

PROYECTO FIN DE CARRERA

**DEFINICIÓN Y OPERATIVA INICIAL PARA
LA INTRODUCCIÓN DE SERVICIOS DE
GRABACIÓN LOCAL EN PLATAFORMAS DE
IPTV**



Universidad de Sevilla

Autor: Javier García Villard.
Tutor: Carmen Serrano Gotarredona.

MAYO 2007

- 0 -

A mis padres,

1. MOTIVACIÓN Y OBJETIVOS.	- 8 -
1.1. GUÍA ELECTRÓNICA DE PROGRAMACIÓN.	- 9 -
1.2. SERVICIOS DE PUBLICIDAD INTERACTIVA.	- 9 -
1.3. E-LEARNING.	- 9 -
1.4. JUEGOS EN TELEVISIÓN.	- 10 -
1.5. GESTIÓN DEL CORREO ELECTRÓNICO Y ACCESO A INTERNET.	- 10 -
1.6. E-COMMERCE.	- 10 -
1.7. VIDEOCONFERENCIA.	- 11 -
1.8. VÍDEO BAJO DEMANDA.	- 11 -
1.9. RESTRICCIÓN DE ACCESO. CONTROL PARENTAL.	- 11 -
1.10. TELEVISIÓN DE ALTA DEFINICIÓN.	- 11 -
1.11. ALMACENAMIENTO EN DISCO DURO.	- 12 -
1.12. INTEGRACIÓN DE LA TELEFONÍA Y LA TELEVISIÓN.	- 12 -
1.13. GRABACIÓN DE CONTENIDOS.	- 12 -

CAPITULO I. INTRODUCCIÓN A LA IPTV.

2. ESTADO DE LA IPTV EN EL MUNDO.	- 15 -
3. ARQUITECTURA DE UNA RED IPTV.	- 17 -
3.1. CABECERA CENTRAL DE TV.	- 17 -
3.1.1. CENTRO DE PROCESO DIGITAL.	- 18 -
3.1.2. SERVIDOR CENTRAL DE VÍDEO BAJO DEMANDA.	- 19 -
3.1.3. SERVIDOR DE GESTIÓN CENTRAL DE VÍDEO BAJO DEMANDA.	- 19 -
3.1.4. REMULTIPLEXOR IP.	- 20 -
3.1.5. SERVIDOR DE ARRANQUE.	- 20 -
3.2. CÉLULA DE SERVICIOS CENTRAL.	- 20 -
3.2.1. CENTRO DE SERVICIOS CENTRAL.	- 20 -
3.2.1.1. BASE DE DATOS DE ABONADOS	- 21 -
3.2.1.2. SERVIDOR DE APLICACIONES	- 21 -
3.2.1.3. BASE DE DATOS DE CONTENIDOS	- 21 -
3.2.2. CENTRO DE GESTIÓN CENTRAL	- 22 -
3.2.2.1. SISTEMA DE PROVISIÓN DE RED DE ACCESO	- 22 -
3.2.2.2. SISTEMA DE SUPERVISIÓN	- 22 -
3.2.2.3. SISTEMA DE GESTIÓN DE ELEMENTOS DE RED.	- 22 -
3.2.2.4. SISTEMA DE TARIFICACIÓN	- 23 -
3.3. RED DE TRANSPORTE	- 23 -
3.3.1. RED DE TRANSPORTE DE VÍDEO.	- 24 -
3.3.2. RED DE TRANSPORTE DE DATOS.	- 24 -
3.3.3. RED DE GESTIÓN.	- 24 -

3.3.4.	RED DE DISTRIBUCIÓN DE CONTENIDOS.	- 24 -
3.3.5.	RED DE TRANSPORTE DE AUTORIZACIÓN	- 24 -
3.3.6.	RED DE BACKUP	- 25 -
3.4.	CÉLULA DE SERVICIOS LOCAL	- 25 -
3.4.1.	SERVIDORES DE CONTENIDOS BAJO DEMANDA	- 26 -
3.4.1.1.	CONFIGURACIÓN EN RAID	- 28 -
3.4.1.2.	NIVELES RAID	- 29 -
3.4.1.2.1.	RAID 1	- 29 -
3.4.1.2.2.	RAID 2	- 29 -
3.4.1.2.3.	RAID 3	- 30 -
3.4.1.2.4.	RAID 4	- 30 -
3.4.1.3.	DIMENSIONAMIENTO DE SERVIDORES LOCALES DE VOD.	- 31 -
3.4.1.4.	EL PROBLEMA DEL ENVÍO	- 33 -
3.4.1.5.	PROTOCOLOS PARA TRANSMISIÓN DE FLUJOS DE VÍDEO.	- 34 -
3.4.1.5.1.	PROTOCOLOS DE NIVEL DE ENLACE.	- 34 -
3.4.1.5.2.	IP (INTERNET PROTOCOL).	- 34 -
3.4.1.5.3.	UDP Y TCP.	- 34 -
3.4.1.5.4.	RTP (REAL TIME PROTOCOL)	- 35 -
3.4.1.5.5.	RTCP	- 35 -
3.4.1.5.6.	RTSP (REAL TIME STREAMING PROTOCOL)	- 36 -
3.4.1.5.7.	MPEG-2 TRANSPORT STREAM	- 36 -
3.4.2.	SISTEMA DE GESTIÓN Y CONTROL	- 37 -
3.4.3.	NODO DE ACCESO.	- 38 -
3.5.	UBICACIÓN DE ABONADO	- 38 -
3.5.1.	INTERFAZ DE RED	- 39 -
3.5.2.	DECODIFICADOR	- 39 -
3.5.3.	BUFFER	- 39 -
3.5.4.	HARDWARE DE SINCRONIZACIÓN	- 39 -

CAPITULO II. DESCRIPCIÓN DE LOS SERVICIOS DE GRABACIÓN DE CONTENIDOS.

4.	GRABACIÓN DE CONTENIDOS	- 41 -
4.1.	TIME SHIFTING	- 41 -
4.1.1.	MODOS DE OPERACIÓN.	- 41 -
4.1.1.1.	PAUSA DE LA EMISIÓN	- 41 -
4.1.1.2.	REANUDACIÓN DE LA EMISIÓN	- 42 -
4.1.1.3.	AVANCE RÁPIDO	- 42 -
4.1.1.4.	RETROCESO RÁPIDO	- 43 -
4.1.1.5.	SALTO AL DIRECTO	- 43 -
4.1.1.6.	REANUDACIÓN DE LA GRABACIÓN	- 44 -
4.1.1.7.	PICTURE IN PICTURE	- 44 -

4.1.2.	MODO NORMAL DE FUNCIONAMIENTO. TIME SHIFTING BÁSICO.	- 45 -
4.1.3.	MODO DE GRABACIÓN CONTINUA. TIME SHIFTING AVANZADO.	- 46 -
4.1.4.	IMPLICACIONES EN EL DISCO DURO DEL STB	- 47 -
4.2.	DIGITAL VIDEO RECORDER	- 48 -
4.2.1.	MODALIDADES DE GRABACIÓN	- 48 -
4.2.1.1.	MÉTODO DE GRABACIÓN IMPULSIVA.	- 49 -
4.2.1.2.	GRABACIÓN PROGRAMADA DESDE LA EPG.	- 50 -
4.2.1.3.	GRABACIÓN POR HORA ABSOLUTA DE INICIO Y FIN.	- 50 -
4.2.2.	FUNCIONALIDADES AÑADIDAS DEL SERVICIO DVR	- 51 -
4.3.	INTEROPERABILIDAD DE DVR Y TIME SHIFTING.	- 51 -

CAPITULO III. ESTUDIO DEL IMPACTO DE LOS SERVICIOS DE GRABACIÓN.

5.	IMPACTO EN EL SET-TOP-BOX	- 54 -
5.1.	CÁLCULO DE CAPACIDAD DE DISCOS DUROS	- 55 -
5.2.	NECESIDAD DE SINTONIZADORES DE TDT.	- 56 -
5.3.	ELECCIÓN DEL SET-TOP-BOX	- 57 -
5.3.1.	SAGEM	- 57 -
5.3.2.	KREATEL	- 57 -
5.3.3.	ADB	- 57 -
6.	DERECHOS LEGALES SOBRE LA GRABACIÓN DE CONTENIDOS	- 59 -
7.	IMPACTO EN LA ARQUITECTURA Y SISTEMAS DE LA RED	- 60 -
7.1.	IMPACTO EN EL SERVIDOR DE APLICACIONES	- 60 -
7.2.	IMPLICACIONES EN LA EPG	- 60 -
7.2.1.	MARCA DE DERECHOS DE GRABACIÓN DVR.	- 61 -
7.2.2.	MARCA DE DERECHOS TIME SHIFTING.	- 61 -
7.2.3.	MARCA DE CONTENIDO PPV.	- 61 -
7.2.4.	MARCA DE CONTENIDO VBD.	- 61 -
7.2.5.	MARCAS DE CONTENIDO EN FORMATO MPEG-2 Y MPEG-4.	- 62 -
7.2.6.	MARCAS DE CONTENIDO EN FORMATO SDTV Y HDTV.	- 62 -
7.2.7.	MARCA DE CONTENIDO REPETIDO.	- 62 -
7.2.8.	MARCA DE PERIODICIDAD DEL CONTENIDO	- 62 -
8.	IMPACTO EN SISTEMAS COMERCIALES	- 63 -

CAPITULO IV. ESTUDIO DE RENTABILIDAD

9. VIABILIDAD ECONÓMICA.	- 66 -
9.1. IMPACTO EN INGRESOS DEL OPERADOR.	- 66 -
9.2. ESTUDIO DE VIABILIDAD ECONÓMICA.	- 66 -

CAPITULO V. TENDENCIAS EN LA GRABACIÓN DE CONTENIDOS.

10. GRABACIÓN DE CONTENIDOS EN RED.	- 72 -
--	---------------

APENDICES

11. VOCABULARIO ESPECÍFICO.	- 75 -
12. REFERENCIAS.	- 76 -

ANEXOS.

13. ESTÁNDARES MPEG.	- 78 -
14. ESTÁNDAR MPEG-1.	- 80 -
14.1. SECCIÓN DE VÍDEO	- 80 -
14.1.1. LAS IMÁGENES I (INTRA).	- 81 -
14.1.2. LAS IMÁGENES P (PREVISTAS).	- 81 -
14.1.3. LAS IMÁGENES B (BIDIRECCIONALES)	- 82 -
14.1.4. DESCOMPOSICIÓN EN CAPAS DE UNA SECUENCIA DE VIDEO MPEG	- 84 -
14.1.4.1. BLOQUE (BLOCK)	- 86 -
14.1.4.2. MACROBLOQUE (MACROBLOCK)	- 86 -
14.1.4.3. REBANADA (SLICE)	- 86 -
14.1.4.4. IMAGEN DE TIPO I, P O B	- 87 -
14.1.4.5. GRUPO DE IMÁGENES (GROUP OF PICTURES O GOP)	- 87 -
14.1.4.6. SECUENCIA	- 88 -
14.1.5. CODIFICACIÓN INTRACUADRO	- 88 -
14.1.5.1. TRANSFORMACIÓN DEL ESPACIO DE COLOR	- 89 -
14.1.5.2. SUBMUESTREO	- 90 -

14.1.5.3. TRANSFORMACIÓN DISCRETA DEL COSENO	- 90 -
14.1.5.4. CUANTIZACIÓN	- 92 -
14.1.5.5. CODIFICACIÓN ENTRÓPICA	- 93 -
14.1.5.6. DECODIFICACIÓN	- 94 -
14.1.6. CODIFICACIÓN DIFERENCIAL CON COMPENSACIÓN DEL MOVIMIENTO INTERCUADRO.	- 94 -
14.1.7. CONTROL DE BUFFERING	- 98 -
14.2. SECCIÓN DE AUDIO	- 99 -
14.2.1. MPEG-1 CAPA I.	- 100 -
14.2.1.1. MAPEO DEL TIEMPO A LA FRECUENCIA	- 100 -
14.2.1.2. MODELO PSICOACÚSTICO	- 100 -
14.2.1.3. CUANTIZADOR / CODIFICADOR	- 101 -
14.2.1.4. EMPAQUETAMIENTO DE TRAMA	- 101 -
14.2.2. MPEG-1 CAPA II.	- 102 -
14.2.3. MPEG-1 CAPA III.	- 103 -
14.3. SECCIÓN DE SISTEMA	- 104 -
14.3.1. PACKETIZED ELEMENTARY STREAM (PES)	- 105 -
14.4. SECCIÓN DE PRUEBAS DE LA CONFORMIDAD.	- 111 -
14.5. SOFTWARE DE REFERENCIA.	- 111 -
15. ESTANDAR MPEG-2	- 112 -
15.1. SECCIÓN DE VÍDEO	- 112 -
15.1.1. PERFILES Y NIVELES MPEG-2	- 114 -
15.1.1.1. NIVELES	- 114 -
15.1.1.2. PERFILES	- 115 -
15.1.2. MODOS ESCALABLES	- 116 -
15.1.3. MODOS DE PREDICCIÓN ESPECÍFICOS MPEG-2 (IMÁGENES ENTRELAZADAS)	- 120 -
15.1.4. LA ESTRUCTURA FRAME	- 121 -
15.1.5. LA ESTRUCTURA FIELD	- 122 -
15.1.5.1. EL MODO FRAME	- 123 -
15.1.5.2. EL MODO FIELD	- 123 -
15.1.5.3. EL MODO MIXTO	- 123 -
15.1.6. SCANNING (EXPLORACIÓN)	- 123 -
15.1.7. DESCRIPCIÓN DEL PROCESO DE CODIFICACIÓN MPEG-2	- 125 -
15.1.8. DESCRIPCIÓN DEL PROCESO DE DECODIFICACIÓN MPEG-2	- 127 -
15.2. LA SECCIÓN DE AUDIO MPEG-2	- 128 -
15.3. PARTE DE SISTEMA MPEG 2 PARA MULTIPLEXIÓN Y TRANSPORTE	- 130 -
16. ESTÁNDAR MPEG4	- 134 -
16.1. PARTES DE MPEG-4	- 134 -
16.2. NUEVAS FUNCIONALIDADES	- 136 -
16.2.1. ACCESO A DATOS MULTIMEDIA BASADO EN CONTENIDO	- 136 -
16.2.2. EDICIÓN Y MANIPULACIÓN DEL FLUJO DE DATOS CODIFICADOS BASADOS EN CONTENIDO	- 137 -

16.2.3.	CODIFICACIÓN DE DATOS HÍBRIDOS NATURALES Y SINTÉTICOS	- 137 -
16.2.4.	ACCESO TEMPORAL ALEATORIO MEJORADO	- 137 -
16.2.5.	EFICIENCIA DE COMPRESIÓN MEJORADA	- 137 -
16.2.6.	CODIFICACIÓN DE MÚLTIPLES FLUJOS DE DATOS CONCURRENTES	- 137 -
16.2.7.	ROBUSTEZ EN ENTORNOS CON PROPENSIÓN A ERROR	- 137 -
16.2.8.	ESCALABILIDAD BASADA EN CONTENIDO	- 138 -
16.3.	CODIFICACIÓN ORIENTADA A OBJETOS	- 138 -
16.4.	EL MODELO DE VERIFICACIÓN	- 139 -
16.5.	EL LENGUAJE DE DESCRIPCIÓN DE ALGORITMOS	- 140 -
16.6.	CODIFICACIÓN DE AUDIO.	- 141 -
16.6.1.	CODIFICACIÓN ORIENTADA A LA TELEFONÍA	- 142 -
16.6.1.1.	G.728	- 143 -
16.6.1.2.	G.729	- 143 -
16.6.1.3.	G.723.1	- 143 -
16.6.2.	SISTEMAS DE CODIFICACIÓN DE AUDIO ORIENTADOS A CONTENIDOS DE TIPO GENÉRICO	- 143 -
16.7.	PROTECCIÓN DE DERECHOS DE AUTOR. CONTROL DE ACCESO A LA INFORMACIÓN.	- 145 -
16.8.	EL MPEG-4 SP Y ASP	- 146 -
16.9.	H.264 O MPEG-4 AVC MPEG-4	- 146 -
16.10.	COMPARACIÓN DE FORMATOS	- 146 -
17.	DATA SHEETS DE SET-TOP-BOXES	-148-

1. MOTIVACIÓN Y OBJETIVOS.

Las siglas IPTV se corresponden con Internet Protocol Televisión, es decir, lo que se conoce comúnmente como Televisión sobre IP. Este protocolo describe la manera a través de la cual podemos recibir una señal de vídeo a través de una red IP.

La Televisión IP está íntimamente ligada al desarrollo del ancho de banda de las comunicaciones. IPTV se muestra como el resultado del gran auge de conexiones a Internet y del avance tecnológico que ha permitido ofrecer a los usuarios un mayor ancho de banda a un menor precio. Esta solución hace posibles nuevas opciones de entretenimiento y servicios para los usuarios, y la generación de mayores ingresos para los operadores que los ofrecen explotando las infraestructuras ya existentes.

La libre competencia obliga a los operadores (no sólo de IPTV, sino también de TV por cable y televisión satélite) a ofrecer nuevos servicios innovadores que contribuyan a la captación de clientes. La evolución natural es ofrecer a los usuarios una experiencia totalmente personalizable. El cliente demanda una televisión inteligente, hecha a su medida, en la que él mismo pueda:

- Acceder a una mayor oferta de programación.
- Elegir qué es lo que quiere ver y cuándo desea verlo.
- Disfrutar de una mayor calidad de imagen y sonido.
- Obtener información adicional sobre la programación.
- Realizar grabaciones de sus emisiones favoritas.
- Desechar toda la publicidad que no le interesa y obtener más información sobre la que sí le interesa.
- Recibir formación específica.
- Realizar videoconferencias
- Disfrutar de juegos en su tiempo libre.
- Restringir el acceso a ciertos contenidos.
- Gestionar su correo electrónico.
- Acceder a Internet.
- Realizar compras.
- Almacenar información personal.
- Integrar la telefonía en el televisor.
-

Para satisfacer esta demanda de mercado, aparecen nuevos servicios ^{[1] [2] [3]}:

1.1. Guía Electrónica de Programación.

La guía electrónica de programación (EPG) permite al usuario acceder de manera rápida, sencilla y organizada a todos los canales que le ofrece su proveedor de televisión. La EPG representa la evolución a la era digital del tradicional servicio de programación que nos ofrece el teletexto. Así el usuario puede hacer una elección de lo que desea ver por televisión sin necesidad de recurrir al habitual zapping.

