCD o un SVCD). El único aspecto negativo del DivX;) es que no puede ser reproducido en un DVD Player, a diferencia de los VCDs.

Dolby Digital, Dolby Surround: sistema de codificación de audio desarrollado por Dolby Laboratories y aceptado como un estandar internacional. Es la forma más común de

decodificar audio para el DVD Video y es el sistema de codificación obligado para discos con sistema NTSC (525/60).

Dolby ProLogic: Técnica (o circuito que aplica la técnica) de extraer audio de sonido surround de una señal de audio de codificación matricial. El Dolby ProLogic es tan solo una técnica de decodificación, pero usualmente es , erróneamente, utilizado para referirse al sistema de codificación Dolby Surround.

Dual Layered: (dos capas) se refiere a discos de DVD con dos capas en un mismo lado. Los reproductores de DVD pueden leer datos de video u audio en ambas capas con tan solo reenfoque del laser. UN DVD de una cara y de doble capa puede almacenar hasta 4 horas de video y audio y es conocido como DVD-9. Un DVD de doble cara y doble capa puede almacenar hasta 8 horas de video y audio (DVD-18).

DTS: Digital Theater Sound: Sistema de codificación de audio desarrollado para salas de cine. Un competidor de Dolby Digital y un formato de audio opcional para el DVD-Video y el DVD-Audio.

DVD: son las siglas correctas de Disco Versátil Digital (originalmente Digital Video Disc). Los DVDs pueden contener una o dos caras, cada una con una o dos capas. Dependiendo de su construcción, pueden tener diversas capacidades. Existen varios formatos de DVD, como el DVD-Video para películas y otros títulos de video; el DVD-Audio que posee múltiples canales de audio digital, y el DVD-ROM para almacenamiento de datos en el computador. Los reproductores de DVD que reproducen tanto DVD-Video como DVD-Audio son conocidos como DVD-Universal.

EDTV: Enhanced Definition Television, un subconjunto del nuevo formato de televisión digital (Digital TV). El formato EDTV define 480 líneas de resolución en escaneo progresivo (o 480p).

EHF Extreme High Frequency. Extra Alta Frecuencia, desde 30 hasta 300 GHz.

Encode: (*codificar*) Método que permite transformar los datos para su almacenamiento o transmisión, eliminando redundancias y reduciendo su complejidad. La mayoría de los métodos de compresión se basan en uno o más métodos de codificación.

FBAS Farb Bild Austast Sincronsignal (alemán). Señal de color, imagen, borrado y sincronismo. En resumen, estas siglas denominan a la señal de vídeo compuesto.

FCC Federal Communications Commission. Comisión Federal de Comunicaciones. Es una agencia independiente del gobierno de los Estados Unidos, directamente ligado al Congreso. El FCC se creó en 1934 y ha ido modificándose para regular el desarrollo de la tecnología de las comunicaciones: radio, televisión, redes, satélite y cable.

FIT Frame Interline Transfer. Sensor de transferencia de cuadro interlineal. Se trata de un sensor CCD que recoge las ventajas de los modelos anteriores FT e IT. En este sensor los problemas de 'lag' y de 'smear' son mínimos.

FM Modulación en frecuencia

FPS Cuadros por segundo (Frames per second).

FSK Frecuence Shift Key. Cuando se utilizan moduladoras digitales y éstas modulan señal senoidal en frecuencia.

FT Frame Transfer. Sensor de transferencia de cuadro. En estos primeros sensores las áreas de captación y de almacenamiento se dispusieron separadas. Este sistema adolece de varios problemas entre ellos el 'lag' y el 'smear'.

Frecuencia: El espectro de sonido es caracterizado por su rango de frecuencias. Esta puede medirse en ciclos por segundo o hertz. El oído humano puede oír entre los 20Hz (bajos) a 20.000Hz. A medida que envejecemos, el rango de frecuencias se vuelve más estrecho. La siguiente tabla ilustra los rangos de frecuencia en una banda sonora:

Tambores - menos de 30 Hz

Guitarra (Bajo) - 30 Hz a 200 Hz

Voz Humana - 100 Hz a 1,700 Hz

Piano - 25 Hz a 3,600 Hz

Címbalos - sobre los 15,000 Hz

Los equipos de audio y video tienen frecuencias que van de los 20 a los 20 mil hertz, con mejor respuesta a las anchas frecuencias. Frecuencias menores a los 20 Hz se traducen normalmente en vibraciones.

Full motion video: video que se ejecuta a treinta cuadros por segundo (29,97 para ser exacto) en NTSC o 25 frames por segundo (en PAL).

HF High Frequency. Alta Frecuencia, desde 3 hasta 30 MHz.

HDTV: Hight Definition Television o Televisión de Alta Definición, parte de la nueva especificación de la ATSC de Televisión Digital (Digital TV). HDTV se refiere normalmente a los formatos 1080i o 720p. El formati 1080i consiste en 1080 líneas de resolución de escaneo interlazado, mientras que el formato 720p se refiere a 720 líneas de resolución en escaneo progresivo. El estandar integro que define a la televisión digital posee no menos de 18 formatos de imagen.