Desde la EPG además se ofrece la posibilidad de realizar una búsqueda exhaustiva seleccionando diferentes temáticas: deportes, series, películas, informativos, magazines, etc, que recogen el conjunto de programas que resultan de interés al usuario ^[4].

También se presenta al usuario información adicional sobre la programación. La EPG puede presentar, además de la duración de un programa, programas siguientes y anteriores, progreso del programa actual, información sobre el director de una película, actores principales, sinopsis, fecha de estreno, premios, crítica, etc.

La EPG se convierte por todo esto en una facilidad que se considera ya imprescindible para cualquier operador de televisión digital.

1.2. Servicios de publicidad interactiva.

El usuario de televisión puede definir cuáles son sus preferencias y recibir únicamente publicidad sobre aquello que le interese. Del mismo modo, puede recibir más información de aquellos productos que consigan llamar su atención, e incluso realizar pedidos. Se podría incluir publicidad interactiva no sólo durante las pausas para espacios publicitarios, sino también durante la emisión de sus programas favoritos.

1.3. e-Learning.

A través de la televisión pueden recibirse cursos de formación para todos los niveles. Es la adaptación del ritmo de aprendizaje al alumno y la disponibilidad de

las herramientas de aprendizaje, independientemente de límites horarios o geográficos. El usuario podría recibir cursos de inglés interactivos, documentales sobre ciencia o naturaleza, o la formación que una empresa estime conveniente dar a sus empleados.

1.4. Juegos en Televisión.

La demanda por parte de los usuarios de servicios de juegos en televisión es un reflejo de la necesidad que sienten por integrar en un único elemento todas sus necesidades de ocio y tiempo libre. Este es un servicio que proporciona un valor añadido y fidelizador del cliente con el proveedor de servicios de televisión.

1.5. Gestión del correo electrónico y acceso a Internet.

Aprovechando la conexión de banda ancha de los operadores de cable e IPTV, el usuario puede acceder a Internet y navegar por páginas Web, así como enviar y recibir correos electrónicos, todo esto desde su televisor.

Si bien la experiencia ha demostrado que la televisión no resulta un medio atractivo para el acceso a sitios de Internet, es posible utilizar la red para diseñar portales específicos para el formato televisivo, como por ejemplo portales de banca electrónica, que introduzcan servicios para un segmento de la población que no está habituado ni predispuesto al uso de Internet.

1.6. e-Commerce.

Con este servicio se ofrece al usuario la posibilidad de realizar compras a través de la televisión desde la comodidad de su sofá. La pérdida de la desconfianza por los sistemas de compra electrónica hace que esta clase de servicios estén en auge. Además de acceder a portales específicos de televenta, este servicio puede combinarse con el de publicidad interactiva. A modo de ejemplo, durante el transcurso de un evento deportivo, podría presentarse al usuario la posibilidad de comprar el balón del mundial de fútbol, o entrar en una subasta por el casco de un conductor de Fórmula1.

1.7. Videoconferencia.

El servicio de videoconferencia se ofrece como un servicio de doble propósito, puede estar enfocado tanto a necesidades de particulares como de empresas. Mediante este servicio, un usuario puede mantener una conexión de audio y vídeo con otro locutor distante.

1.8. Vídeo bajo Demanda.

El usuario puede solicitar al operador un contenido y tener control sobre el, es decir, puede parar y reanudar a discreción el material que está visionando, y también puede congelar imágenes, posicionarse en un punto del vídeo, retroceder y avanzar rápidamente, etc. como si de un vídeo doméstico se tratara, si bien en general se consigue una calidad superior.

El servicio permite el acceso a un catálogo de contenidos audiovisuales con las mismas prestaciones que si el contenido estuviese disponible en un DVD adquirido o alquilado. Este tipo de servicio es el más diferenciador, en la cual la comunicación entre el cliente y la red no es compartida con ningún otro cliente. Es un gran exponente en la personalización de la experiencia televisiva, pues sigue la filosofía de "lo que quieras y cuando quieras".

1.9. Restricción de acceso. Control parental.

Mediante esta facilidad, el usuario puede solicitar autenticación para un grupo cerrado de contenidos. Típicamente, esta facilidad es utilizada por padres de familia (de ahí su nombre) para limitar el acceso a contenidos de pago o clasificados.

1.10. Televisión de alta definición.

Gracias a nuevos estándares en la codificación de vídeo y audio, el usuario puede recibir en su ubicación televisión en formato de alta definición. Actualmente este servicio experimenta un gran auge. No obstante, no siempre es posible ofrecer al usuario televisión en calidad HDTV (High Definition Television), puesto que implica un aumento sustancial del ancho de banda de transmisión, y existen bucles de abonado en los que, por la distancia entre el usuario y el nodo que le presta servicio, no puede satisfacerse dicha demanda. Del mismo modo, sería

necesario un televisor que sea capaz de sacarle el máximo rendimiento y mostrar una mayor definición de la imagen. Aún así, este servicio se considera el siguiente escalón en la Televisión Digital, algo parecido a lo que ocurre entre la Televisión Digital Terrestre y la emisión analógica. Del mismo modo, el ya clásico euroconector necesita ser sustituido por un conector HDMI (High Definition Multimedia Interface).

1.11. Almacenamiento en disco duro.

Los últimos Set-Top-Boxes incorporan un disco duro para realizar grabaciones de audio y video. Del mismo modo, ese disco duro también se puede utilizar como unidad de almacenamiento externa al PC. Los STB dotados de disco duro incluyen herramientas que permiten, entre otros, visualizar las imágenes almacenadas, reproducir música en formato MP3, gestionar archivos y carpetas, etc. La conexión entre el STB y el PC se realiza mediante un puerto USB, y una vez desconectado, pueden accederse a los contenidos y realizar las acciones ya mencionadas.

1.12. Integración de la telefonía y la televisión.

En este afán de conseguir la personalización de la experiencia televisiva y la integración con todos los elementos que conforman su actividad diaria, se le presenta al usuario la posibilidad de interactuar con el teléfono, tanto fijo como móvil, desde la televisión. El usuario recibe indicadores en su televisor de llamadas entrantes, mensajes recibidos, llamadas perdidas... Asimismo puede aceptar o rechazar llamadas entrantes, o personalizar su agenda de teléfono.

1.13. Grabación de contenidos.

Pero sin duda alguna, el servicio que es mejor valorado por los usuarios y que despierta un mayor interés es la grabación del vídeo y audio procedentes de la televisión ^[5].

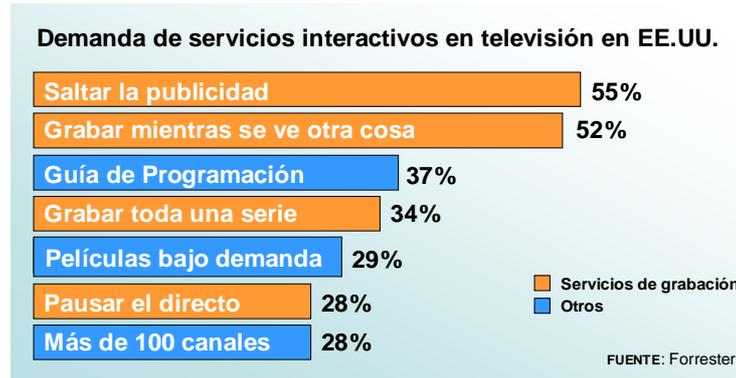


Ilustración [1] Demanda de servicios interactivos en televisión EE.UU. [6]

Los grandes operadores de TV Digital del mundo se han lanzado a una carrera por ofrecer el mejor de estos servicios. Se plantean dos formas de realizar las grabaciones.

Por un lado, dotando al STB de un disco duro pueden realizarse grabaciones en local. El usuario puede programar su STB, como si de un vídeo doméstico se tratase, y grabar un programa de televisión en su terminal. Este servicio se conoce como DVR, Digital Video Recording, y va de la mano con otro servicio, conocido como Time Shifting o gestión del directo. Gracias a este último, el usuario puede detener la emisión en curso de un programa, para retomarla más tarde, mientras la emisión se va almacenando en el disco duro. De este modo, el usuario puede pausar, avanzar o retroceder en la emisión actual.

La otra forma que se plantea es la de realizar la grabación de contenidos en red, no en el terminal del usuario. Esta forma de grabación implica grandes cambios en la arquitectura de red del operador, y por el momento, sólo algunos operadores han sido capaces de llevarla a cabo, a costa de sacrificar la disponibilidad de sus servidores para otras aplicaciones o con una enorme inversión en nuevos equipos.

El objeto de este Proyecto Fin de Carrera es la definición inicial para la introducción de los servicios de grabación de contenidos en local para una plataforma de IPTV. La metodología que seguiremos es la siguiente: en primer lugar, se presentará la arquitectura de red IPTV; a continuación, se realizará la definición de los servicios de grabación de contenidos en local; y por último, se estudiarán las implicaciones que tiene introducir servicios de grabación en local en el producto de IPTV que ofrece el operador, así como la rentabilidad de los servicios de grabación en local.

Capítulo I.

INTRODUCCIÓN A LA IPTV.

2. ESTADO DE LA IPTV EN EL MUNDO.

La IPTV es una realidad en Europa, Asia y Estados Unidos ^[7]. Sólo en Hong Kong, por ejemplo, existen más de 350 mil abonados a IPTV. En Estados Unidos existen cerca de 250 comunidades de IPTV y más de 300 mil hogares tienen contratado el servicio de televisión con empresas telefónicas. Se calcula que para finales de 2007, habrá en todo el mundo cerca de 27 millones de hogares con servicios de IPTV y que el 10% de la TV de pago en Europa será IPTV a final del año 2009.

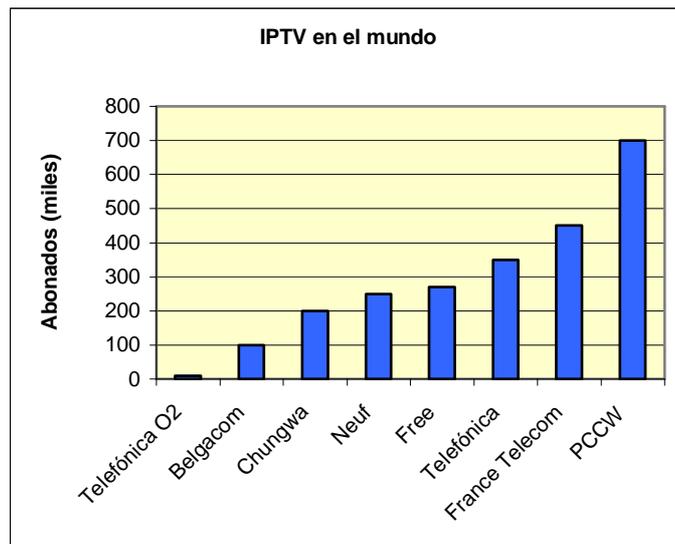


Ilustración [2] Suscriptores de IPTV en el mundo a final de 2006. ^[8]

La IPTV tuvo en 2004 una penetración del uno por ciento en el mercado de la Unión Europea, lo que movió 62 millones de euros, y se estima que en 2009 la cifra sea de 25.000 millones ^[9].

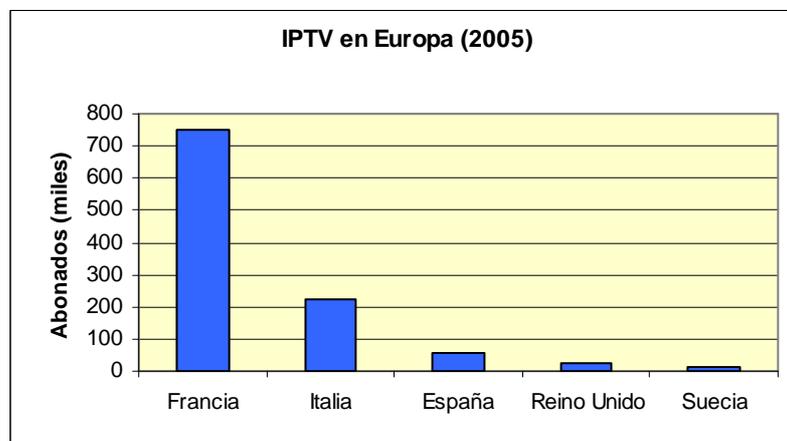


Ilustración [3] Suscriptores de IPTV en Europa a finales de 2005.

Según informe de la consultora Datamonitor ^[10], se cree que la IPTV continuará con un relativo nicho de mercado en Europa. Espera que haya entre 9,5 y 10 millones de suscriptores a plataformas IPTV en Europa para el final de 2010 ^[11] ^[12], con aproximadamente el 10% de los hogares digitales conectados por esta vía. Francia continuará con una fuerte posición con IPTV en Europa, con el 28% de los hogares suscritos para el final de 2010 .

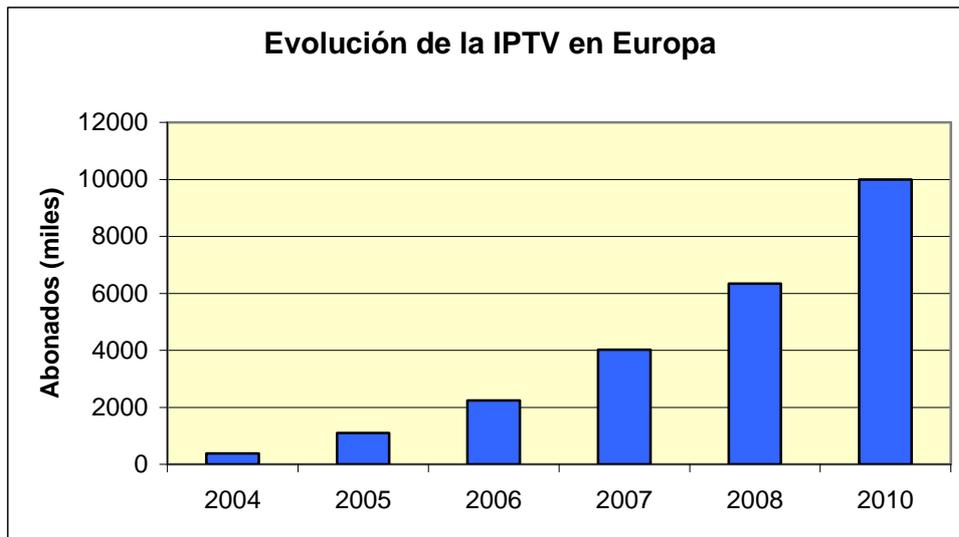


Ilustración [4] Evolución de la IPTV en Europa ^[13]

La extensión de los servicios de DVR es central para la estrategia de los operadores de la televisión. Sin embargo, la penetración sólo ha alcanzado niveles significativos en Estados Unidos, donde se estimó que sería del 18% de los hogares para finales de 2006. En Europa, las estimaciones rondaban el 2%, lo cual hace esperar un gran crecimiento en este servicio en los próximos años.

3. Arquitectura de red IPTV.

La red de cualquier operador de IPTV necesita incluir, al menos, los siguientes elementos ^[2] ^[14].

- **Cabecera de TV:** donde se reciben y codifican los canales de TV y audio.
- **Célula de Servicios Central:** donde se gestiona el funcionamiento de toda la red y se almacena una copia de los contenidos bajo demanda (en caso de que el operador ofrezca dicho servicio)
- **Red de distribución:** hace llegar los canales de TV y audio a los abonados.
- **Célula de Servicios Local:** es el punto de acceso de los usuarios a la plataforma y a contenidos bajo demanda.
- **Ubicación de cliente.**

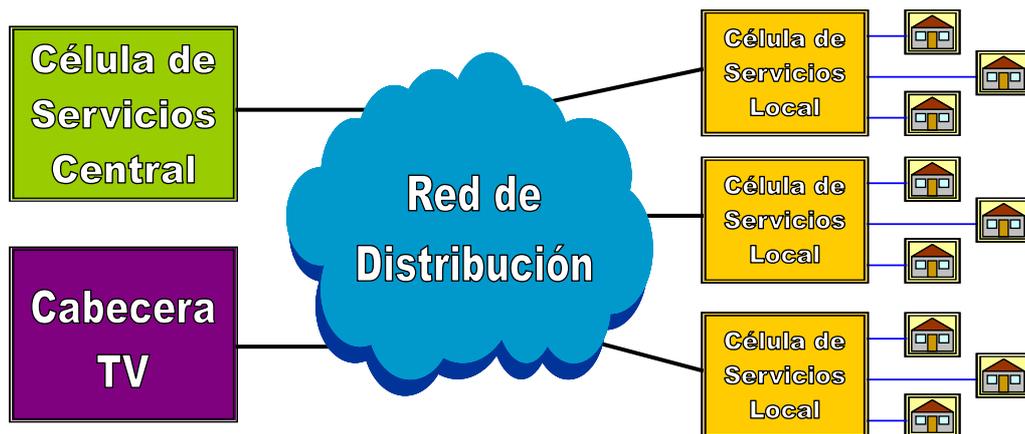


Ilustración [5] Arquitectura IPTV

3.1. Cabecera Central de TV.

En la Cabecera Central de TV ^[15] es donde se reciben las contribuciones de los distintos canales de TV y audio, que se codifican y preparan para su distribución. A continuación se describen los elementos más notables que la conforman.

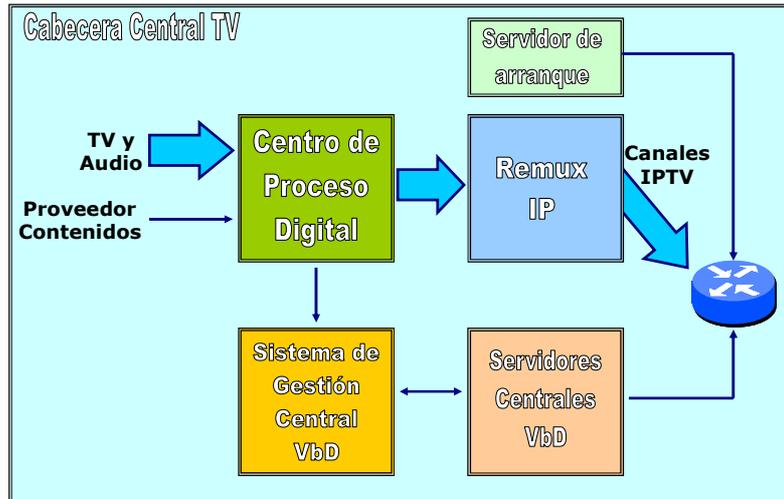


Ilustración [6] Cabecera de TV

3.1.1. Centro de Proceso Digital.

El Centro de Proceso Digital convierte las contribuciones de los canales de TV tradicional a la codificación empleada en la plataforma (típicamente MPEG-2 o MPEG-4) para reducir los requerimientos de ancho de banda de transmisión y almacenamiento. Del mismo modo, también codifica los contenidos que se vayan a poner a disposición del abonado bajo demanda desde las células de servicio locales.

Cada canal de TV digital se compone de un canal de vídeo (MPEG-2 o MPEG-4), uno o dos canales de audio (MPEG-1 Layer II) y un canal opcional de teletexto, de modo que:

- El componente de vídeo se codifica para que el bitrate total esté entre 4 y 4.5 Mbps, dependiendo del número de canales de audio utilizados y el ancho de banda reservado para el canal de teletexto. Un bitrate típico es 3.9 Mbps.
- El componente de audio consta de uno o dos canales estéreo MPEG-1 Layer II (ISO/IEC 11173-2), con un bitrate entre 32 y 384 Kbps, aunque normalmente se utilizará un bitrate en torno a 128 o 192 Kbps por cada canal estéreo.
- El componente de teletexto consta de un canal DVB-teletext ETS300472. El bitrate asignado variará según el contenido de teletexto que se transmita por cada canal, y típicamente será alrededor de 350 Kbps.

La trama de transporte MPEG-2 monoprograma (SPTS según ISO/IEC 13818-1) tiene un bitrate constante (CBR) de hasta 4.8 Mbps.

Una vez codificados, los distintos canales de TV son enviados al remultiplexor IP para su distribución por la plataforma. Los contenidos bajo demanda son enviados al Servidor de Gestión Central de Vídeo bajo demanda, y posteriormente almacenados en el Servidor Central de Vídeo bajo Demanda.

3.1.2. Servidor Central de Vídeo bajo Demanda.

El Servidor Central de Vídeo bajo Demanda almacena los contenidos bajo demanda codificados por el Centro de Proceso Digital. Desde este servidor central, los contenidos se envían al remultiplexor IP para su distribución y almacenamiento en los servidores de vídeo de las células locales, donde serán accesibles a los usuarios finales que soliciten contenidos bajo demanda.

3.1.3. Servidor de Gestión Central de Vídeo bajo Demanda.

El Servidor de Gestión Central de Vídeo bajo Demanda almacena en su base de datos la relación de contenidos bajo demanda que han sido inyectados en los servidores de las células locales mediante la red de distribución (anillo unidireccional SDH). Recibe el contenido codificado por el Bloque de Proceso Digital y genera los archivos de índice para las funcionalidades de avance rápido y retroceso, propias del Vídeo bajo Demanda, y posteriormente los envía al Servidor Central de Vídeo bajo Demanda.

Una vez que el nuevo contenido se encuentra disponible en el Servidor Central de Vídeo bajo Demanda, el Servidor de Gestión Central de Vídeo bajo Demanda ordena su distribución a los servidores de las células locales. Cuando el contenido se encuentra correctamente almacenado en los servidores de vídeo de las células locales, el Sistema de Gestión y Control de las células locales envía confirmación al Servidor de Gestión Central de Vídeo bajo Demanda. En caso de que un contenido no haya sido correctamente distribuido, el Servidor de Gestión Central de Vídeo bajo Demanda recibe la orden de reenviar el contenido. Si una vez en los servidores de vídeo de las células locales resulta corrompido, el Sistema de Gestión y Control de las mismas solicita reenvío.

El Servidor de Gestión Central de Vídeo bajo Demanda almacena también la información relativa a la caducidad de los contenidos. Cuando vence el periodo de

vigencia que la plataforma ha negociado con el suministrador del contenido, ordena el borrado del contenido en los servidores centrales y locales de vídeo bajo demanda.

3.1.4. Remultiplexor IP.

En el frecuente caso de que el operador de IPTV haya optado por codificar el vídeo en MPEG-2, cada uno de los canales que se distribuyen a través de la red de distribución de TV estará formado por una trama de transporte MPEG-2 monoprograma que se encapsula sobre IP utilizando el protocolo UDP/ multicast.

El remultiplexor IP asigna una dirección IP (multicast) a cada uno de los canales de TV y audio. El usuario final selecciona qué canal recibe a través del dial mediante el protocolo IGMP (join, leave)

3.1.5. Servidor de Arranque.

Cada vez que un usuario enciende su STB, el servidor de arranque proporciona el software de arranque al STB, de modo que el firmware del STB esté siempre actualizado.

3.2. Célula de Servicios Central.

La Célula de Servicios Central monitoriza y controla el comportamiento de toda la plataforma, suministrando la aplicación de navegación propia del operador, provisionando los equipos en la red y recogiendo información para la tarificación de los servicios. La componen el Centro de Servicios Central y el Centro de Gestión Central.

3.2.1. Centro de Servicios Central.

El Centro de Servicios Central realiza funciones de gestión de abonados y proporciona la aplicación de navegación del operador. Está compuesto por la Base de Datos de Abonados, el Servidor de Aplicaciones y la Base de Datos de Contenidos.

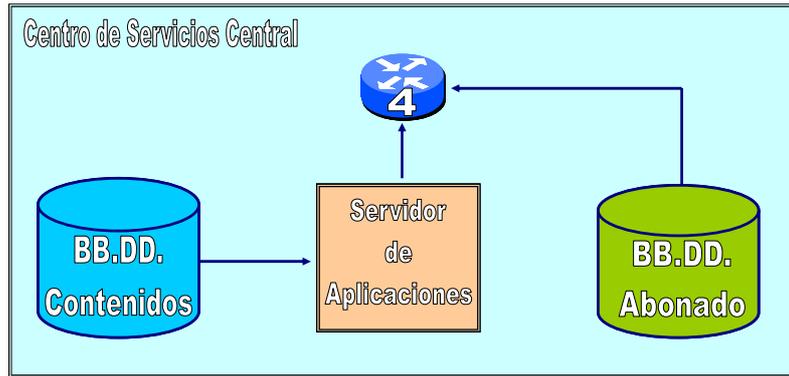


Ilustración [7] Centro de Servicios Central

3.2.1.1. Base de Datos de Abonados

En la Base de Datos de Abonados se almacena información particular de cada abonado, y permite:

- Autenticar al usuario (dar acceso a la plataforma).
- Autorizar el acceso a contenidos bajo demanda (comprobación del PIN de compra y autorización DRM, Digital Rights Management)
- Almacenar los datos de las compras en el Registro de Datos de Compras.

3.2.1.2. Servidor de aplicaciones

El servidor de aplicaciones envía a cada usuario la aplicación de navegación cada vez que el usuario se conecta a la plataforma, así como la información de la Guía Electrónica de Programación generada en la Base de Datos de Contenidos.

3.2.1.3. Base de Datos de contenidos

La Base de Datos de contenidos tiene por objeto:

- Almacenar la información referente a los servicios (canales TV, radio, etc) ofrecidos por los proveedores, así como su programación
- Definir de Paquetes de Canales de TV por demarcaciones
 - Canales Nacionales: Se ofrecen a todas las ciudades y con la misma programación en todas ellas.
 - Canales Nacionales con desconexiones locales: Se ofrecen a todas las ciudades pero con diferente programación en ciertos tramos horarios.
 - Canales Autonómicos: Sólo se ofrecen en un conjunto reducido.

- Definición de Canales de PPV
- Generar los ficheros de programación de los Canales de TV para la Guía y Miniguía
- Carga de logos de Canales.

3.2.2. Centro de Gestión Central

El Centro de Gestión Central está conformado por los siguientes núcleos:

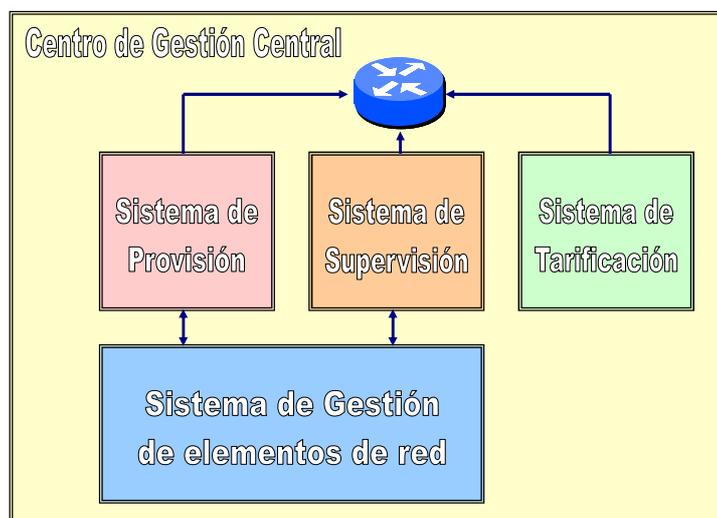


Ilustración [8] Centro de Gestión Central

3.2.2.1. Sistema de Provisión de Red de Acceso

El Sistema de Provisión de Red de Acceso establece o libera el circuito que permite el acceso a la plataforma entre el DSLAM de acceso y la ubicación de abonado. Da acceso a la aplicación de navegación en función de los servicios que el usuario haya suscrito.

3.2.2.2. Sistema de Supervisión

El Sistema de Supervisión recoge el servicio que el abonado suscribe para que el Sistema de Gestión pueda comprobar que realmente coincide con el servicio asignado.

3.2.2.3. Sistema de Gestión de elementos de red.

El Sistema de Gestión de elementos de red recoge las peticiones de provisión de nuevos servicios contratados, así como movimientos de altas, bajas, etc. Envía

las órdenes de provisión al Sistema de Provisión de Red de Acceso y comprueba la correcta asignación del servicio.

3.2.2.4. Sistema de Tarificación

El Sistema de Tarificación se encarga de emitir una factura con los servicios contratados y todos los consumos VoD y PPV realizados.

3.3. Red de Transporte

La red de transporte de un operador de IPTV está soportada sobre redes tanto a nivel físico ^[16] como a nivel lógico. A nivel físico, se distinguen las siguientes redes:

- Red de distribución de TV.
- Red multipropósito.

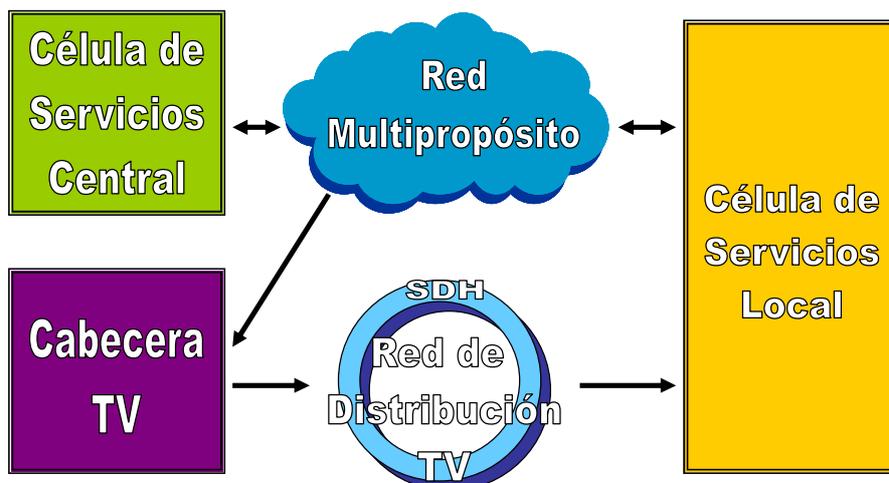


Ilustración [9] Cabecera de TV

Sobre estas redes se implementan otras a nivel lógico:

- Red de transporte de vídeo.
- Red de transporte de datos.
- Red de gestión.
- Red de distribución de contenidos.
- Red de transporte de autorización.
- Red de backup.

3.3.1. Red de transporte de vídeo.

La red de transporte de vídeo distribuye los canales de TV, desde la Cabecera de TV hasta las células locales, donde serán accesibles por los usuarios (multicast). Debido a la cantidad de información que es necesario enviar (un canal de TV codificado en MPEG-2 tiene una tasa de bit típica en torno a 4.5 Mbps), esta red es soportada sobre un anillo SDH.

La red de transporte de vídeo también distribuye el software de arranque de los descodificadores.

3.3.2. Red de transporte de datos.

Esta red proporciona comunicación entre la Célula de Servicios Central y los nodos de acceso. Es una red lógica a través de la red multipropósito.

3.3.3. Red de gestión.

La red de gestión comunica el Centro de Gestión Central con los diferentes elementos de la plataforma que deben ser gestionados o supervisados.

Esta red se realiza a través de la red multipropósito mediante túneles para garantizar la privacidad.

3.3.4. Red de distribución de contenidos.

Es una red unidireccional entre el Servidor Central de Contenidos bajo Demanda y las células locales para la distribución de contenidos de audio y vídeo. Tiene sentido descendente a través de un anillo unidireccional SDH (multicast), y el retorno se realiza a través de la red multipropósito mediante túneles GRE.

3.3.5. Red de transporte de autorización

La red de transporte de autorización conduce el flujo bidireccional de información de autorización entre las células de servicio locales y el Centro de Servicios Central. Es una red lógica a través de la red multipropósito.

3.3.6. Red de backup

Es una red dedicada al backup de los sistemas de la plataforma que comunica la célula central de gestión con la célula central de servicios, la cabecera de TV y las células locales.

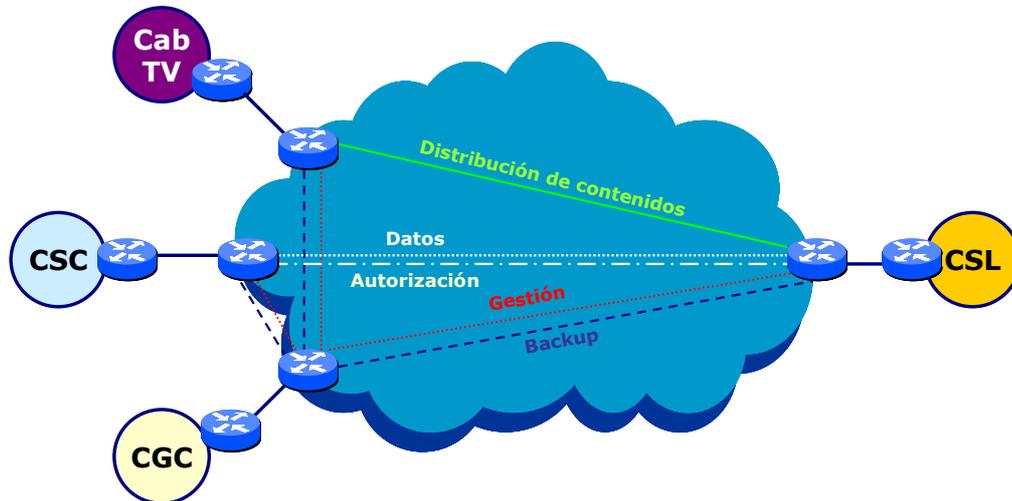


Ilustración [10] Distribución de VLANs en la red multipropósito.

3.4. Célula de servicios local

En la célula de servicios local, se encuentran los siguientes elementos:

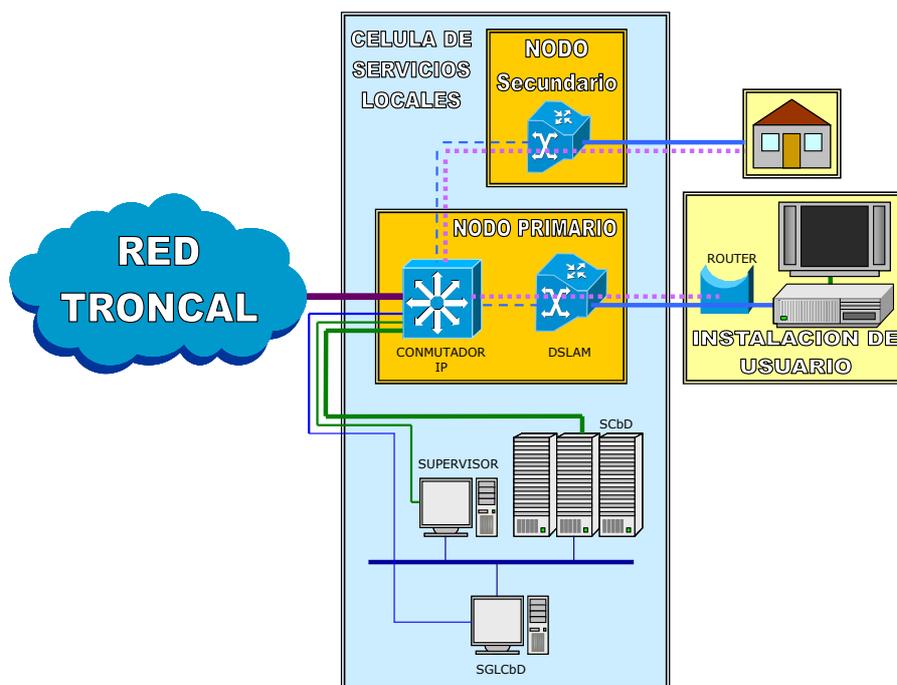


Ilustración [11] Célula de Servicios Local

- Servidores de Contenidos bajo Demanda: almacenan los contenidos que el usuario puede solicitar bajo demanda.
- Sistema de gestión y control: Servidor de Gestión Local de Contenidos bajo Demanda y Supervisor de TV.
- Nodos de acceso: punto de acceso de los usuarios a la red.

3.4.1. Servidores de Contenidos bajo Demanda

En una arquitectura IPTV distribuida, se inyecta el contenido bajo demanda una sola vez desde el servidor central del operador, y a través de la red de transporte (multicast), llega al Servidor de Gestión Local de Contenidos bajo demanda. Posteriormente este lo envía a los servidores de vídeo local de cada una de las Células de Servicios Local del operador, donde son accesibles a los usuarios. Los servidores locales de vídeo almacenan el contenido y se encargan de enviar el flujo de información a los usuarios que lo soliciten.