Hi-Fi VHS: se refiere al formato VHS estandard con Dolby Surround Pro-Logic codificado en canales de audio stereo.

Horizontal lines of resolution: líneas de resolución horizontales: el número de líneas de escaneo horizontal que define a una imagen de video. La mayoría de los reproductores de DVD poseen una salida de 480 líneas de resolución horizontal. En comparación, el VHS standard ofrece 260 líneas, y la televisión análoga estandard aproximadamente 330 líneas de resolución horizontal. Un mayor número de líneas horizontales generalmente significa mayor nitidez de la imagen.

Todo el mundo se confunde con el término "líneas de resolución horizontal", también llamadas LoHR o TVL, puesto que es una medida subjetiva. Siendo un sobrante del video analógico, es medido y anunciado inconsistentemente por los fabricantes, pero hemos de aguantar con esto hasta que el video sea digital y la resolución se mida en píxeles.

IEEE 1394. Interfase mediante la cual se pueden conectar a un bus de transferencia de datos varios dispositivos simultáneamente, con una alta velocidad de transferencia.

Fue desarrollada inicialmente por Apple (firewire) y Texas Instruments, pasando a ser más tarde un estándar aceptado por el IEEE (Institute of Electrical and Electronics Engineers). Sony lo llama "i.link".

Su primera versión, IEEE-1394, proporciona velocidades de acceso de hasta 400 Mbps (megabits por segundo), llegando en la actualidad a velocidades muy superiores. Sus principales características son:

- Velocidad de transferencia de 3'2 GB por segundo.
- Hasta 63 dispositivos en la misma conexión, permite la conexión en caliente al ordenador.
- Permite la captura directa de imágenes desde cámaras digitales que tengan este interfaz al ordenador, sin necesidad de convertir las imágenes y sin perder calidad.

El bus de alta velocidad IEEE 1394 complementa al USB proporcionando una conectividad de PC mejorada para un amplio rango de dispositivos: dispositivos de electrónica de consumo de audio / video (A/V), periféricos de almacenamiento, otros ordenadores y dispositivos portátiles.

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 (anteriormente CCIR) se ocupa de los temas de gestión y regulación del espectro de radio mientras que la ITU-T (anteriormente CCITT) se ocupa de la normalización de las telecomunicaciones.

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 y se deriva del SMPTE RP125 y del EBU Tech. 3246-E. 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 - ver ITU-R 656.

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 para Y (luminancia), Cr y Cb (señales diferencias de color digitalizadas) = $224 = 16.777.216$ 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, que deriva de las normas SMPTE RP125 y EBU Tech 3246-E. 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.

Interlace: es un sistema de escaneo de video basado en la transmisión de líneas alternadas, de modo que con cada barrido de pantalla se muestra la mitad de la imagen. Un frame "entrelazado" está generado, por ende, por dos campos (o *fields*).

Interleave: es un sistema utilizado para ordenar datos en trozos alternados, de modo de poder seleccionar las partes a ser extraídas y visualizadas y seleccionar aquellas que se van a "saltar". En DVD, es utilizado para implementar cámaras multi-ángulo o el material de selección alternativa para una misma escena.

Interpolate : *interpolación*; sistema que permite crear píxeles adicionales al escanear líneas o imágenes, promediando píxeles, líneas o frames adyacentes. Generalmente, esto causa un efecto de suavizado en imágenes detenidas y un efecto de blur en imágenes en movimiento.

Inverse Telecine: *Telecine Reverso* (IVTC); proceso de edición de video que consiste, básicamente, en volver a adecuar el framerate original de una película NTSC de 29,97 a 24 cuadros por segundo.

JPEG Grupo de Expertos Fotográficos Unidos (Joint Photographic Experts Group), ISO/ITU-T. JPEG es una norma para la compresión de datos de imágenes fijas (intra-campo). Concretamente, su trabajo tiene que ver con imágenes codificadas de acuerdo con la norma ITU-R 601. JPEG utiliza DCT, ofrece compresión de datos con una relación entre dos y cien veces y se definen tres niveles de procesamiento: codificación básica, extendida y "sin pérdidas".

En general, es de esperar que la compresión introduzca algún tipo de pérdida o degradación en la imagen, dependiendo su grado del algoritmo utilizado así como de la relación de compresión.

LF Low Frequency. Baja Frecuencia, desde 30 hasta 300 KHz.

LNB Low Noise Booster. Amplificador de bajo ruido. También denominado LNC, Conversor de bajo ruido. Es el elemento de la antena parabólica encargado de transformar las señales de alta frecuencia (SFH), de 10 a 12 GHz, a una banda de frecuencias más baja, denominada primera frecuencia intermedia, que se extiende desde 1'75 hasta 2'055 GHz, a la vez que se le aplica una amplificación.

LNE Lineal No Edition. Edición No Lineal. Sistema de edición de vídeo utilizando plataformas informáticas, donde el acceso a la información de vídeo puede ser aleatorio y rápido.