Cuando un usuario solicita un contenido bajo demanda, tiene lugar el siguiente flujo:

- El usuario se encuentra navegando por la aplicación del operador de IPTV y solicita un contenido bajo demanda.
- En este momento, se le pide al usuario autenticación para permitir la compra del contenido (PIN de compra).
- Esta petición llega hasta la Base de Datos de Abonados, en la Célula de Servicios Central, que realiza una consulta para saber si el usuario tiene autorización para reproducir el contenido (comprueba el PIN de compra), e igualmente comprueba si el contenido ya figura en el Registro de Datos de Compras.
- En caso de que el PIN de compra sea correcto, la base de datos de abonado devuelve al servidor de gestión local de contenidos bajo demanda una orden para que facilite al usuario un acceso exclusivo al contenido (envío de la licencia DRM al usuario) y el servidor local de vídeo bajo demanda inicia la transmisión del contenido al usuario.
- Este movimiento queda también reflejado en la Base de Datos de Compras del núcleo central para su facturación.

La política que se usa para el acceso a contenidos bajo demanda es la de streaming, frente a la clásica descarga directa. Las ventajas del streaming frente a la descarga son principalmente dos:

Tiempo de reproducción

La descarga de un contenido implica que, hasta que el contenido no esté descargado en el disco duro local, no se puede empezar su reproducción. Por ejemplo, para un vídeo de 2 horas de duración, codificado en MPEG-2 a una tasa de 4 Mbps, y con una velocidad de descarga de 5 Mbps, tenemos que:

$$\text{Tamaño del vídeo: } 2 \text{ horas} * 3600 \frac{\text{s}}{\text{hora}} * 4 \text{ Mbps} = 28800 \text{ Mb}$$

$$\text{Tiempo de descarga: } \frac{28800 \text{ Mb}}{5 \text{ Mbps}} = 5760 \text{ s} = 1 \text{ hora y } 36 \text{ minutos}$$

Necesitaríamos 1 hora y 36 minutos para la descarga de un vídeo completo, lo cual es inaceptable si lo que buscamos es la reproducción en el momento.

En el streaming, únicamente necesitamos descargar un mínimo buffer para asegurar la reproducción continua aunque existan pequeñas fluctuaciones en la velocidad de la red, y el contenido se va reproduciendo a medida que se va descargando, pasados los segundos iniciales en los que el buffer se llena.

Seguridad del contenido

Mediante la descarga directa, el contenido queda almacenado en el disco duro local, y podría ser manipulado por el usuario final, lo cual va en contra del modelo de negocio del operador y de la política de privacidad del suministrador del contenido. Mediante el streaming, el contenido es eliminado del buffer de reproducción a medida que se visualiza, garantizando que no puedan hacerse copias ilegales del mismo y dando sentido al modelo de cobro por visualización del contenido bajo demanda.

A diferencia de la distribución de programación multicast desde la cabecera central de TV, el flujo que se ofrece al usuario que solicita un contenido VoD es unicast. Los servidores de vídeo bajo demanda se encargan del streaming del contenido hasta el usuario.

Por lo tanto, los servidores de vídeo realizan las funciones de almacenamiento y streaming del contenido. En el dimensionamiento de los servidores deberán tenerse en cuenta estas dos funciones.

Para asegurar la integridad de los contenidos en los núcleos de servicios local, suele optarse por disponer de varios discos que actúen a modo de copia de seguridad (o al menos paridad) de los contenidos bajo demanda. Lo habitual es montar los servidores de vídeo en configuración RAID (Redundant Array of Independent Disks). La relación entre el número de discos de datos y discos dedicados a copia de seguridad es el RSS (Raid Set Size).

3.4.1.1. Configuración en RAID

RAID es la creación de un gran disco lógico a partir de discos individuales, repartiendo la información de manera homogénea.

Los servidores de Vídeo bajo Demanda requieren una gran capacidad de acceso a los contenidos. Dado que existen riesgos por cuellos de botella (pocos contenidos con muchas peticiones), los contenidos se trocean y se distribuyen por los discos homogéneamente. De este modo se logran mayores tasas de entrada y salida a los discos y balanceo de cargas (repartición de peticiones entre los discos).

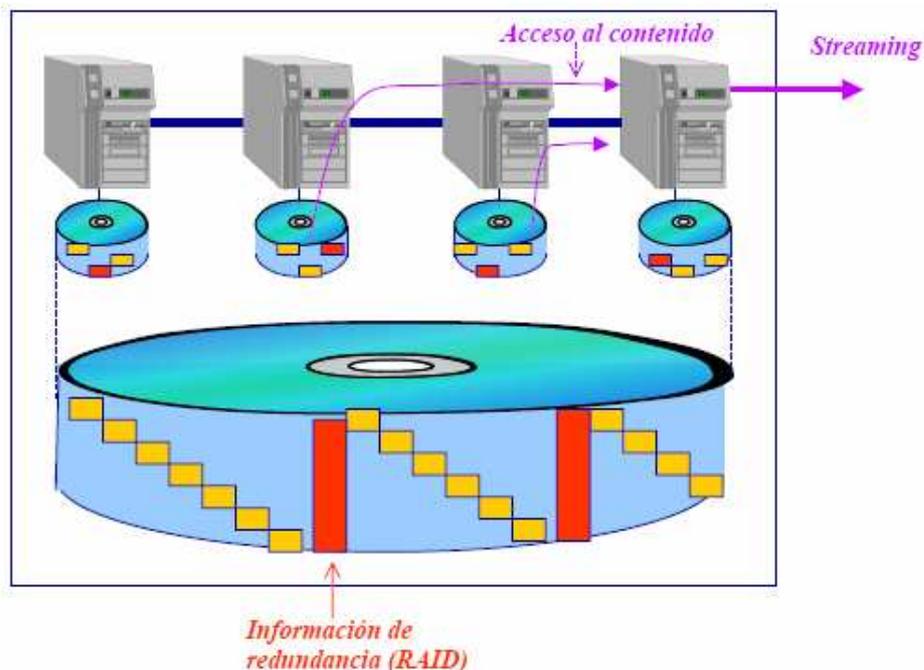


Ilustración [12] Configuración en RAID

3.4.1.2. Niveles RAID

3.4.1.2.1. RAID 1

Un RAID 1 crea una copia exacta de un conjunto de datos en dos o más discos (array). Esto resulta útil cuando el rendimiento en lectura es más importante que la capacidad y también desde el punto de vista de la seguridad, ya que tolera el fallo en uno de los discos, al disponer de la misma información en otro disco.

Un RAID 1 clásico consiste en dos discos en espejo, lo que incrementa exponencialmente la fiabilidad respecto a un solo disco; es decir, la probabilidad de fallo del conjunto es igual al producto de las probabilidades de fallo de cada uno de los discos (pues para que el conjunto falle es necesario que lo hagan todos sus discos).

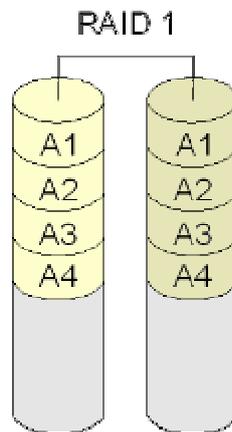


Ilustración [13] Configuración en RAID RSS 1

Adicionalmente, dado que todos los datos están en dos o más discos, con hardware habitualmente independiente, el rendimiento de lectura se incrementa aproximadamente como múltiplo lineal del número de copias; es decir, un RAID 1 puede estar leyendo simultáneamente dos datos diferentes en dos discos diferentes, por lo que su rendimiento se duplica.

3.4.1.2.2. RAID 2

Un RAID 2 divide los datos a nivel de bits en lugar de a nivel de bloques y usa un código de Hamming para la corrección de errores. Los discos son sincronizados por la controladora para funcionar al unísono. Éste es el único nivel RAID original que actualmente no se usa. Permite tasas de transferencias extremadamente altas.

3.4.1.2.3. RAID 3

Un RAID 3 usa división a nivel de bytes con un disco de paridad dedicado. Uno de sus efectos secundarios es que normalmente no puede atender varias peticiones simultáneas, debido a que por definición cualquier simple bloque de datos se dividirá por todos los miembros del conjunto, residiendo la misma dirección dentro de cada uno de ellos. Así, cualquier operación de lectura o escritura exige activar todos los discos del conjunto.

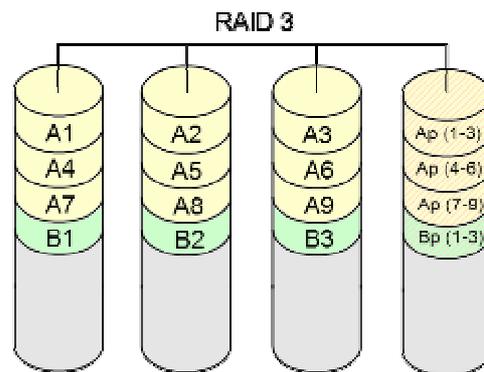


Ilustración [14] Configuración en RAID RSS 3

3.4.1.2.4. RAID 4

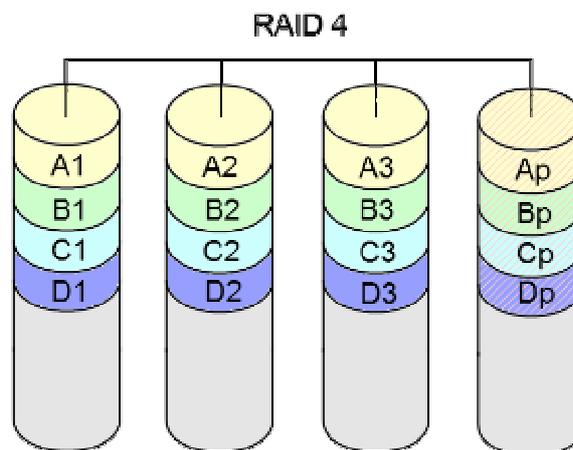


Ilustración [15] Configuración en RAID RSS 4

Un RAID 4 usa división a nivel de bloques con un disco de paridad dedicado. El RAID 4 es parecido al RAID 3 excepto porque divide a nivel de bloques en lugar de a nivel de bytes. Esto permite que cada miembro del conjunto funcione independientemente cuando se solicita un único bloque. Si la controladora de

disco lo permite, un conjunto RAID 4 puede servir varias peticiones de lectura simultáneamente.

La configuración más utilizada en IPTV es la configuración en RAID 4 (RSS 4, por cada 3 discos de datos, un disco de paridad).

3.4.1.3. Dimensionamiento de servidores locales de VoD.

Supongamos por ejemplo que un operador de IPTV necesita cubrir las siguientes necesidades:

- Ofrecer contenidos bajo demanda a un máximo de 500 usuarios al mismo tiempo (500 streams), por célula de servicios local.
- Disponer de 2500 horas de almacenamiento de contenidos bajo demanda (codificado en MPEG-2, a una velocidad de 4,5 Mbps), por célula de servicios local.

El operador ha elegido para sus servidores de vídeo bajo demanda un servidor de vídeo con una serie de discos duros que tienen las siguientes características:

- Capacidad máxima de bombeo streaming: 192 Mbps.
- Capacidad máxima de almacenamiento: 300 GB.

El operador debe evaluar la cantidad de discos duros de estas características que necesita instalar para satisfacer su necesidad de almacenamiento y streaming.

En base a los requerimientos de ancho de banda para streaming:

$$500 \text{ streams} \times 4,5 \frac{\text{Mbps}}{\text{stream}} = 2250 \text{ Mbps necesarios para streaming}$$

$$\frac{2250 \text{ Mbps}}{192 \frac{\text{Mbps}}{\text{disco}}} = 11,71 \rightarrow 12 \text{ discos necesarios}$$

Este error en exceso de capacidad de streaming se utiliza para garantizar la posibilidad de enviar el stream desde el Servidor Central de Contenidos bajo Demanda.

En base a la capacidad de almacenamiento necesaria:

$$4,5 \text{ Mbps} \times \frac{3600 \text{ s}}{\text{hora}} \times \frac{1 \text{ byte}}{8 \text{ bits}} = 2025 \text{ MB} = 1,98 \text{ GB por hora}$$

$$2500 \text{ horas} \times 1,98 \frac{\text{GB}}{\text{hora}} = 4950 \text{ GB de capacidad necesarios}$$

$$\frac{4950}{300 \frac{\text{GB}}{\text{disco}}} = 16,5 \rightarrow 17 \text{ discos necesarios}$$

Vemos que la condición más restrictiva es la de almacenamiento, que necesita 17 discos de datos, frente a los 12 discos de datos necesarios para streaming.

Este cálculo ideal resulta quedarse lejos de lo necesario, ya que no podemos disponer el 100% de la capacidad de los discos para almacenamiento:

- Cada disco duro necesita reservar aproximadamente un 1% para el sistema operativo.
- Los archivos de paginación (tricky play) requieren espacio un espacio en disco que depende de la velocidad de tricky.

Los archivos de tricky play almacenan las imágenes Intra del archivo de vídeo codificado en formato MPEG, y son las imágenes de referencia que se muestran cuando el usuario realiza un avance rápido o un rebobinado. Para permitir velocidades de reproducción de 8x y 16x (típicas en VoD), debe elegirse dicho valor como distancia entre imágenes I. Este proceso ha tenido lugar previamente en el Servidor de Gestión Central de Vídeo bajo Demanda.

Por tanto, para generar un tricky de Fast Forward a 8x es necesario reservar 1/8 extra del tamaño de cada archivo de vídeo. Del mismo modo, para generar un tricky de Fast Forward a 16x es necesario un extra de 1/16 del tamaño del archivo de video. Lo mismo ocurre para Rewind a 8x y 16x.

- Es necesaria una configuración en RAID que aporte seguridad.
- Debería reservarse en torno a un 25% para ampliación de contenidos.

En la siguiente tabla se resumen estos hechos:

Discos necesarios	17	Discos
Capacidad de Almacenamiento neta	5199	GB

Reserva de 1% para el S.O.	51,99	GB
----------------------------	-------	----

Tricky plays		
Fast Forward 8x	649,875	GB
Fast Forward 16x	324,9375	GB
Rewind 8x	649,875	GB
Rewind 16x	324,9375	GB

25% para ampliación	1299,75	GB
---------------------	---------	----

Capacidad de Almacenamiento necesaria	8517,365	GB
Número de discos necesarios	29	discos

Configuración en RAID RSS 4	38,66667	Discos
Configuración en RAID RSS 4 válida	40	Discos

3.4.1.4. El problema del envío

El servicio de Vídeo bajo Demanda requiere que a cada cliente se le envíe un flujo de vídeo personalizado, que podrá controlar a voluntad. Como ya se ha explicado, los operadores de IPTV lo realizan mediante el procedimiento de streaming.

El streaming consiste en el envío de flujos de datos, pero con ciertas particularidades. En el caso del vídeo, el streaming supone que el flujo de datos que se envía al cliente cumple una condición: la reproducción es inmediata. Esto conlleva que el vídeo no se almacena en ningún dispositivo, sino que tal y como se recibe es decodificado y reproducido. En el caso de que el cliente solicite un avance rápido o un rebobinado, lo que se producirá es una modificación del modo en que se envían las tramas de vídeo en el streaming, pero no se realizará ningún mecanismo local que implemente estas facilidades. El streaming implica que los equipos de cliente sean elementos pasivos, estando la inteligencia en el equipo del operador.

Este modo de funcionamiento tiene una consecuencia muy importante en las exigencias de red, especialmente en las relativas al retardo. Existe un tipo de retardo que afecta mucho a la calidad de streaming. No nos referimos al retardo que se mantiene constante en todo el flujo sino al retardo de unos paquetes respecto de otros, el jitter, y que hace que en un mismo flujo de datos la distancia entre paquetes sea variable. En este tipo de servicios el jitter es intolerable, ya que los equipos terminales disponen de buffers muy pequeños y las velocidades de transmisión son muy altas. Por ello, las redes sobre las que se prestan deben evitarlo a toda costa. Teniendo en cuenta que los flujos con los que se trabaja en este servicio tienen elevadas tasas binarias, este requisito implica una gran exigencia para las redes de transporte.

El streaming por si solo no implica que se envíen flujos personalizados, de hecho, se puede establecer servicios de difusión de vídeo sobre streaming sin necesidad de que se establezca un flujo de datos por cliente. La personalización es una condición exclusiva de los servicios de vídeo bajo demanda, a la cual se adapta el streaming en esos casos. Los servicios sobre streaming necesitarán una arquitectura que permita establecer los flujos necesarios y controlarlos individualmente.

3.4.1.5. Protocolos para transmisión de flujos de vídeo.

Los protocolos utilizados en el streaming son los siguientes:

3.4.1.5.1. Protocolos de nivel de enlace

Comúnmente, el acceso a los servidores de vídeo se realiza a través de fibra óptica. La tendencia es la evolución de ATM hacia IP, por lo que el protocolo dominante en el acceso a los servidores es Gigabit Ethernet.

3.4.1.5.2. IP (Internet Protocol).

El IP es un protocolo situado en la capa de red del modelo OSI, basado en la conmutación de paquetes y que no está orientado a conexión. Sobre el protocolo IP normalmente se utilizan dos protocolos de control de transmisión: TCP (Protocolo de control de transmisión) y UDP (Protocolo de datagramas).

3.4.1.5.3. UDP y TCP.

UDP es muy utilizado para el transporte de tráfico multimedia en tiempo real. Gracias a su sencillez y su escaso sobrepeso en cabeceras de control y

procedimientos está muy indicado para tipos de tráfico en los que lo más importantes es que los paquetes lleguen a tiempo.

Pese a ser un protocolo fiable, TCP no permite garantizar una calidad de servicio a los usuarios finales, ni permite la reserva de ancho de banda que garantice la transmisión del flujo de datos durante el periodo de visualización de contenidos multimedia. Estas características limitan considerablemente la aplicación de este protocolo para soportar servicios de VoD.

3.4.1.5.4. RTP (Real Time Protocol)

Creado por el IETF, este protocolo de nivel de aplicación, como su nombre indica, está diseñado para tráfico de tiempo real, típicamente audio y vídeo. Por lo tanto, no está orientado específicamente a los datos multimedia ni a un estándar concreto.

El objetivo de RTP es proveer servicios de transporte extremo a extremo al tráfico de tiempo real. Se puede utilizar sobre cualquier protocolo de transporte o red y el propio protocolo se encarga de aislar el tráfico que transporta de todos los niveles inferiores. El tráfico encapsulado en RTP es muy adecuado para las aplicaciones multimedia, ya que añade características como marca de tiempo y numeración de secuencia de los mensajes e identificación del tipo de datos transmitidos. Pese a esto, RTP no asegura la entrega ordenada ni la seguridad de las transmisiones.

3.4.1.5.5. RTCP

Mediante un protocolo adicional, el RTCP (Real Time Control Protocol), se implementa la monitorización de las sesiones: control de la calidad de servicio y control de flujo y congestión. Este protocolo genera la transmisión de informes estadísticos entre el transmisor y el receptor en el protocolo RTP, mediante los cuales se identifica el estado de congestión de la red y que consiguen limitar el número de paquetes perdidos (ajuste automático de ancho de banda para que la velocidad de reproducción sea la misma que la de descarga, y no se vacíe ni rebose el buffer local).

La mayoría de operadores de IPTV optan por no implementar este protocolo para el VoD debido a las características de la red sobre la que se transportan los flujos de streaming. La implantación de una arquitectura distribuida hace la red fiable, ya que los servidores están cerca de los conmutadores de acceso, por lo

que no merece la pena utilizar un protocolo como RTP, orientado a redes no fiables.

3.4.1.5.6. RTSP (Real Time Streaming Protocol)

Este protocolo de nivel de aplicación permite controlar la transmisión de tráfico de tiempo real por parte de un cliente. El tipo de tráfico que se suele controlar con este protocolo es audio y/o vídeo, y se usa especialmente en aplicaciones multimedia.

En una sesión de streaming, RTSP sólo se ocuparía del control; en cambio, el transporte de los flujos de datos multimedia se realizaría mediante otro protocolo, típicamente RTP o Trama de Transporte MPEG-2, ya sea sobre UDP o TCP. Los mensajes RTSP tampoco llevan asociado un protocolo concreto de transporte.

Mediante una sesión RTSP el cliente puede controlar uno o varios flujos multimedia que estén asociados (audio y vídeo de una película, por ejemplo). El control que el cliente puede ejercer sobre la reproducción simula el de un aparato de vídeo doméstico, es decir, mediante RTSP se puede iniciar la reproducción, parar, avanzar o retroceder a mayor velocidad, situarse en un punto concreto y abortar la reproducción. Cada aplicación puede implementar estas acciones u otras diferentes.

Aplicaciones multimedia muy difundidas como Windows Media Player o RealPlayer utilizan RTSP para los diálogos entre los clientes y los servidores.

3.4.1.5.7. MPEG-2 Transport Stream

Este es el encapsulado de transporte por defecto del estándar de vídeo MPEG-2. Permite agrupar distintos flujos en uno sólo, lo que se puede utilizar, por ejemplo, para multiplexar varios canales de audio en un mismo flujo, de modo que se puedan ofrecer contenidos en dual.

Esto no es una capa de transporte en sí misma, por lo que no implementa funcionalidades de control de flujo ni similares. Puede transmitirse sobre RTP o directamente sobre TCP o UDP.

Por tanto, para garantizar el correcto funcionamiento del servicio de vídeo bajo demanda debe implementarse una torre de protocolos similar a la que se muestra a continuación:

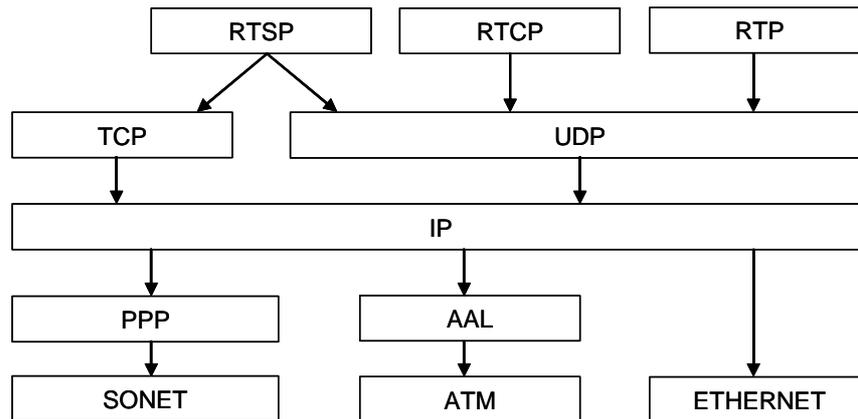


Ilustración [16] Torre de Protocolos

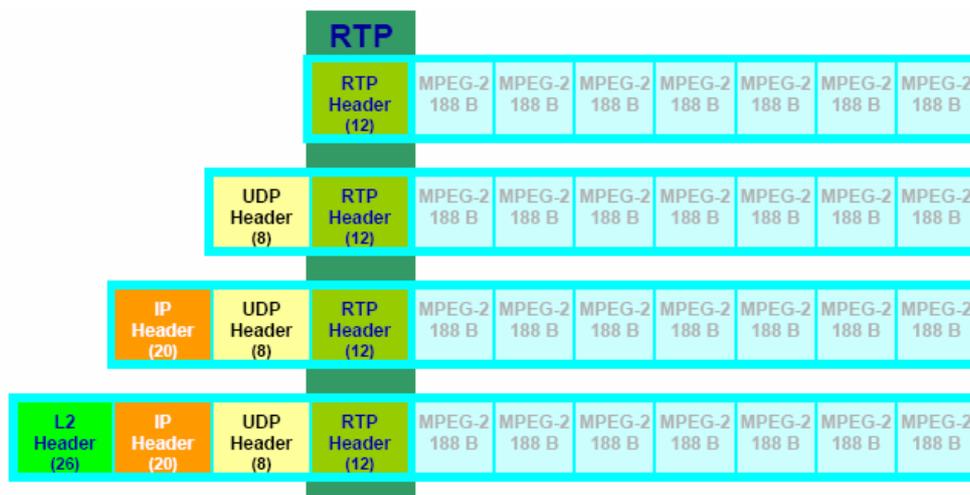


Ilustración [17] Protocolos para la transmisión de vídeo. Formación de tramas.

3.4.2. Sistema de gestión y control

El sistema de gestión y control de la célula de servicios local está formado por:

- El Servidor de Gestión Local de Contenidos bajo Demanda.
- El Supervisor de TV.

El Servidor de Gestión Local de Contenidos bajo demanda recibe los contenidos que le llegan desde el servidor central situado en la Cabecera de TV. El es el que se encarga de procurar el almacenamiento equitativo en los servidores locales. (balanceo de cargas)

El Servidor de Gestión Local de Contenidos bajo Demanda recoge igualmente las peticiones del usuario de contenidos bajo demanda, y las envía a la Base de Datos de Abonado. Si el usuario ha sido autenticado correctamente (PIN de compra), devuelve la licencia DRM al Servidor de Gestión Local de Contenidos bajo

Demanda, y este permite el acceso enviando dicha licencia al usuario. Esta clave viaja cifrada por la red, y permanece en la memoria del STB hasta que termina la visualización del contenido.

El supervisor de TV es un terminal situado en la célula de servicios local, que monitoriza constantemente el estado de las recepciones desde el núcleo central. Este terminal realiza peticiones de canales y los va visualizando, obteniendo información sobre la calidad de la red.

3.4.3. Nodo de acceso.

Corresponde al punto de conexión con el abonado. Este se puede conectar a la red directamente a través de la central, o mediante un nodo remoto. Este nodo no es más que un DSLAM (Digital Subscriber Line Access Multiplexer), que proporciona a los abonados acceso a los servicios DSL sobre cable de par trenzado de cobre. Igualmente, los usuarios que se conectan directamente a la central, encuentran a la entrada de esta otro DSLAM. La comunicación del DSLAM y el Router xDSL se realiza a través de dos interfaces llamadas ATU-R del lado del cliente y ATU-C del lado del proveedor del servicio.

3.5. Ubicación de Abonado

Los operadores de IPTV no sólo ofrecen a los abonados servicios de vídeo y audio a través de una red IP, sino que completan la oferta con servicios de Triple Play ^[2]: televisión, telefonía e Internet, con tarifas distintas en función de la combinación escogida. Por esto, en la ubicación de abonado solemos encontrar los siguientes elementos:

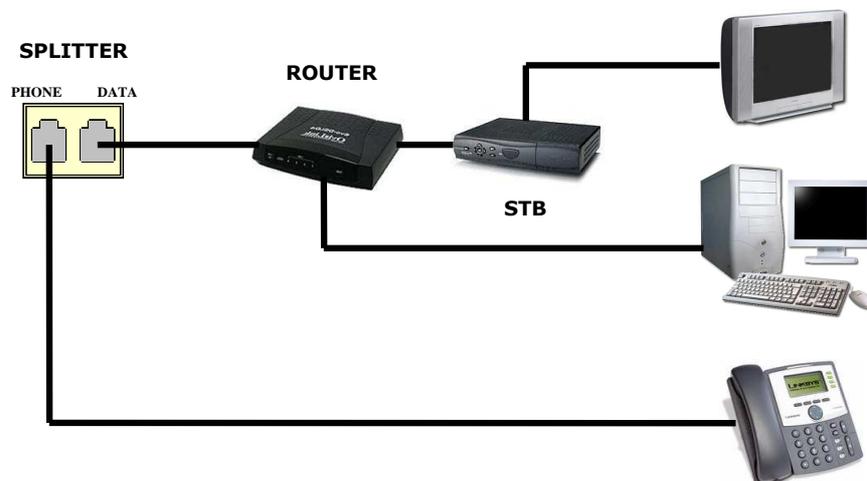


Ilustración [18] Ubicación de abonado

Actualmente, la evolución es hacia los servicios de Quadruple Play, esto es, la integración de la telefonía móvil con los servicios anteriormente mencionados. De todos ellos, el equipamiento que resulta de interés en los servicios de IPTV es el Set-Top-Box.

Los usuarios deben soportar la recepción y visualización sin cortes de los contenidos de TV y audio, así como soportar los comandos clásicos de VCR para VbD. En general, los STB constan de cuatro componentes principales:

3.5.1. Interfaz de red

Permite al cliente recibir y enviar información desde o hacia la red.

3.5.2. Decodificador

Como ya se ha comentado anteriormente, para reducir los requisitos de almacenamiento y ancho de banda, los contenidos están codificados. Así, se necesita un decodificador en la ubicación de abonado para decodificar el vídeo y audio, antes de ser presentado al usuario.

3.5.3. Buffer

Debido a los retrasos introducidos por la red, el tiempo de llegada de la información no puede ser determinado con exactitud. Para conseguir una reproducción fluida sin cortes, debe garantizarse que la siguiente porción de vídeo o audio está disponible antes de que el usuario la quiera reproducir. Para lograr este objetivo, se envía al usuario en adelanto, de forma que se asegure un margen de maniobra que amortigüe los posibles retardos introducidos por la red de comunicación.

3.5.4. Hardware de sincronización

Los contenidos de TV están compuestos por un stream de vídeo y un stream de audio independientes. Para poder realizar una reproducción correcta, ambos tipos de flujos deben ser sincronizados entre sí antes de que puedan ser reproducidos.

El desarrollo de los STB mantiene una continua evolución, no sólo enfocada a reducir su coste, sino también a incrementar su potencia. La tendencia es sobrepasar el mero rol de receptor y decodificador de vídeo, convirtiéndolo en un verdadero centro de entretenimiento familiar, e incrementando su capacidad de almacenamiento y procesamiento.

Capítulo II.

DESCRIPCIÓN DE LOS SERVICIOS DE GRABACIÓN DE CONTENIDOS.

4. GRABACIÓN DE CONTENIDOS

En este capítulo se describirán los dos nuevos servicios de grabación de contenidos que un operador de IPTV podría añadir a su oferta de servicios, objeto de este Proyecto Fin de Carrera. A continuación, también se describirán las modificaciones necesarias en el equipamiento de cliente y en la arquitectura de red para dar soporte a los servicios de grabación de contenidos.

Los servicios de grabación de contenidos se presentarán en este Proyecto como un único servicio de suscripción mensual.

4.1. *Time Shifting*

El servicio de Time Shifting ofrece la posibilidad de detener la emisión en curso de un programa, simplemente pulsando el botón de Pausa del mando a distancia del STB. Desde este momento, se empieza a almacenar en el disco duro del STB la emisión del contenido que se estuviera visualizando, mientras en la pantalla lo que aparece es la imagen del vídeo congelada en el momento en el que se pulsó el botón de Pausa. El usuario puede posteriormente pulsar el botón de Play de su mando a distancia, y seguir visualizando el contenido por donde lo dejó, mientras el STB continua grabando la emisión en curso, para no perder nada. El usuario tiene la sensación de poder detener el tiempo a su antojo.

Existe la posibilidad de implementar un Time Shifting básico o un Time Shifting más avanzado. La mejora de este último consiste en poder rebobinar la emisión sin haber activado de manera explícita la función de Time Shifting, esto es, que el propio STB esté ejecutando un Time Shifting continuo. Ambos modos de operación se describirán a continuación.

4.1.1. Modos de operación.

Las posibilidades que ofrece el servicio de Time Shifting son:

4.1.1.1. Pausa de la emisión

Pulsando el botón de Pausa del mando a distancia, la imagen de la emisión en vivo se congela, y el STB empieza a grabar el canal de TV en cuestión. Debe indicarse en la pantalla que el Time Shifting está activo, mediante un icono de pausa, así como un contador de tiempo de grabación. El STB continua grabando la emisión en curso, de forma transparente al usuario, para que pueda visionarla

más tarde. Es necesario indicar en la pantalla que se está haciendo uso del Time Shifting para evitar dudas o reclamaciones de usuarios (podrían olvidar que están haciendo uso de esta funcionalidad y creer que el servicio tiene algún tipo de problema que provoca la congelación de la imagen).



Ilustración [19] Pausa de la emisión en vivo

4.1.1.2. Reanudación de la emisión

Si la reproducción se encuentra en estado de pausa, con Time Shifting activo, y el usuario pulsa el botón Play del mando a distancia, la visualización del contenido se reanuda desde el lugar en el que se quedó. Si estamos en otro modo de operación (rebobinado, avance rápido o reproducción), pulsar el botón de Play no tendrá efecto alguno.

En la pantalla se indicará durante unos segundos cuanto tiempo de retraso existe respecto a la emisión en tiempo real. Se seguirá grabando la emisión en tiempo real para que el usuario pueda visionarla más tarde, a medida que avanza la reproducción.

4.1.1.3. Avance rápido

Se logra pulsando el botón de Fast Forward del mando a distancia. Durante el avance rápido, se informará al usuario de la velocidad de visionado y del tiempo de desfase que nos queda hasta la emisión en vivo. Igualmente, se mostrará un icono que identifique la acción de avance rápido.

Si mediante el avance rápido el usuario alcanza el momento de emisión "en vivo", se sigue con la emisión en curso.

Durante el avance rápido, la grabación del contenido en el disco duro del STB continúa, para que el usuario no se pierda ninguna parte de la emisión que está grabando.



Ilustración [20] Avance rápido

4.1.1.4. Retroceso rápido

La filosofía es la misma que para el avance rápido. Se logra pulsando el botón de retroceso rápido del mando a distancia, y se permitirá el retroceso hasta el momento en el que se inició el Time Shifting. Si el usuario alcanza este momento inicial, se conmuta a la función de pausa, deteniéndose el rebobinado y mostrándose en la pantalla la imagen congelada, así como los indicadores de Time Shifting activo y el retraso con respecto a la emisión en tiempo real, mientras el buffer de grabación sigue creciendo.

Del mismo modo que en el avance rápido, al retroceder en la emisión se informará al usuario de la velocidad del retroceso y del tiempo de desfase que existe con respecto a la emisión en curso. La emisión se sigue grabando mientras se hace retroceso rápido.

4.1.1.5. Salto al directo

Pulsando el botón de salto al directo, Stop, se borra el contenido grabado en Time Shifting y se vuelve a la emisión en curso. En este caso, se presentará al usuario un mensaje avisándole de que si realiza un salto al directo, perderá todo lo almacenado en Time Shifting. Este mensaje se deberá presentar en un cuadro de texto semitransparente, para no molestar la visualización. Una vez confirmada la acción, se borra el contenido de Time Shifting del disco duro y se sigue con la emisión en curso.

El disco duro continúa grabando mientras se presenta al usuario el mensaje de aviso, para que en caso de rechazar el salto al directo, no haya perdido partes del programa que está grabando.

4.1.1.6. Reanudación de la grabación

Si durante un visionado de un contenido grabado en Time Shifting el usuario vuelve a hacer pausa, se congelará la imagen y se mostrarán igualmente los indicadores de Time Shifting activo y el retraso con la emisión en tiempo real. El disco duro continúa grabando el contenido para su posterior reproducción.

4.1.1.7. Picture in Picture

Pulsando el botón de PiP, se detiene la visualización actual de Time Shifting (se conmuta al modo de operación pausa, el disco duro del STB sigue grabando), y se presenta en la pantalla, además de la imagen congelada en pausa, un recuadro con la emisión en curso. De este modo, el usuario podrá decidir, según lo que observe en este recuadro, realizar un salto al directo. Al pulsar los botones de play, fast forward, rewind, pause, salto al directo, PiP, menú, o al cambiar de canal, desaparece la ventana PiP.



Ilustración [21] Picture in Picture

La funcionalidad de Time Shifting no estará disponible en los casos en los que se muestre video enventanado desde el menú de navegación, PPV ó VoD. Para no permitir Time Shifting en los canales de PPV ó VoD se discriminará por rango de canales en el dial, impidiendo el Time Shifting en el rango reservado a PPV y VoD. El menú de navegación se considera a todos los efectos como un canal de televisión, y por tanto es posible discriminar la grabación en él.

4.1.2. Modo normal de funcionamiento. Time Shifting básico.

Por la reducción de la complejidad en los desarrollos que hay que realizar en la plataforma de IPTV, así como por la reducción de costes en equipamiento, el operador puede optar por implementar un Time Shifting que no permita la función de retroceso rápido sin haber realizado previamente un Pause explícito.

En este caso, si un usuario está haciendo Time Shifting en un canal de TV (ha pulsado el botón de Pausa), y cambia a otro canal distinto del actual, o pulsa la tecla para acceder al menú de navegación, la grabación realizada para el Time Shifting se perderá (se considera que el acceso al menú de la aplicación es a todos los efectos un cambio de canal). En esta situación el servicio dará un mensaje de notificación al usuario, indicando que está actualmente funcionando en Time Shifting y que si cambia de canal, perderá la grabación del desfase del canal correspondiente. Si el cliente confirma esta acción, se efectuará el cambio de canal. El mensaje se presentará en un formato y tipología poco intrusiva (semitransparente sobre la imagen o en un margen de la superficie de TV) para no molestar el visionado de la TV.