LTC Código de tiempo longitudinal (Longitudinal Timecode). Código de tiempo grabado en una pista lineal en una cinta y leído por una cabeza estática. Se puede leer fácilmente cuando la cinta se mueve hacia delante o hacia atrás pero no en un cuadro parado, en el que se puede utilizar el código de tiempo grabado con la imagen (VITC).

LaserDisc: es un formato de video análogo basado en discos grandes (8 u 12 pulgadas). Las últimas versiones de LaserDiscs soportan audio codificado con sonido Dolby Digital o DTS surround. El formato DVD es superior al LaserDisc, puesto que es digital, posee video codificado en MPG-2 de mayor calidad y mayor espacio de almacenamiento. El LaserDisc a sido desplazado por la popularidad del DVD. Muy pocos reproductores de DVD ofrecen compatibilidad con LaserDisc.

Low Frequency Effects (LFE): se refiere a los efectos de baja frecuencia (bajos profundos), representados por el ".1" de los 5.1 o 6.1 canales de sonido surround. En palabras sencillas, hace la explosiones más explosivas y dan a las pisadas de dinosaurios en *Jurassic Park* un efecto más impactante. Para lograr reales efectos de baja frecuencia, se recomienda usualmente la utilización de subwoofers. Sin embargo, puedes obtener resultados similares si posees parlantes que puedan reproducir frecuencias entre los 20 y 30 Hz, y tu receptor puede dividir la señal LFE a los parlantes principales.

MPEG: Moving Pictures Expert Group. Comité internacional que desarrolló la familia de sistemas de compresión de audio y video MPEG.

MF Médium Frecuency. Media Frecuencia, desde 300 KHz hasta 3 MHz

MJPEG Compresión de imágenes utilizando la compresión intracuadro al igual que JPEG. Se puede utilizar directamente en la edición no lineal.

MPEG-1: algoritmo de compresión de video, que es parte del standard del VideoCD. Este algoritmo fue desarrollado por el Motion Pictures Experts Group (MPEG). MPEG-1 efectivamente, comprime una imagen de video a cerca del 1/40 de su tamaño original.

MPEG-2: algoritmo de compresión de video, que es parte del standard del DVD-Video, Digital Broadcast Satellite y de Digital TV (incluyendo HDTV). Este algoritmo fue desarrollado por el Motion Pictures Experts Group (MPEG). MPEG-2 comprime una imagen de video a cerca del 1/40 de su tamaño original, pero su calidad es superior a la lograda por MPEG-1.

MP3: MPEG-1 Layer III audio. Algoritmo perceptual de codificación de audio. No es soportado por los formatos DVD-Video o DVD-Audio.

MPEG audio: audio comprimido de acuerdo al sistema de codificación perceptual MPEG. El audio MPEG-1 provee dos canales, que pueden ser formato Dolby Surround. El MPEG-2 audio agrega datos para proveer audio de multicanales discretos. El MPEG stereo es el sistema de compresión de audio obligado para el sistema de DVD-Video 625/50 (PAL/SECAM).

MPEG video: compresión de video de acuerdo al sistema de codificación MPEG. Típicamente, el MPEG-1 es utilizado para bajas tasas de transmisión de datos (*data rate*) como en el Video CD. El MPEG-2 es utilizado para obtener mejor calidad en el video, especialmente para video de tipo entrelazado (interlaced), como el DVD o la HDTV.

Multiplexar: combinar múltiples señales o cadenas de datos en una única señal o cadena. Usualmente se utiliza para entrelazar (interleaving) a bajo nivel.

NLE. EDICION NO LINEAL. Quiere decir que el medio de grabación utilizado no es lineal - no es cinta. Se utiliza con frecuencia para describir un entorno de grabación en que hay acceso rápido (directo) a los "clips" fuente y al espacio de grabación - normalmente utilizando discos. Esto elimina el rebobinado y los pre-rolls de las operaciones lineales (VTR), acelerando así el trabajo, pero no implica la mucha mayor flexibilidad del acceso aleatorio en tiempo real a cualquier cuadro (acceso aleatorio real). El término se ha

utilizado mucho asociado a sistemas de edición off-line que almacenan imágenes muy comprimidas, pero cada día aumenta el número de sistemas no lineales on-line. Hay una gran cantidad de sistemas que aseguran ofrecer calidad on-line con compresión de vídeo. El presunto usuario debe juzgar si los resultados son los adecuados para su aplicación específica. Los sistemas no comprimidos, que no comprometen la calidad de la imagen, se utilizan cada vez más.

Non-drop (frame time code) Código de tiempos en que no se utiliza la eliminación de cuadros y siempre identifica 30 cuadros por segundo. De esta manera, la velocidad del código de tiempos no será exactamente igual a la del tiempo real (29.97 cuadros por segundo).

El desajuste es de 18 cuadros más cada 10 minutos. Esto se aplica solo a los sistemas 525/60 ya que los sistemas 625/50 tienen un número exacto de cuadros por segundo.