Asimismo, si un usuario está haciendo Time Shifting de un canal, y ocupa la cuota total reservada en el disco duro para Time Shifting, se procederá a ir eliminando el contenido más antiguo e ir grabando el contenido nuevo, sin previo aviso. De este modo, si el usuario dispone de 3 horas continuas de Time Shifting, sólo podrá recuperar las últimas 3 horas de emisión, aunque llevase más tiempo haciendo Time Shifting.

Para prevenir situaciones de olvido por parte del usuario durante una grabación y que esté siempre visualizando el canal de TV en modo diferido Time Shifting, una vez se haya grabado el máximo espacio de disco asignado a Time Shifting, la grabación se mantendrá durante un tiempo preconfigurado en el equipo desde fábrica y fijado por el operador, que en principio podría ser del tiempo dedicado a Time Shifting + 2 horas. Este valor se podrá modificar remotamente con la telecarga de software del descodificador.

Por tanto, y como ejemplo, si un usuario tiene configurada una cuota de Time Shifting de 3 horas, y llega a este máximo, el servicio irá manteniendo en el disco duro siempre las 3 últimas hora de emisión del canal. Si no hay ninguna actividad, el servicio irá grabando y borrando de manera cíclica (manteniendo 3 horas de grabación) durante 2 horas más, y después de este tiempo se borrará todo el contenido en Time Shifting y se volverá la TV a su emisión en vivo. En resumen,

tras cinco horas de inactividad, se desactiva la función de Time Shifting, sin previo aviso.

Es de vital importancia detener la funcionalidad de Time Shifting en caso de que no se esté haciendo uso de la misma. Si se está ejecutando Time Shifting de manera continua, el disco duro del STB se ve obligado a trabajar continuamente. Aquí aparece la problemática tecnológica de este servicio a nivel de equipamiento de cliente. El sobreesfuerzo del disco duro limita la vida útil del STB.

4.1.3. Modo de grabación continua. Time Shifting avanzado.

La otra modalidad de grabación permite la función de retroceso rápido, sin necesidad de haber realizado un Pause anterior. Existe la posibilidad de que el STB esté grabando constantemente el canal sintonizado y guardando en disco duro los 15 o 30 últimos minutos de emisión, de manera que el funcionamiento del Time Shifting sea ligeramente diferente a lo descrito anteriormente. Al hacer Pause sobre la emisión en curso no iniciamos la grabación, pues ya estaba en marcha, sino lo que hacemos es incrementar ese buffer de grabación del canal. Esto es lo que se conoce como grabación continua.

El usuario siempre podrá volver 15 o 30 minutos atrás de la emisión en vivo, como una facilidad adicional del servicio, sin necesidad de haber dado previamente al botón pausa, y reproducir así la última jugada de su equipo favorito o la mejor escena de la película. De este modo, se hace posible el hacer Rewind de un contenido, con las mismas velocidades de tricky play definidas para el Fast Forward, sin haber pulsado previamente el botón de activación de Time Shifting, "Pausa".

En caso de que el usuario cambie a un canal distinto o entrase en el menú de navegación, se borraría todo el contenido de Time Shifting del disco duro, esta vez sin previo aviso, pues al usuario le resultaría engorroso recibir un mensaje de notificación cada vez que cambie de canal. Cuando se permanezca un tiempo conveniente en el mismo canal, se entiende que el usuario está interesado en el contenido que está viendo, y se comienza la grabación de Time Shifting.

Por otra parte, si se realizase un Time Shifting explícito, esto es, pulsando el botón de "Pausa", si debe notificarse al usuario de que perderá el contenido almacenado en caso de querer cambiar de canal o entrar en el menú de navegación de la aplicación de IPTV.

Si un usuario de grabación continua pulse el botón play, habiendo hecho previamente un pause, este podrá siempre volver atrás y recuperar los últimos 15 o 30 minutos anteriores al momento en el que hizo el primer pause. En el modo de operación normal, sólo se permitía retroceso rápido hasta el momento en el que se hizo el primer pause.

El buffer no se borra cuando se alcanza el momento de emisión "en vivo". A partir de ese momento se anexa la emisión en curso al buffer existente. Si se alcanza la cuota de 3 horas destinada a Time Shifting, se continua grabando de manera cíclica, borrando lo más antiguo y almacenando la emisión en curso. Si se realiza un cambio de canal, se borra el buffer anterior, sin previo aviso, y se inicia un nuevo buffer que almacene los 15 o 30 minutos previos.

4.1.4. Implicaciones en el disco duro del STB

Si bien antes era importante limitar el tiempo que se está haciendo Time Shifting para no someter el disco duro a esfuerzos continuos, ahora resulta ser crítico. En la grabación continua, el disco duro es sometido a un esfuerzo continuado que reducirá mucho su vida útil. Existen varias soluciones, todas ellas de compromiso, para paliar este problema.

Por un lado, podría dotarse al STB de un ventilador y aperturas que refrigeren el disco duro, sin suponer un gran incremento en el coste final del mismo. El problema que aparece con esta solución es el ruido acústico, perceptible por el usuario, que introduce el ventilador.

Otra solución podría ser implementar el STB sobre un chasis metálico, dotado de rejillas o aperturas para la circulación del aire, que mantenga el dispositivo dentro de unos márgenes de temperatura adecuados. El problema es que el STB disiparía calor a través del chasis, calentándose el mismo y calentando todo su entorno en sobremanera.

También podría utilizarse un disco duro de calidad Premium de altas prestaciones, cuya vida útil cumpla los requisitos especificados por el operador para su servicio. Obviamente, el problema que aparece es la repercusión en costes.

En ambos modos de funcionamiento, normal y grabación continua, el operador podrá configurar en la aplicación del servicio el espacio de disco que quiere reservar para el Time Shifting del usuario. Este espacio, cuyo tamaño es configurable por el operador, está limitado en el disco duro, pues los contenidos

grabados en Time Shifting se almacenan en una partición especial más rápida, preparada para paginar archivos de tricky play. Como ya se ha comentado, estos archivos índice son los que se visualizan cuando se hace un Fast Forward a cualquier velocidad. El tamaño de disco duro del descodificador, el espacio máximo reservado para Time Shifting y el espacio de disco asignado a Time Shifting deben poder ser verificados remotamente por el operador para ofrecer un correcto soporte técnico. En ningún momento el usuario podrá tener acceso a esta partición especial del disco duro, ni podrá gestionar archivos ni carpetas en ella.

Por último, mencionar que si forzamos el STB al estado de stand-by, se detiene la función de Time Shifting y se elimina todo el contenido de Time Shifting almacenado en el disco duro.

4.2. Digital Video Recorder

Esta aplicación corresponde en términos generales a la funcionalidad de grabación de contenidos de TV que ya conocemos por los vídeos domésticos, los grabadores de vídeo en formato DVD u otros descodificadores de TDT preparados con disco duro ^[17]. Para el caso de operadores de IPTV, y por cuestiones de derechos legales sobre los contenidos, no siempre se permitirá grabar sobre todos los canales de TV ni sobre todos los contenidos de los mismos. En tal caso siempre se le mostrará al cliente un mensaje informativo indicando la imposibilidad de grabar dicho contenido por problemas de derechos legales ajenos al operador.

Cabe mencionar que el servicio expuesto aquí como DVR es popularmente más conocido como PVR, Personal Video Recorder (evitaremos en la medida de lo posible el uso de esta terminología pues PVR es una marca registrada de grabación de contenidos).

4.2.1. Modalidades de grabación

Nunca se permitirá grabar eventos de PPV ni contenidos de VbD, pues atentan contra el negocio del operador y contra las restricciones legales que el proveedor de contenidos pueda establecer. Del mismo modo que en el caso del Time Shifting, se anulará la posibilidad de grabar DVR en el rango de canales correspondiente al VoD y PPV. La grabación de un determinado programa de TV se podrá operar de tres maneras distintas.

4.2.1.1. Método de grabación impulsiva.

Una de las formas de grabación sería directamente pulsando el botón de REC del mando a distancia. De esta forma se recoge la información de dicho programa desde la EPG y se adopta la hora de finalización de grabación que indique la EPG para el programa en cuestión. Se informará por pantalla durante unos segundos del inicio de la grabación, además de la hora prevista de finalización de dicho programa. Se podrá parar la grabación pulsando en el mando a distancia el botón de Stop, y el contenido quedará almacenado en el disco duro del STB para su administración.



Ilustración [22] Grabación instantánea pulsando RECORD

La hora de finalización de la grabación podrá variar de manera dinámica según las actualizaciones que aparezcan en la EPG durante la grabación, por los motivos que sean, ya que el servicio se basará en la grabación de programas completos, y no en horas de inicio y fin.

Si pulsamos nuevamente el botón de grabación, se añade a la lista de grabación el programa siguiente, informando de la nueva hora de finalización de la grabación, y siempre que se disponga de derechos legales de grabación. Este comportamiento es limitado, no podemos seguir pulsando REC para ir añadiendo programas a la lista de grabación, debido a que es necesario consultar los derechos de grabación para cada canal y contenido. La información sobre el programa actual y el siguiente se encuentra en la miniguía de navegación. Para consultar información sobre programas posteriores, habría que hacer consultas a la base de datos de la EPG, y estas interrumpirían momentáneamente el servicio de TV (congelarían la imagen y el sonido). Una vez añadido a la lista de

programación el programa siguiente, si el usuario pulsa de nuevo el botón de REC, se elimina de la lista de grabación dicho programa, informando de este hecho.

4.2.1.2. Grabación programada desde la EPG.

Otra modalidad de grabación sería desde la EPG: Se podrá seleccionar el programa a grabar de los disponibles o permitidos, así como cancelar grabaciones previstas. Sólo se permitirá la programación de una grabación de un contenido en un determinado momento, no se permitirán dos grabaciones simultáneas. Los contenidos sobre los que no esté permitido grabar por problemas de derechos podrían aparecer difuminados en gris en la EPG.

Tanto en el caso en que el cliente quiera programar desde la EPG dos grabaciones simultáneas o quiera seleccionar un programa sin derechos, se le mostrará al cliente el mensaje de error o notificación correspondiente. Al finalizar una grabación programada, se elimina de la lista de grabaciones pendientes

4.2.1.3. Grabación por hora absoluta de inicio y fin.

La tercera y última modalidad de grabación sería la programación directa de la hora de inicio y fin de grabación. Como se puede hacer actualmente en los vídeos domésticos, también será posible que el cliente pueda programar las grabaciones desde una aplicación en la que se especifique hora de inicio y final, así como el canal de TV en cuestión. Si la grabación la programa el usuario directamente por hora de inicio y fin, se efectuará una consulta a los derechos de los contenidos marcados en dicha franja horaria, y si aparece como inviable por motivos de derechos, o porque está intentando solapar dos grabaciones simultáneas sobre la misma franja horaria, se le informará al usuario de este hecho, indicado el motivo, si es por no derechos sobre un programa determinado (indicado dicho programa) o si es por solape de dos programas, informando de los dos espacios solapados. De esta manera el usuario decidirá la acción más conveniente para solucionar el conflicto.

En cualquiera de las tres modalidades de grabación, si cambiamos de canal durante una grabación, ésta sigue transcurriendo, y a la vuelta al canal de grabación, se mostrará información de tiempo transcurrido y previsto de final durante unos segundos en la pantalla de TV.

4.2.2. Funcionalidades añadidas del servicio DVR

- **Control parental:** Se podrá programar un control parental de acceso por carpeta (de la librería personal) o por grabación de manera individual.
- **Funcionalidades de VbD:** El acceso y la interacción con los contenidos grabados seguirán la misma filosofía que los contenidos de VbD, permitiéndose facilidades de visualización del tipo pausa, avance, retroceso rápido, etc.
- **Programación de grabaciones periódicas:** existirá la posibilidad de habilitar una nueva facilidad de programación de grabaciones consistente en la selección de un determinado programa de la guía (cuya emisión pueda ser periódica, y atendiendo a criterios lógicos que identifiquen el programa, no mediante la hora de inicio y fin que podría variar), pudiendo programar su grabación por períodos de 1 semana, 1 mes o sin límite temporal. Esto permitiría grabar, por ejemplo, todos los capítulos de una serie de forma automática, y evitando grabar repeticiones o reposiciones de un mismo programa. Desde la EPG se podría obtener información para grabar también, por ejemplo, todas las películas de su actor favorito o programas deportivos sobre motor.
- **Permisión del funcionamiento del descodificador en "stand-by":** Cuando una grabación programada llegue a su hora de inicio, el descodificador empezará a grabar. En caso de que éste estuviera en "stand-by", se conectará automáticamente y comenzará a operar, y al finalizar la grabación, se desconectará si no se ha detectado ninguna actividad (mando a distancia) sobre él durante toda la grabación. Por el contrario, si el descodificador detecta alguna interacción con el usuario durante este tiempo, terminará de grabar pero no se irá a estado "stand-by" por sí solo.

4.3. Interoperabilidad de DVR y Time Shifting.

En todo momento, la función de DVR prevalecerá sobre la del Time Shifting. Por ejemplo, si un usuario se encuentra haciendo Time Shifting en un momento anterior a la emisión en curso y pulsa el botón REC para hacer una grabación, debe consultarse si se disponen de derechos para la grabación en curso. Si esto es así, el Time Shifting se desactivará, perdiéndose el contenido almacenado, y

además, se conmutará a la función de DVR programando la grabación de dicho espacio hasta su hora de finalización de acuerdo a la operativa ya especificada. En caso de no disfrutar de derechos para la grabación, se presentará al usuario un mensaje de información, y se proseguirá con el funcionamiento normal de Time Shifting.

Capítulo III.

ESTUDIO DEL IMPACTO DE LOS SERVICIOS DE GRABACIÓN.

5. IMPACTO EN EL SET-TOP-BOX

La introducción de nuevos servicios de grabación de contenidos en una plataforma de IPTV tiene repercusiones directas en los terminales de usuario. Es necesario cambiar los STB por otros STB avanzados, dotados de disco duro, y que además deben ser compatibles con la plataforma IPTV.

Los descodificadores básicos de una plataforma IPTV que no ha ofrecido facilidades de grabación o gestión del directo no incorporan disco duro. Es por tanto necesario dotar al STB de un disco duro local, con capacidad de almacenar contenidos Time Shifting y DVR, y dotado de algún software que permita la gestión de los contenidos procedentes de DVR: crear carpetas, renombrar archivos, proteger archivos, borrar archivos...

Existe la posibilidad de añadir un disco duro externo USB al STB, pero esto se evitará en la medida de lo posible, debido principalmente a dos motivos:

- Resulta más rentable el STB con disco duro interno, ya que, desde siempre, resulta más económico comprar un dispositivo final que comprar los subsistemas por separado. Aquellos usuarios con STB sin disco duro que deseen contratar un servicio de grabación de contenidos deberán sustituir su STB por un STB con disco duro interno.
- El almacenamiento en dispositivos extraíbles hace imposible controlar el uso que se le pueda dar a los contenidos. Los suministradores de contenidos podrían acordar no proveer de programación a aquellos operadores que no puedan garantizar la privacidad y exclusividad del contenido.

La implementación del servicio de grabación continua hace necesario que el disco duro sea un HDD de altas prestaciones, ya que como se explicó anteriormente, se somete al disco duro a un esfuerzo excesivo. También es necesario que el disco duro, integrado en el STB, venga convenientemente particionado para soportar Time Shifting.

También es necesario dotar al STB de un doble sintonizador IP. Para realizar una grabación DVR de un programa, necesitamos que el sintonizador esté recibiendo el canal que queremos grabar. Por tanto, para poder seguir navegando por otros canales, es necesario otro sintonizador, ya que el primero está fijo grabando un canal. Uno de ellos se usará para sintonizar las grabaciones DVR y el

otro para navegar por los canales y recoger las grabaciones Time Shifting. La lógica entre los distintos sintonizadores que permita la grabación y la visualización de distintos canales debe venir fijada por el fabricante.

Como hemos dicho anteriormente, a un STB que ofrece facilidades de grabación de contenidos se le conoce como STB avanzado. El precio de un STB avanzado, según fabricante, supone un incremento de en torno a un 80% del precio con respecto al STB básico. Este hecho condiciona fuertemente el modelo de negocio del servicio, como se verá más adelante.

5.1. Cálculo de capacidad de discos duros

A la hora de elegir el disco duro, debemos encontrar una relación de compromiso entre el tamaño del disco y el precio. En función del formato de grabación, podremos almacenar mayor o menor número de horas. Es importante que el STB sea compatible multi-bitrate, ya que los avances en la codificación del vídeo harán que las tasas de bit sean distintas. A modo de ejemplo, la mayoría de operadores de IPTV empezaron transmitiendo en formato MPEG-2, con una tasa de en torno 4 ó 4,5 Mbps. Actualmente, los que aún transmiten en MPEG-2, están preparando la migración a MPEG-4, ya que al transmitir a una menor tasa de bit (2,5 Mbps), pueden llegar más lejos y aumentar así la zona de cobertura del operador. El futuro será la televisión digital de alta definición, HDTV, que se prevé emita en torno a 12 Mbps. Puesto que se va a cambiar el equipamiento de cliente de la plataforma, el operador de IPTV debe ser capaz de predecir otros cambios futuros y obtener así un resultado consistente en el tiempo. De nada serviría integrar un disco duro en un STB que únicamente grabara formato MPEG-2.

La siguiente tabla muestra el número de horas que se podría grabar en función de la capacidad del disco duro y la codificación:

		Régimen binario (Mbps)		
		2,5 (MPEG-4)	4,5 (MPEG-2)	12 (HDTV)
HDD (GB)	40	36,41	20,23	7,59
	60	54,61	30,34	11,38
	80	72,82	40,45	15,17
	160	145,64	80,91	30,34

A la vista de esta tabla, el operador debería integrar discos duros con capacidad de al menos 80 GB, siendo recomendable que la capacidad del disco fuera de 160 GB. Este disco permite además lectura y escritura simultánea, de modo que nos permite realizar una grabación DVR o ShiftTV al tiempo que reproducimos otra.

5.2. Necesidad de sintonizadores de TDT.

Actualmente, los operadores de IPTV ofrecen la posibilidad de abonar una cuota mensual en concepto de alquiler del STB o de comprar el mismo. Para que el modelo en compra tenga sentido, es necesario dotar al STB de sintonizadores de TDT.

El cliente que suscribe los servicios de grabación de contenidos y que adquiere el STB en propiedad desea poder utilizarlo siempre, incluso en el caso de que curse baja del operador.