NTSC: Comité de Normas de Televisión Nacional (National Television Standard Committee) Es un grupo de asesores de un comité establecido en los años 40 que recomendó el sistema de transmisión actual de 525 líneas, 60 campos. Este sistema es usado en América del Norte y Sur mientras que en Europa y Asia se utiliza el sistema PAL. Cabe mencionar que estos 2 sistemas no son compatibles.

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 (Fase alternada en cada línea) 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.

PCM: "Modulación por Código de Pulsos" (Pulse Code Modulation) Un método para representar una señal de audio como una serie de muestras digitales. Este modo proporciona grabación digital de alta fidelidad sin ninguna señal de compresión.

Progressive scan: el escaneo progresivo es la habilidad para generar una imagen en una pasada (como lo hace el monitor del computador). El escaneo entrelazado convencional requiere dos pasadas para generar una imagen, la primera para las líneas impares. La televisión analógica convencional que conocemos, posee escaneo entrelazado (*interlaced*) con una imagen completa que se refresca 30 veces por segundo.

QPSK Quadrature Phase Shift Key. Es una versión muy extendida de la modulación en fase, que utiliza cuatro fases posibles para la portadora. Las cuatro coinciden con los ejes de coordenadas. La transmisión se efectúa tomando los bits de información por parejas, y asignándoles las 4 posibles fases: 00 = 0°, 01 = 90°, 10 = -90°, 11 = 180°.

RGB Abreviatura de las señales rojo (Red), verde (Green) y azul (Blue), los colores primarios en televisión. Las cámaras y los telecines tienen receptores rojo, verde y azul, las pantallas de TV tienen fósforos rojo, verde y azul iluminados por cañones rojos, verdes y azules. Gran parte de la monitorización en un centro de producción se realiza en RGB.

RCA: El conector más popular en el mercado, creado por RCA Victor.

Reverberación: (Reverberation) Es el lapso de tiempo en el que un sonido es audible por reflexiones dentro de un lugar cerrado. Es una serie de ecos espaciados pero suficientemente cercanos en tiempo que se mezclan con una señal de audio.

S/N o SNR RELACION SEÑAL/RUIDO. (Signal to noise ratio). La relación entre el ruido y la información útil de la imagen (señal) se suele expresar en dB. Los equipos digitales son capaces en teoría de generar imágenes puras libres de ruido, que tendrían una relación señal/ruido infinita. Pero éstas, debido precisamente a su pureza, pueden causar artificios de "contorneado" si se procesan sin emplear técnicas especiales - una razón para utilizar el Redondeo Dinámico (Dynamic Rounding).

Una regla general para expresar la relación señal/ruido real de un sistema digital se define con la expresión: S/N (Relación señal/ruido en dB) = $6N + 6$ donde N es el número de bits. Por lo tanto un sistema de 8 bits dará 54 dB S/N. Este sería el nivel de ruido de un 'dither' LSB continuo y sólo se produciría sobre la imagen completa al digitalizar un campo

uniforme (es decir un gris uniforme sobre toda la pantalla) ajustado a un nivel situado en medio de dos LSBs. Con otros métodos de prueba se obtienen diferentes resultados, produciendo la mayoría cifras S/N más elevadas.

SECAM SECuencial A Memoria. Sistema de TV instaurado en Francia, Europa Oriental, el Magreb y algunos países de Oriente Próximo. Se envían las señales R-Y y B-Y de forma alternada, por lo que es necesaria una memoria para la síntesis de la señal de color. Estas señales se modulan en frecuencia La señal Y se envía en todas las líneas consiguiendo así la retrocompatibilidad.

SHF Super High Frequency. Súper alta frecuencia, desde 3 hasta 30 GHz.

Surround Sound: sonido surround, generalmente se refiere a canales de audio multicanal y múltiples parlantes, que permiten recrear un sonido tridimensional para reproducir la banda sonora de cine y música.

S-Video: (Y/C Video): es un tipo de señal de video utilizado en formatos de videocintas S-VHS y Hi-8. Transmite la luminancia y el color como componentes separados, evitando la combinación de ambos y la consecuente pérdida de calidad al componer un video.

S-VHS: "Super VHS" Es una mejora del formato VHS que produce una mejor resolución de imagen al separar la chroma, luminancia y los pulsos de sincronía vertical y horizontal por separado. La resolución de imagen llega hasta 400 líneas cuando en un VHS normal es aproximadamente 240.

TRC Cathode Ray Tube. Tubo de rayos catódicos. la mayoría de las pantallas de TV y PC usan este logro técnico, un bombardeo de electrones en un tubo vacío que permite ver las imágenes en el monitor. Actualmente está en clara sustitución por las pantallas de plasma o TFT.

Television Standards: Actualmente existen tres estándares para señales de televisión analógica, tanto para broadcast y aplicaciones de videotape. Los tres formatos son: PAL, NTSC y SECAM, que tienen significantes diferencias. Todas comparten una característica en común: cada frame o imagen completa se compone de dos fields o barrido de líneas

alternadas. PAL se mueve a 25 frames por segundo (fps) y 625 líneas de escaneo; utiliza amplitud modulada (AM) en un ancho de canal de 8Mhz. El NTSC se mueve a 29.97 fps. con 525 líneas de escaneo y utiliza amplitud modulada (AM) en un ancho de canal de 6Mhz. SECAM utiliza el mismo frame rate y número de líneas que PAL pero utiliza frecuencia modulada (FM) sujeta a una banda de 8Mhz. En algunos países de SudAmérica, existen sistemas híbridos (*PAL-N*, *PAL-M*), que combinan características de NTCS y PAL. NTSC es el estándar en la mayor parte de América Central, mientras que PAL predomina en el oeste de Europa. Algunos ejemplos:

NTSC: Estados Unidos, Canada, México, Japon, Korea del Sur, Filipinas, Chile.

PAL: Inglaterra, Alemania, España, India, China, Australia.

SECAM: Francia, Hungría, Polonia, Rusia, Arabia Saudita, Egipto.

THX:"Tomlinson Holman Experiment" Conjunto de patentes, tecnologías y criterios desarrollados por Lucas Film para la reproducción de películas, basada en el control del ambiente acústico y visual y en el sistema Dolby Surround. No es un decodificador, no necesita software THX (de hecho no existe) es un reproceso a la información decodificada, el dolby digital, pro-logic (MPEG en su caso) y DTS se benefician de este reproceso que logra dar más realismo al home theater creando un sonido trasero difuso, sin localización (multicanal) incrementa el rango dinámico, crea diálogos claros, el movimiento de sonido es natural y la respuesta de frecuencias es plana, el circuito THX se compone de Cinema Re-EQ, Timbre Match y decorrelación dinámica.

UHF Ultra High Frequency. Ultra Alta Frecuencia, desde 300 MHz hasta 3 GHz.

VCD Los CD de vídeo contienen secuencias de vídeo MPEG-1. El estándar para la creación de CD de vídeo fue establecido por Philips y JVC en el White Book en 1993. La primera pista contiene el programa de reproducción CD-i para el CD de vídeo, así como los directorios CDI, MPEGAV y VCD basados en el sistema de archivos ISO Level 1, ISO Level 2, Joliet. Las pistas siguientes tienen el formato de sectores CD-ROM/XA modo 2/formato 2 y contienen las secuencias de audio/vídeo codificadas MPEG que están incluidas en MPEGAV.

Los CD de vídeo se pueden reproducir en reproductores CD-i, reproductores de CD de vídeo especiales o bien en un PC en unidades de CD-ROM con soporte de CD-ROM/XA y decodificador MPEG (estándar en Windows 98 y versiones superiores) o con el software de CD de vídeo propio del usuario.

El CD de vídeo admite los formatos siguientes:

- 352x288 con 25 imágenes por segundo (PAL)
- 352x288 con 29,97 imágenes por segundo (NTSC)
- 704x576 como imagen única
- 704x480 como imagen única
- 352x288 con 29,976 imágenes por segundo (resolución "FILM")

Variable Bit Rate: (VBR): Es una técnica de compresión de datos que produce streams de datos variables entre un rango máximo y un mínimo a través del tiempo. En palabras sencillas, esta técnica permite ir variando la calidad mientras la tasa de transmisión se mantiene constante.

VGA (Video Graphic Adapter). Placa gráfica cuya resolución permite ver imágenes de calidad en la pantalla de la computadora. Válida para la mayoría de las creaciones multimedia

VHF Very High Frequency. Muy Alta Frecuencia, desde 30 a 300 MHz.

VHS Video Home Service. Formato de vídeo para uso doméstico cuya calidad no es suficiente para el terreno profesional. Desarrollada por JVC. Competió con sus homónimos BETA de Sony y V2000 de Philips, imponiéndose definitivamente.

Virtual surround: Cualquier algoritmo que simula un sistema de sonido surround con sólo los dos parlantes frontales.

VLf Very Low Frequency. Muy baja frecuencia, desde 3 hasta 30 KHz.

Widescreen: se refiere a aspect ratios mayores que 1,33:1. Por ejemplo, 1,78:1; 1,85:1 y 2,35:1.

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.

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: Método que define a una señal de video que separa los componentes de luminancia (Y) y crominancia (UV). La visión humana es mucho más sensitiva a las variaciones de intensidad que a las variaciones de color. El proceso de codificación YUV toma ventajas de este fenómeno y provee un ancho de banda mayor para la información de luminancia que para la de crominancia. Individualmente, las letras YUV significan Intensity, Hue, y value. (Intensidad, matíz y valor).