Si el cliente cursa baja del operador de IPTV, dejará de recibir los canales por esa vía. Puesto que en 2010 únicamente existirá la difusión en abierto en formato TDT, debe preverse que para que un usuario compre el STB, debe poder recibir la señal de TDT, para que en caso de cursar baja pueda seguir disfrutando de sus ventajas. Este hecho abriría, además, la posibilidad de venta del STB como sintonizador de TDT y grabador, aumentando las posibilidades de negocio.

La señal de TDT que recibe el cliente del operador de IPTV suele proceder de la cabecera central de TV, no de una antena receptora en el domicilio de cliente. Para que en un futuro pueda recibir los canales, aunque curse baja del operador, debe dotarse al STB de sintonizadores de TDT.

En el caso de que el STB incluya sintonizadores de TDT, el operador debe integrar en su aplicación de grabación de contenidos los canales TDT, de modo que se puedan realizar grabaciones DVR y Time Shifting de los mismos.

Debe estudiarse la viabilidad económica de este hecho. La integración de un sintonizador de TDT en el STB aumenta el coste de este último en torno a 5€ por sintonizador. Si se opta por integrar dos sintonizadores de TDT, uno para la visualización y otro que sintonice otro canal distinto para grabación DVR, se incrementa el coste del STB en torno a 10€. El beneficio que se obtiene es, por un lado, la posibilidad de vender el STB, que sería útil para operadores con alta tasa de rotación de planta (bajas del servicio elevadas), que no obtienen beneficio del alquiler del STB debido a que la corta permanencia del usuario en el operador no

es suficiente para amortizar los gastos de compra del STB. Por otro lado, se abre el negocio de venta de sintonizadores de TDT con grabador.

5.3. Elección del SET-TOP-BOX

En función de los actuales STB dotados de disco duro que existen en el mercado, se explica a continuación cuál sería la mejor elección para un operador de IPTV.

5.3.1. SAGEM

Empezamos la selección por uno de los mayores fabricantes de STB's de Europa: Sagem. Este fabricante provee de STB a distintos operadores en Francia, Reino Unido, Alemania, Italia, Suecia y España.

De los STB que actualmente ofrece SAGEM, elegimos el modelo IAD5116. Este modelo tiene un disco duro de 40 GB (opcional 80GB), y nos permite un Time Shifting de hasta 30', con grabación continua. Desgraciadamente, este modelo sólo permite grabaciones en formato MPEG-2, y atarse a este estándar supondría quedarse atrás en la carrera de mercado por la IPTV.

5.3.2. KREATEL

Kreatel fue la compañía elegida por Telefónica para sus STB de Imagenio. El STB de Kreatel que más se ajusta a nuestras necesidades es el VIP1970. Este STB tiene un disco duro de 80 GB, permite Time Shifting y además incluye un sintonizador de TDT. Permite asimismo la grabación en formatos MPEG-2, MPEG-4 y HDTV.

5.3.3. ADB

De la compañía ADB elegimos el modelo 5810TWX. Ofrece un disco duro de 160 GB, incluye sintonizador de TDT, soporta DRM y es capaz de grabar contenidos en MPEG-2, MPEG-4 y HDTV.

En la siguiente tabla se resume la comparativa entre los tres fabricantes:

	SAGEM IAD5116	KREATEL VIP1970	ADB 5810TWX
16:9	s.i	X	X
4:3	s.i	X	X
ACCESO CONDICIONAL	s.i	s.i.	X
DRM	s.i	s.i.	X
EPG	X	X	X
HDD	40 GB	80 GB	160 GB
MP3 PLAYER	s.i	X	X
MPEG1	s.i	X	X
MPEG1 LAYER 1	X	X	X
MPEG1 LAYER 2	X	X	X
MPEG1 LAYER 3	s.i	X	X
MPEG2 AAC	s.i	X	X
MPEG2 HDTV	s.i	X	X
MPEG2 SDTV	X	X	X
MPEG4 AAC	s.i	X	X
MPEG4 HDTV	s.i	X	X
MPEG4 SDTV	s.i	X	X
PHOTO ALBUM	s.i	X	X
PIP	s.i	X	X
SINTONIZADOR TDT	s.i	X	X
SW UPGRADABLE	X	X	X
VoIP	s.i	s.i.	X
WMA	s.i	s.i	X

s.i.: sin información del fabricante.

A la vista de los resultados, el operador debería elegir como descodificador el ADB 5810 TWX. El fabricante asegura que el disco duro es de alta calidad, garantizando su uso por al menos 5 años y hasta 500.000 horas.

6. DERECHOS LEGALES SOBRE LA GRABACIÓN DE CONTENIDOS

Los proveedores que suministran contenidos a la plataforma de IPTV manifiestan su necesidad de evitar la copia ilegal; podrían restringir el uso del servicio de DVR.

Los servidores realizan diversas funciones, entre ellas el almacenamiento de contenido, la administración del video bajo demanda y la gestión de licencias sobre derechos de contenidos digitales, DRM (*Digital Rights Management*).

El servidor DRM codifica el contenido y lo encapsula en un contenedor para evitar su uso no autorizado. Existen sistemas de primera y segunda generación. Los de primera generación no suelen imponer restricciones de uso sobre los contenidos una vez abiertos (esto se conoce como acceso condicional, si tengo acceso al contenido, podría tener acceso a su manipulación). Los de segunda generación especifican a través de una licencia qué usos se puede dar al objeto.

De este modo, el proveedor de contenido puede permitir o no la grabación del contenido en DVR. En principio, el proveedor debería permitir la gestión de directo mediante Time Shifting, ya que el uso que se le da al contenido escapa de la manipulación ilegal por parte del cliente (el contenido se borra una vez visualizado).

Los proveedores que permiten la grabación de sus contenidos no desean que estos sean de libre acceso e intercambio. Esto se podría asegurar por triplicado.

Por un lado, la propia plataforma del operador exige autenticación al encender el STB. No será posible acceder a los contenidos ya grabados en el disco duro sin estar conectado a la red de IPTV, evitando así el libre acceso a los mismos. El segundo mecanismo es la licencia DRM, que limita el uso que se puede dar a los contenidos. El tercer mecanismo es la codificación del contenido en el disco duro del STB, que se realiza mediante una encriptación propia, evitando que se puedan reproducir en otros sistemas. El STB podría conectarse a un PC, vía USB, y usarse para almacenar imágenes o música. De este modo se asegura la privacidad del contenido. Hay que estar abonado a la plataforma IPTV y gozar de licencia DRM sobre el contenido para poder reproducirlo, y además, sólo se puede reproducir desde el propio STB.

7. IMPACTO EN LA ARQUITECTURA Y SISTEMAS DE LA RED

7.1. Impacto en el servidor de aplicaciones y provisión de red.

El sistema de provisión de red debe ordenar al servidor de aplicaciones el acceso al menú de grabación de contenidos para aquellos usuarios que hayan contratado el servicio. Para ello debe, debe permitir el acceso a aquellas IP privadas estáticas correspondientes a los STB cuyos usuarios hayan contratado el servicio.

7.2. Implicaciones en la EPG.

La fiabilidad de los servicios de grabación de contenidos depende de la fiabilidad de la Guía de Programación Electrónica, en los modos de grabación instantánea y programada desde la EPG. La EPG debe suministrar una información fiable y actualizada en tiempo real sobre la programación ^[18], capaz de responder a cambios en la misma, para que las grabaciones de un contenido completo se realicen de forma correcta.

La EPG debe suministrar a la aplicación de grabación de contenidos información sobre si se tienen derechos para la grabación del contenido, el formato de grabación, etc, y añadir nuevas marcas adicionales que sirvan para mejorar el servicio en un futuro. Al tal efecto, es necesario que la Guía incluya una serie de marcas, algunas por contenido y otras por canal, como se explica en la siguiente tabla:

Marca	Necesaria por contenido	Necesaria por canal
Derechos grabación DVR	X	X
Derechos de Time Shifting	X	X
Contenido PPV	X	X
Contenido VbD		X
Contenido en MPEG-2	X	X
Contenido en MPEG-4	X	X
Contenido en SDTV	X	X
Contenido en HDTV	X	X
Contenido repetido	X	
Periodicidad del contenido	X	

7.2.1. Marca de derechos de grabación DVR.

La marca de derechos de grabación DVR es necesaria para deshabilitar la grabación de los contenidos para los cuales el operador de IPTV no haya llegado a un acuerdo con el suministrador de contenidos.

La marca identificará aquellos canales para los cuales se haya negociado la posibilidad de grabación. De todos modos, es posible que un suministrador de contenidos no permita la grabación de su programación, a excepción de una serie de contenidos considerados de interés general. A tal efecto, es necesaria una marca individual que identifique estos contenidos y permita su grabación.

7.2.2. Marca de derechos Time Shifting.

El funcionamiento de esta marca es similar a la marca de derechos de grabación DVR, aunque resulta mucho menos crítica, pues independientemente de que el contenido pueda ser grabado en Time Shifting, este escapa de la manipulación por parte del usuario. Por este hecho, se presupone que los suministradores no tendrán inconveniente en permitir el Time Shifting de sus contenidos.

7.2.3. Marca de contenido PPV.

La marca de contenido PPV es necesaria para evitar la grabación de contenidos de pago, pues esta idea va en contra del modelo de negocio del operador. No obstante, del mismo modo que ocurría en la marca de grabación DVR, podría permitirse la grabación de eventos PPV de interés general, con un coste adicional en la compra del contenido PPV. Por defecto, la marca PPV permitirá el Time Shifting del evento.

7.2.4. Marca de contenido VbD.

Bajo ningún concepto el operador estará dispuesto a permitir la grabación de contenidos en vídeo bajo demanda, pues suponen una importante fuente de ingresos. En el caso de que la EPG no sea capaz de soportar todas las marcas de grabación necesarias, podría prescindirse de esta marca y discriminar la grabación de eventos VbD deshabilitando la aplicación de grabación al acceder al menú de vídeo bajo demanda.

Nótese que el Time Shifting no tiene sentido en el caso del vídeo bajo demanda, ya que el propio servicio permite las funciones de avance rápido, rebobinado, pausa y reproducción.

7.2.5. Marcas de contenido en formato MPEG-2 y MPEG-4.

Estas marcas son necesarias para habilitar la grabación a distintas tasas de bits, y realizar una grabación correcta. Por defecto, el formato de los contenidos será el mismo para un canal completo. No obstante, puede ser que ciertos contenidos se codifiquen en otro formato, lo que hace necesario marcas para canal y contenido individual.

7.2.6. Marcas de contenido en formato SDTV y HDTV.

Su funcionamiento es similar a las marcas de contenido en formato MPEG-2 y MPEG-4, y su finalidad es la misma.

7.2.7. Marca de contenido repetido.

Esta marca se introduce para permitir la mejora del servicio de grabación de contenidos anteriormente descrito.

Algunos operadores de IPTV como Tivo (EE.UU.) y empresas especializadas en servicios de grabación de contenidos (InOut TV) ofrecen la posibilidad de programar grabaciones de eventos periódicos, como series y documentales. Con un único botón, podría programarse la grabación de todos los capítulos de una serie. Para que esta facilidad pueda llevarse a cabo de manera eficiente, es necesario identificar los contenidos repetidos, evitando así grabaciones innecesarias, que supondrían una visión imperfecta del servicio de grabación por parte del usuario.

7.2.8. Marca de periodicidad del contenido

La marca de periodicidad del contenido informa a la aplicación de grabación sobre la periodicidad del mismo, para hacer posible la facilidad de grabación de eventos periódicos anteriormente descrita.

8. IMPACTO EN SISTEMAS COMERCIALES

De acuerdo con los requisitos del servicio, se contempla la posibilidad de grabación de 1 canal de IPTV mientras se visualiza otro diferente. Esto implica que al cliente le tendrán que llegar simultáneamente 2 canales de IPTV.

Cuando un usuario interesado en los servicios de grabación de contenidos se ponga en contacto con un sistema de contratación, deberá comprobarse si el usuario dispone de cobertura suficiente como para poder ofrecerle dos canales de TV. El operador deberá elaborar un nuevo mapa de cobertura que identificará en qué zonas se podrá dar el servicio de DVR / Time Shifting y en cuales no por limitaciones de ancho de banda.

Para que el servicio tenga una penetración mayor, se puede optar por ofrecerlo en aquellas ubicaciones en las que se disponga de ancho de banda suficiente para únicamente un canal, e informar al usuario de que, por las características y la atenuación del bucle, puede realizar grabaciones pero no podrá ver otro canal al mismo tiempo. De este modo, sería necesario generar dos códigos comerciales para contratación y facturación, ambos con el mismo precio, que distingan esta situación, de modo que ambos fueran accesibles desde los sistemas de contratación.

Los sistemas de contratación del operador deben adaptarse para presentar estos servicios de suscripción mensual. Al iniciar la contratación, se comprueba de manera automática la disponibilidad del bucle. Si la central a la que pertenece el número de abonado puede ofrecer servicios de IPTV, y la calidad del bucle de abonado es suficiente para poder prestar dicho servicio sobre dos canales, debe presentarse al operador la posibilidad de contratarlo, con el código comercial correspondiente a DVR/ Time Shifting sobre dos canales de IPTV. En caso de que sólo se pueda prestar el servicio DVR/ Time Shifting sobre un canal, debe presentarse al comercial la posibilidad de contratarlo, indicando que debe notificar al cliente de que únicamente podrá realizar grabaciones del canal que esté visualizando, sin posibilidad de ver otro canal de IPTV a la vez. En el caso de esta contratación debe registrarse la misma con el código comercial correspondiente a DVR/Time Shifting sobre un único canal de IPTV.

Puesto que los servicios de grabación de contenidos implican que el equipo de cliente sea un STB dotado de disco duro, son dos las acciones de contratación que deben llevarse a cabo en el sistema: servicio de suscripción mensual DVR/ Time Shifting y alquiler de STB avanzado.

En cuanto a los impactos en la gestión de los descodificadores, se pueden dar 2 casos:

- Alta en el servicio de manera simultánea a el alta en la plataforma: se asignan directamente al cliente los códigos del servicio de DVR/ Time Shifting y del descodificador avanzado en alquiler.
- Alta en el servicio de un cliente ya existente (con descodificador estándar): En este caso será necesario proceder a la recogida del descodificador que tiene el cliente (estándar) y sustituirlo por el descodificador avanzado.

En la contratación se deberá articular este mecanismo de sustitución de equipamiento para este caso.

Capítulo IV.

ESTUDIO DE RENTABILIDAD.

9. VIABILIDAD ECONÓMICA.

9.1. Impacto en ingresos del operador.

Los servicios de grabación de contenidos generan dos efectos colaterales, uno de ellos beneficioso, y otro perjudicial.

Por un lado, adelantarse al mercado y ofrecer servicios de grabación harían del operador un operador de mayor prestigio. Los servicios de grabación se presentan como servicios diferenciadores, que aumentan notablemente la calidad del operador. El cliente se sentiría partícipe de una experiencia selecta, que no ofrecen otros operadores, y este hecho tiene un efecto fidelizador con el operador, actuando como gancho y reduciendo la tasa de bajas del operador, del mismo modo que aumentaría el número de abonados.

Por otro lado, experiencias anteriores en Europa y EE.UU. demuestran que, al ofrecerse la posibilidad de grabar contenidos, se reduce el número de consumos en vídeo bajo demanda.

Estos dos efectos deben tenerse en cuenta en el análisis económico de viabilidad del servicio para el operador, y debe evaluarse si el efecto global es positivo o si la reducción de consumos bajo demanda reduce la rentabilidad del negocio.

9.2. Estudio de viabilidad económica.

A continuación, se hará un estudio de la viabilidad económica de los servicios de grabación. El horizonte de la planificación será de cuatro años. En primer lugar, debemos resumir la situación de la planta del operador de IPTV en el año anterior a la implantación, y realizar una estimación de la ampliación de planta durante los años siguientes.

Supondremos que la planta nuestro operador de IPTV es actualmente (año 0) de 250.000 abonados, y se espera un crecimiento de unos 100.000 abonados al año, hasta estabilizarse en una planta de 500.000.

Datos de planta	Año 0	Año 1	Año 2	Año 3	Año 4
Planta año anterior	250.000				
Estimación planta fin de año		300.000	400.000	500.000	500.000
Planta media		275.000	350.000	450.000	500.000
Churn anual	30,00%	30,00%	30,00%	30,00%	30,00%
Bajas		82.500	105.000	135.000	150.000
Altas totales		107.500	155.000	185.000	150.000

Se sabe que los servicios de grabación de contenidos resultan de gran atractivo, y se espera para ellos una penetración del 30% total sobre la planta del operador. De este modo:

Penetración del servicio	Año 0	Año 1	Año 2	Año 3	Año 4
Índice de penetración del servicio		10,00%	20,00%	30,00%	30,00%
Planta STB HDD final año		30.000	80.000	150.000	150.000
Planta STB HDD media		15.000	55.000	115.000	150.000

Los parámetros financieros que rigen el proyecto son:

- Periodo de amortización: adoptamos una amortización a 3 años.
- Tasa Impositiva (Impuesto de sociedades): para el año 2007, la tasa impositiva fijada por el Ministerio de Economía y Hacienda es del 32.5%, y se prevé sea del 30% para el año 2008.
- Tasa de actualización: nuestro operador de IPTV fija la tasa de actualización en el 10%.

Después de un estudio de mercado, y basándose en experiencias anteriores en otros países, nuestro operador de IPTV estima conveniente fijar la cuota mensual de suscripción al servicio en 7€.

Cuota mensual alquiler	Año 0	Año 1	Año 2	Año 3	Año 4
Cuota mensual alquiler		7,00 €	7,00 €	7,00 €	7,00 €

De este modo, con los datos de abonados a los servicios de grabación de contenidos y la cuota mensual de suscripción, obtenemos los ingresos derivados de la cuota de suscripción mensual:

Ingresos	Año 0	Año 1	Año 2	Año 3	Año 4
Ingreso por cuota mensual		1.260.000 €	4.620.000 €	9.660.000 €	12.600.000 €

Las inversiones que se realizarán para llevar a cabo el proyecto vienen derivadas, por un lado, de la inclusión de un nuevo equipo de cliente dotado de disco duro, y por otro lado, de los desarrollos software que son necesarios para incluir las funcionalidades de grabación en la plataforma del operador.