BIBLIOGRAFIA

[1] Fredy Fabián Cuello Rojas, Juan Carlos Rueda Erazo. 2 FUNDAMENTOS DEL VIDEO ANÁLOGO Y DIGITAL
http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/video_analogo.htm

[2] Formatos de vídeo, FORMATOS DE VÍDEO Y CONECTORES
<http://www.telepolis.com/cgi-bin/web/DISTRITODOCVIEW?url=/mundopc/doc/audio/video.htm>

[3] R.G. Bosco, Sistemas de televisión color
<http://www.monografias.com/trabajos5/siste/siste.shtml>

[4] Forma del Pixel
<http://www.video-computer.com/Pixels.htm>

[5] Juan Manuel López Sánchez, televisión
<http://dfists.ua.es/es/asignaturas/tv/>

[6] Fredy Fabián Cuello Rojas, Juan Carlos Rueda Erazo. 9 COMPRESIÓN DE VIDEO EN EL ESTÁNDAR MPEG-1 (APLICACIONES MULTIMEDIA)
<http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/MPEG1.htm>

[7] Fredy Fabián Cuello Rojas, Juan Carlos Rueda Erazo. 10 COMPRESIÓN DE VIDEO EN EL ESTÁNDAR MPEG-2 (APLICACIONES BROADCAST)
<http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/MPEG2.htm>

[8] Ramón Sendra, MPEG-4: el estándar de video para multimedia

<http://www.dvdenlared.com/cineencasa/20040824100258.html>

[9] Manuel Hernán Villa-Roel G. Estándar MPEG-7 y sus Aplicaciones

<http://dis.eafit.edu.co/cursos/st780/material/bdigital/articulos/2003/MPEG7%20y%20sus%20aplicaciones.pdf>

[10] Pablo Barrera González, MPEG-21: Distribución de contenidos en un mercado digital

<http://gsync.info/~barrera/documentos/articulos/mpeg21.pdf>

[11] 3 MPEG Audio

<http://www.hispamp3.com/tallermp3/tutoriales/mp3profundidad/3.shtml>

[12] Yolanda Lucas Baldó, Nuria Higuera Ruiz, Marta Solaz Hernández, Montse Celorrio Jiménez, Sistema de compresión MPEG

<http://www.upv.es/satelite/trabajos/pracGrupo21/mpeg.htm>

[13] Fredy Fabián Cuello Rojas, Juan Carlos Rueda Erazo. 7 COMPRESIÓN DE AUDIO

http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/compresion_audio.htm

[14] HDTV Básico

http://www.hdtv.video-computer.com/0_HDTV_Basico.htm

[15] USO DE LA ALTA DEFINICION

http://www.hdtv.video-computer.com/1_Usos_HDTV.htm

[16] ALTA DEFINICION. LOS ESTANDARES.

http://www.hdtv.video-computer.com/2_Estandars_HDTV.htm

[17] AUDIO PARA ALTA DEFINICIÓN

http://www.hdtv.video-computer.com/3_Audio_HDTV.htm

[18] Balanceo Inteligente de vides

<http://www.commspecial.com/deuceguidespa.htm>

[19] Video sobre redes

<http://www.monografias.com/trabajos10/vire/vire.shtml>

[20] La FAQ del DVD

<http://club.idecnet.com/%7Emodegar/video/dvdfaq.html>

[21] Evelio Martínez, conversión analógico digital

<http://www.eveliux.com/fundatel/analogdigital.html>

[22] ESTÁNDAR VIDEO, ANÁLISIS COMPRENSIVO.

<http://translate.google.com/translate?hl=es&sl=en&u=http://www.altvetmed.com/face/32720videostandard.html&prev=/search%3Fq%3Dstandars%2Bde%2Bvideo%26start%3D10%26hl%3Des%26lr%3D%26sa%3DN>

Fredy Fabián Cuello Rojas, Juan Carlos Rueda Erazo. COMPRESION DE VIDEO DIGITAL, BAJO LOS ESTANDARES MPEG.

http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/doc_b.htm

<http://www.terra.es/tecnologia/articulo/html/tec10394.htm>

<http://www.diac.upm.es/docencia/plan2000/sonido%20e%20imagen/STV2002.htm>

<http://www.domotica.net/Cable.htm>

<http://translate.google.com/translate?hl=es&sl=en&u=http://www.cs.cf.ac.uk/Dave/Multimedia/node255.html&prev=/search%3Fq%3Dmpeg%2B3%2Bvideo%26hl%3Des%26lr%3D>

<http://translate.google.com/translate?hl=es&sl=en&u=http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg4.htm&prev=/search%3Fq%3Dmpeg%2B4%2B%26hl%3Des%26lr%3D>

www.ucm.es/info/multidoc/multidoc/revista/cuadern5/blesa.htm

<http://www2.netexplora.com/sinopsis/tv1.htm>

<http://ict.pue.udlap.mx/people/raulms/compression.html>

dolby laboratories inc. broadcast distribution with dolby e

http://www.dolby.com/professional/pro_audio_engineering/solutions_dolbye.html

<http://mpeg-industry.com>

<http://news.sel.sony.com/pressrelease/5118>

<http://www.areacanopus.com>

www.mediasoftware.sonypictures.com

www.hdvxdv.com

www.adamwilt.com/HDV

www.blackmagic-design.com

<http://www.wmvhd.com/>

InterActual Player

<http://player.interactual.com/>

Microsoft Windows Media 9 Series

www.microsoft.com/windows/windowsmedia

Microsoft Windows Media : High Definition Video: WMV HD

<http://www.wmvhd.com/>

Microsoft Windows Media: Professional AV

www.microsoft.com/windows/windowsmedia/proav

Sonic Solutions DVD Producer

www.sonic.com/products/dvdproducer

Títulos:

Standing In The Shadows of Motown: Site

<http://www.standingintheshadowsofmotown.com/>

Step Into Liquid: Support FAQ

www.synccast.com/faq/stepintoliquid

Mpeg-3: Originalmente para HDTV (1920 x 1080), consiguieron doblado en Mpeg-2

Mpeg-4: Apuntado originalmente en kb/sec de velocidad de transmisión de bites muy bajo de la comunicación (4,8 a 64). Ahora tratando el vídeo que procesa...

G.K. Wallace, *El JPEG Todavía Representa Estándar De la Compresión*

CCITT, *Recomendación H.261*

D. Le Gall, *MPEG: Un estándar video de la compresión para los usos de los multimedia*

K. Patel, et. al., *funcionamiento de un decodificador del vídeo del MPEG del software*

P. Cosman, et. al., *usando el quantization del vector para el proceso de imagen*

"codificación de cuadros móviles y del audio asociado para los medios de la memoria numérica en hasta cerca de 1,5 Mbit/s," ISO/IEC 11172-2: Vídeo (Noviembre De 1991).

"codificación genérica de cuadros móviles y de la información audio asociada: Vídeo, "ISO/IEC 13818-2: Proyecto de norma internacional (Noviembre De 1994).

Barry G. Haskell, Atul Puri, Arun N. Netravali, Vídeo Digital: Una introducción a Mpeg-2, a Chapman y a Pasillo, 1997.

K.R. Rao, P. Yip, Discreto Coseno Transformar Algoritmo, Ventaja, Uso, Académico Presionar, Inc., 1990.

Majid Rabbani, Paul W. Jones, Técnicas De la Compresión De la Imagen Digital, Prensa Óptica De la Ingeniería de SPIE, 1991.

Joan L. Mitchell, Guillermo B. Pennebaker, Sábalo E. Fogg, Didier J. LeGall, compresión video Standard, Chapman y Pasillo, 1997 del MPEG.

Medios Micro Que procesan, Sociedad De la Computadora de IEEE, Volumen 16 Número 4, Agosto De 1996 Del Del Compartimiento de IEEE.

ANEXO

COMPARACION ENTRE FORMATOS DE VIDEO.

Formatos de codificación video análogos

- NTSC
- Amigacho
- SECAM
- Formatos eléctricos video análogos
- Rf
- vídeo compuesto
- vídeo componente
- S-Vi'deo
- Rgb
- Formatos análogos de la cinta video
- Ampex
- VERA (BBC)
- U-matic (Sony)
- Betamax (Sony)
- Betacam
- SP De Betacam
- VHS (JVC)
- S-vhs (JVC)
- Vhs-c (JVC)
- Video 2000 (Philips)
- Hi8
- Formatos de la cinta video Digital
- D1 (Sony)
- D2 (Sony)
- D3
- D4
- D5

- Digital Betacam (Sony)
- Betacam IMX (Sony)
- DV
- MiniDV
- Digital8 (Sony)
- Formatos de almacenaje ópticos de disco:
- DVD
- Laserdisc
- Formatos de codificación seleccionados del vídeo Digital:
- CCIR 601
- Mpeg-2
- H.261
- H.263
- H.264

Comparacion entre formatos de Video

Columnas

- P DVC: PAL Digital Video
- DVD: Digital Video Disc
- P Ld: PAL Laserdisc
- N Ld DD: NTSC Laserdisc con 5.1 canales de audio es Dolby Digital (AC-3)
- N Ld PCM: NTSC Laserdisc con audio PCM digital convencional.
- P Bcast: PAL Emision
- N Bcast: NTSC Emision
- P SVHS: PAL Super-VHS
- P Hi-8: PAL Hi-8
- P Beta: PAL Beta
- P VHS: PAL VHS
- N VHS: NTSC VHS

- P V8: Pal Video-8. Ten en cuenta que el audio digital no es típico de los modelos de gran público. Mira en Hi-8 para ver información de sonido analógico
- CD-I: Philips CD-I o MPEG-1

Filas

- Video
 - Dominio: Esta la señal de video en el dominio digital o el analógico?
 - Compresión: Si es digital, cuánto se comprime la señal? Si es analógico, cómo se almacena la señal?
 - Comp: Video compuesto: Es responsabilidad del receptor recobrar la luminancia y la crominancia de la señal.
 - Y/C: Video separado. La crominancia y la luminancia de la imagen se almacenan en el haz de video por separado. Por lo tanto, pueden ser enviadas al receptor por separado. (A menudo no se da el caso, por ejemplo, aun tengo que ver un video VHS casero con salidas Y/C)
 - Luminancia Res. Horiz.: Luminancia resolución horizontal (líneas)
 - Luminancia Res. Vert.: Luminancia resolución vertical (líneas)
 - Luminance Relac. S/N: Luminancia, relación señal/ruido (dB)
 - Crominancia Res Horiz.: Crominancia, resolución horizontal (líneas)
 - Crominancia Res. Vert.: Crominancia, resolución vertical (líneas)
 - Crominancia Relac. S/N : Crominancia, relación señal/ruido (dB)
 - Cuadros/s: Cantidad máxima de planos completos por segundo (p/s)
 - Intrelacado: Se muestra la imagen en modo entrelazado?
 - Artefactos de Compr. : Si la imagen es comprimida, produce la compresión artefactos visibles? (los puntos parece que se arrastran cuando un aparato de TV no puede separar bien la luminancia de la crominancia)
- Audio
 - Dominio: Esta la señal de audio en el dominio digital o analógico?
 - Compresión: Si es digital, cuánto está el sonido comprimido? Si es analógico, qué método de compresión/expansión se ha usado?

- Bits Muestreo: Si es digital, la cantidad de bits usados en muestrear antes de la compresion.
- Velocidad de Muestreo: Si es digital, la velocidad de muestreo (muestras/seg)
- Relacion S/N: La relacion Señal/ruido del sonido con reduccion de ruido (dB)
- Baja Frecuencia: La frecuencia mas baja que el sistema puede reproducir sin atenuar mas de s dB (Hz)
- Alta Frecuencia: La frecuencia mas alta que el sistema puede reproducir sin atenuar mas de 3 dB (Hz)
- Canales (Calidad + baja) Los canales que pueden usarse. En parentesis estan los canales analogicos de los laserdisc NTSC y las pistas lineales de video VHS.

Calculando la Resolucion Horizontal

Aunque calcular la resolucion vertical es facil (Cantidad de lineas de escaneo horizontales), la resolucion horizontal es una cosa un poco engañosa.

En la industria del video, la resolucion horizontal se da como si la pantalla de video fuese tan ancha como alta. De este modo, si un dispositivo de video tiene la misma resolucion vertical, se puede pensar que tiene "pixels cuadrados". Si quieres saber cunatas lineas verticales puede mostrar realmente el dispositivo (en el ancho total de la pantalla, digamos), tienes que multiplicar la resolucion dada por 4/3 que es la relacion normal para dispositivos de video normales.

Por ejemplo, si tienes un ordenador Amiga que puede mostrar graficos (en modo entrelazado baja resolucion) en una resolucion de ordenador de 368x564, la resolucion de video es $368 \cdot \frac{3}{4} \times 564 = 276 \times 564$, lo cual es un poco mejor que el VHS. El paso inverso: Si tienes un reproductor laserdisc NTSC, que tiene una resolucion de 420 lineas, puedes tener $420 \cdot \frac{4}{3} = 560$ lineas verticales, o pixels en horizontal, en tu pantalla de television.

Por esta causa, por ejemplo, la resolución de la luminancia horizontal del CD-I no se da como 320, sino como 240 ($320 \cdot 3/4$). El número podría incluso ser más bajo si se usase un filtrado paso-bajo apropiado, pero en el CD-I no se hace nada para prevenir que los píxeles salgan mal.

Por otra parte, la resolución teórica del DVD podría ser de 540 líneas de video según mis cálculos. Sin embargo, puesto que la señal es filtrada paso-bajo (como debe ser) se pierde un poco menos del 10% (40 líneas) en el proceso. [22]

COMPARACION DE FORMATOS DE VIDEO

Audio		N DVD DD	P Ld	N DD	N Ld DD	N Ld PCM	P Bcast	N Bcast	P SVHS	P Hi-8	P Beta	P VHS	N VHS	P V8	CD-I
Dominio	Digital	Digital	Digital	Digital	Digital	Digital	Digital	Analog	Analog	Digital	Analog	Analog	Analog	Digital	Digital
Compresión	No	10:1	No	10:1	No	No	14:10	-	DBX	N-lin.		DBX	DBX	N-lin.	5:1
Bits Muestreo	16	18..20	16	18..20	16	14	14	-	-	8	-	-	-	8	16
Velocidad de muestreo	48000	48000	44100	48000	44056	32000	32000	-	-	32000	-	-	-	32000	
Relación . S/N	96	90	96	90	96	84	84	65	90	85	90	90	90	85	96
Baja Frecuencia	5	5, 2	5	5, 2	5	5	5	60	20	20	20	20	20	20	20
Alta Frecuencia	22k	20k, 120	20k	20k, 120	20k	15k	15k	15k	20k	15k	20k	20k	20k	15k	20k
Canales (calidad + baja)	2	n x 5.1	2	2 (+2)	2 (+1)	1..2	1..2	2 (+1)	2 (+2)	2	2 (+1)	2 (+1)	2 (+2)	2	