La inversión en STB dotados de disco duro se considera la diferencia entre el precio del STB básico y el STB con disco duro, que es realmente el coste extraordinario que tendría que asumir el operador. Nuestro operador decide por no incorporar al STB descodificadores de TDT para disminuir el coste del STB, y desarrollar el modelo en venta del mismo una vez empiecen a obtenerse beneficios.

Inversiones	Año 0	Año 1	Año 2	Año 3	Año 4
Inversión en nuevos STB					
Coste STB básico		120,00 €	110,00 €	100,00 €	100,00 €
Coste STB HDD		200,00 €	190,00 €	170,00 €	170,00 €
Diferencia		80,00 €	80,00 €	70,00 €	70,00 €
Inversión en STB HDD		2.400.000 €	4.000.000 €	4.900.000 €	0,00 €
Inversión en desarrollos SW		250.000 €	100.000 €	50.000 €	0,00 €
Inversión total		2.650.000 €	4.100.000 €	4.950.000 €	0,00 €

Se pretende hacer una amortización agresiva a 3 años. De este modo:

Amortizaciones	Año 1	Año 2	Año 3	Año 4	Año 5	Año 6
Nueva Inversión	2.650.000 €	4.100.000 €	4.950.000 €	0 €	0 €	0 €
Inversión Acumulada	2.650.000 €	6.750.000 €	11.700.000 €	9.050.000 €	4.950.000 €	0 €
Gastos Amortización	883.333 €	2.250.000 €	3.900.000 €	3.016.667 €	1.650.000 €	0 €
Amortización Acumulada	883.333 €	3.133.333 €	4.383.333 €	3.300.000 €	0 €	0 €
Valor Neto contable de la Inversión	1.766.667 €	3.616.667 €	7.316.667 €	5.750.000 €	4.950.000 €	0 €

Los costes del proyecto son los propios de la plataforma: mantenimiento, comercialización, marketing y provisión por insolvencias. Modelaremos estos costes como un tanto por ciento sobre los ingresos totales.

En el caso de que un abonado a la plataforma desee suscribir los servicios de grabación de contenidos, su STB básico debería ser sustituido por un STB dotado de disco duro. La visita del instalador al domicilio del cliente supone un gasto añadido no contemplado por el negocio del operador, sino que es propio de este servicio.

Costes	Año 0	Año 1	Año 2	Año 3	Año 4
% Clientes desde STB básico		90,00%	70,00%	30,00%	5,00%
Coste Instalación STB		30,00 €	29,00 €	28,00 €	27,00 €
Clientes cambian STB		27.000	56.000	45.000	7.500
Coste Instalación total		810.000 €	1.624.000 €	1.260.000 €	202.500 €
% ingresos para mantenimiento		2,00%	2,00%	2,00%	2,00%
Coste mantenimiento		25.200 €	92.400 €	193.200 €	252.000 €
% ingresos previsión insolvencias		2,00%	2,00%	2,00%	2,00%
Provisión por insolvencias		25.200 €	92.400 €	193.200 €	252.000 €
% ingresos costes comercialización		0,50%	0,50%	0,50%	0,50%
Costes Comercialización		6.300 €	23.100 €	48.300 €	63.000 €
% ingresos costes marketing		1,00%	1,00%	1,00%	1,00%
Costes de Marketing		12.600 €	46.200 €	96.600 €	126.000 €
Amortizaciones		883.333 €	2.250.000 €	3.900.000 €	3.016.667 €
Coste total		1.762.633 €	4.128.100 €	5.691.300 €	3.912.167 €

Además de estos ingresos y costes, también es necesario evaluar los efectos de disminución de consumos en vídeo bajo demanda, y el efecto fidelizador sobre la planta de abonados a los servicios de grabación de contenidos. Es conveniente tener en cuenta que la reducción de ingresos en el vídeo bajo demanda también implica una reducción de costes en el mismo, y que el aumento de planta debido al efecto fidelizador conlleva también un aumento de costes, al aumentar el número de abonados.

Efectos colaterales	Año 0	Año 1	Año 2	Año 3	Año 4
Reducción ARPU mensual VoD		2,00 €	2,00 €	2,00 €	2,00 €
Reducción costes/ mes usuario VoD		1,20 €	1,20 €	1,20 €	1,20 €
Resultado mensual por cliente VoD		-0,80 €	-0,80 €	-0,80 €	-0,80 €
Resultado disminución VoD		-12.000 €	-44.000 €	-92.000 €	-120.000 €
ARPU mensual del operador	50,00 €	51,00 €	52,00 €	53,00 €	53,50 €
Costes por usuario/ mes	40,00 €		39,00 €	38,50 €	38,00 €
Resultado mensual por cliente	10,00 €	11,50 €	13,00 €	14,50 €	15,50 €
Reducción del Churn para grabación		10,00%	15,00%	20,00%	20,00%
Churn clientes grabación		27,00%	25,50%	24,00%	24,00%
Clientes fidelizados por la grabación		450	2.475	6.900	9.000
Efecto de la fidelización		62.100 €	386.100 €	1.200.600 €	1.674.000 €
Resultado efectos colaterales		50.100 €	342.100 €	1.108.600 €	1.554.000 €

Para obtener el resultado tras los primeros 4 años, es necesario hacer liquidación de todos los flujos de caja en ese momento. De este modo, tomamos el amortizado pendiente como un ingreso de caja extraordinario a final del cuarto año, ponderado al 80% para modelar la depreciación que supone la no amortización. Los ingresos finales serían:

Ingresos Totales	Año 0	Año 1	Año 2	Año 3	Año 4
Valor residual amortización					4.600.000 €
Ingreso por cuota mensual		1.260.000 €	4.620.000 €	9.660.000 €	12.600.000 €
Ingresos Totales		1.260.000 €	4.620.000 €	9.660.000 €	17.200.000 €

Liquidamos los flujos para obtener los indicadores de rentabilidad del proyecto:

Resultados	Año 0	Año 1	Año 2	Año 3	Año 4
EBITDA		430.800 €	3.084.000 €	8.977.300 €	17.858.500 €
EBIT		-452.533 €	834.000 €	5.077.300 €	14.841.833 €
Resultado tras impuestos		-305.460 €	583.800 €	3.554.110 €	10.389.283 €
Resultado acumulado		-305.460 €	278.340 €	3.832.450 €	14.221.733 €
Cash Flow Libre		577.873 €	2.833.800 €	7.454.110 €	13.405.950 €
Cash Flow Real		-2.072.127 €	-1.266.200 €	2.504.110 €	13.405.950 €
Cash Flow Real acumulado		-2.072.127 €	-3.338.327 €	-834.217 €	12.571.733 €
Cash Flow R.A. actualizado		-1.883.752 €	-2.930.198 €	-1.048.823 €	8.107.621 €
Nº Años Pay Back		0	0	0	3,11

El resultado neto de la explotación a final del cuarto año es:

VAN	615.465,86 €
TIR	16%

Por tanto, este proyecto resulta viable desde el punto de vista económico.

Capítulo V.

TENDENCIAS EN LA GRABACIÓN DE CONTENIDOS.

10. GRABACIÓN DE CONTENIDOS EN RED.

Los servicios de grabación DVR y Time Shifting han sido presentados como servicios de grabación en local, a partir de mejoras en los equipos de usuario. Sin embargo, las mismas funciones pueden ser llevadas a cabo por la red, sin la necesidad de incluir discos duros en los STB.

Dado el comportamiento que se persigue en estos servicios (grabación y facilidades de Vídeo bajo Demanda), el almacenamiento de contenidos se realizaría en los servidores de vídeo de las Células de Servicio Locales.

En el servicio DVR, el usuario podría programar una grabación y esta se almacenaría en los servidores de vídeo locales. La repercusión es clara: sería necesaria una gran inversión en almacenamiento que actualmente sólo se podría afrontar si se elevara notablemente la cuota de suscripción al servicio (se entiende que sería un servicio de suscripción mensual), desechando este servicio desde el punto de vista de rentabilidad.

El caso del Time Shifting es todavía más problemático. En lugar de tratarse de grabaciones puntuales programadas por el usuario, se trata de grabar todas las contribuciones que en un momento dado se estuvieran reproduciendo en alguna de las ubicaciones de abonado pertenecientes a cada Célula de Servicios Local. De este modo, se aseguraría la disponibilidad de los últimos minutos de cada canal visualizado por los abonados (en caso de que se opte por implementar la grabación continua), y se pondría a disposición de los usuarios un buffer de grabación en red para realizar Time Shifting explícito (pulsando Pause).

El problema del equipamiento no se resuelve simplemente añadiendo más discos que aporten una capacidad extra de almacenamiento a los servidores de vídeo locales. El comportamiento de lectura/ escritura simultánea no ha sido reflejado en el diseño inicial de la plataforma, y por tanto se habría optado por elegir discos y controladoras que permitiesen lectura múltiple y la grabación de contenidos únicamente por inyección desde el servidor central de la Cabecera de TV. Sería necesario reemplazar todo el sistema de servidores por unos que permitan multi-lectura y multi-escritura.

No obstante, los principales operadores de IPTV del mundo ya están trabajando en esta línea, para permitir en un futuro la grabación de contenidos en la propia red y no en los terminales de usuario.

Especial mención requiere el caso de algunos operadores como Telefónica O2 en la República Checa. El bajo índice de penetración de la IPTV en el país debido a su reciente implantación (cuenta sólo con 10.000 abonados), y su arquitectura centralizada a nivel de servidores de contenido bajo demanda (existe una única célula de servicios local), han hecho posible la grabación de contenidos DVR en red, para un conjunto reducido de contribuciones, y a costa de dedicar más del 80% de la capacidad de almacenamiento de los servidores de vídeo al servicio de DVR.

Incluso bajo estas condiciones, la grabación personalizada por cada usuario resulta inviable. La solución que ha elegido el operador ha sido la grabación DVR de aquellos contenidos de mayor interés (prime-time), para ponerlos disposición de los usuarios que lo así soliciten bajo demanda. Se trata pues, no de un DVR de usuario como el que ha sido descrito en este Proyecto, sino de un DVR de Operador en red, pero de este modo se ha iniciado el camino hacia la grabación DVR de usuario en red.

APENDICES

11. VOCABULARIO ESPECÍFICO.

ARPU	Average Revenue Per Unit
ATU	ADSL Terminal Unit
Churn	Tasa de rotación anual de planta
DRM	Digital Rights Management
DSLAM	Digital Subscriber Line Access Multiplexor
DVB	Digital Video Broadcast
DVR	Digital Video Recording
EBIT	Earnings Before Interests and Taxes
EBITDA	Earnings Before Interests, Taxes, Depreciation and Amortization
EPG	Electronic Program Guide
GRE	Generis Routing Encapsulation
HDTV	High Definition Televisión
IETF	Internet Ingenieering Task Force
IP	Internet Protocol
IPTV	Internet Protocol Television
JPEG	Joint Picture Expert Group
MPEG	Motion Picture Expert Group
PIN	Personal Identity Number
PiP	Picture in Picture
PPV	Pay per View
PVR	Personal Video Recording
Raid	Redundance Array of Independent (Inexpensive) Disks
RSS	Raid Set Size
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTSP	Real Time Streaming Protocol
SDH	Sinchronous Digital Hierarchy
SDTV	Standard Definition Televisión
STB	Set-Top-Box
TCP	Transmision Control Protocol
TDT	Terrestrial Digital Televisión
TIR	Tasa Interna de Rentabilidad
UDP	User Datagram Protocol
VAN	Valor Actual Neto
VbD	Vídeo bajo Demanda
VoD	Video on Demand
XML	Extensible Markup Language

12. REFERENCIAS.

- [1] www.campusred.net/telos/anteriores/suplementos/supl_41/ponencias/ponencias_03.htm
- [2] www.inad.es/sp/tecnologica/2006/03200601.htm
- [3] www.telefonica.es/sociedaddelainformacion/pdf/publicaciones/telecommultimedia/capitulos/03_usos_y_entornos.pdf
- [4] <http://www.tmtfactory.com/articulos/articuloTVFinder200310.pdf>
- [5] www.accenture.com/Global/Services/By_Industry/Communications/Access_Newsletter/Article_Index/iTVCable.htm
- [6] www.vives.org/agenda/tdt/Canyero.pdf
- [7] www.tvyvideo.com/tv/secciones/TV/ES/MAIN/IN/ARTICULOSREPORTAJESHOME/doc_52559_HTML.html?idDocumento=52559
- [8] www.accenture.com/Microsites/Internet_Protocol_TV/Research_and_Insights/IPTVMonitor3
- [9] www.vertele.com/pistas/detail.php?id=162
- [10] <http://todosobremitele.blogcindario.com/2006/08/03463-la-television-digital-avanza-a-pasos-agigantados.html>
- [11] [http://kc3.pwc.es/local/es/kc3/PrensaR.nsf/V1001/2EC9224909EE6125C125720C002C1ABC/\\$FILE/pdfGlobalOutlookPres17-10-06.pdf](http://kc3.pwc.es/local/es/kc3/PrensaR.nsf/V1001/2EC9224909EE6125C125720C002C1ABC/$FILE/pdfGlobalOutlookPres17-10-06.pdf)
- [12] www.e-mediainstitute.com/it/press/previsionimercatoiptv_it.content
- [13] http://ec.europa.eu/information_society/policy/ecommerce/doc/info_centre/communic_reports/interoperability_idtv/comm_pdf_com_2006_0037_f_es_acte.pdf
- [14] www.cinit.org.mx/articulo.php?idArticulo=34
- [15] www.tid.es/documentos/libros_sector_telecomunicaciones/imagenio.pdf
- [16] www.monografias.com/trabajos14/acceso-atm/acceso-atm.shtml
- [17] www.cablesatelitetele.com/reportaje_5_noviembre2005.htm
- [18] <http://www-gris.det.uvigo.es/~agil/NST/nst3.pdf>

12.1. Otras referencias.

- 1) <http://pub.ufasta.edu.ar/SISD/mpeg/mpeg1.htm>
- 2) <http://societaddelainformacion.telefonica.es/jsp/articulos/detalle.jsp?elem=3474>
- 3) <http://webdelprofesor.ula.ve/humanidades/raymond/computacion2/archivos/VideoDigital.pdf>
- 4) www.adslzone.net/comparativa-adsl-paises.html
- 5) www.ananova.com
- 6) www.arrakis.es/~aikido/interdic/articul8.htm
- 7) www.atc.uniovi.es/teleco/5tm/archives/3comp.pdf
- 8) www.gpi.tsc.uvigo.es/~fmartin/Web-SI/ArticuloTDT-Loreto.pdf
- 9) www.hectormilla.com/categor%C3%ADa/IPTV/
- 10) www.inout.tv/SincroGuia
- 11) www.telefonica.es/societaddelainformacion/pdf/informes/europa_2002/parte_2_8.pdf
- 12) www.telefonica.es/societaddelainformacion/pdf/publicaciones/imagenio/capitulos/imageniocap7.pdf
- 13) www.telefonica.es/societaddelainformacion/pdf/publicaciones/telecommultimedia/capitulos/06_negocios.pdf
- 14) www.televisiondigital.es/Cable/Informacion/OperadorTVCable/aunaono.htm
- 15) www.televisiondigital.es/NR/rdonlyres/0ABF870B-7E9F-43D8-B38C-FD73FD7A7DB8/6042/GT3SG2Serviciosinteractivosb%C3%A1sicosbasadosenMHP.pdf
- 16) www.tid.es/documentos/revista_comunicaciones_i+d/numero38.pdf
- 17) www.titantv.com
- 18) www.tivo.com
- 19) www.upv.es/satelite/trabajos/Grupo8_99.00/mpeg.html

ANEXOS.

ESTÁNDARES MPEG.

13. ESTÁNDARES MPEG.

MPEG es el nombre de un grupo de estándares de codificación de audio y vídeo normalizados por el grupo MPEG (Moving Pictures Experts Group), cuyo nombre formal es ISO/IEC JTC1/SC29 WG11. Desde su primera reunión en 1988, el MPEG ha crecido hasta incluir 350 miembros de distintas industrias y universidades. Los formatos de compresión más conocidos que ha normalizado el grupo MPEG son:

- **MPEG-1:** estándar inicial de compresión de audio y vídeo. Usado después como la norma para el formato VideoCD, incluye popular formato de compresión de audio Capa 3 (MP3).
- **MPEG-2:** normas para audio y vídeo para difusión de calidad de televisión. Utilizado para servicios de TV por satélite, cable, IPTV y los discos de vídeo DVD.
- **MPEG-3:** diseñado originalmente para HDTV (Televisión de Alta Definición), pero abandonado posteriormente en favor de MPEG-2.
- **MPEG-4:** expande MPEG-1 para soportar objetos audio/vídeo, contenido 3D, codificación de baja velocidad binaria y soporte para gestión de derechos digitales (protección de copyright, Digital Rights Management).

14. ESTÁNDAR MPEG-1.

La calidad del vídeo comprimido mediante MPEG-1 con una tasa binaria de 1,2 Mbps se puede comparar con la de un grabador de vídeo VHS. El principal objetivo de MPEG-1 es encontrar un algoritmo genérico, es decir, independiente de la aplicación.

En realidad, el estándar MPEG-1 se compone de cinco partes:

1. La sección de sistema: ISO 11172-1 (1993, última revisión 1999)
2. La sección de vídeo: ISO 11172-2 (1993, última revisión 1999).
3. La sección de audio: ISO 11172-3 (1993, última revisión 1996)
4. La sección de pruebas de conformidad: ISO 11172-4 (1995).
5. La sección de software de referencia: ISO 11172-5 (1998).

14.1. Sección de vídeo

Como se ha presentado anteriormente, MPEG-1 es el estándar utilizado para la codificación del formato VideoCD. La resolución de pantalla de un VCD es de 352x240 píxeles para el formato NTSC, o de 352x288 píxeles para el formato PAL, aproximadamente la cuarta parte de la resolución normal de televisión. El vídeo se almacena a una tasa de 1150 Kbps, el audio a 224 Kbps. Puesto que la tasa de bits conjunta (1374 Kbps) es similar a la tasa de bits de un CD de audio normal, la duración del vídeo que puede ser almacenado es similar a la de un CD de audio (un CD estándar puede almacenar aproximadamente 74 minutos de vídeo en formato VCD).

En MPEG-1, tanto el tamaño de la imagen como la frecuencia de cuadro son variables dentro de un rango. Existe un conjunto de límites de esos valores que forman un subconjunto del margen total posible.

Máximo número de píxeles / línea	720 píxeles
Máximo número de líneas / imagen	576 líneas
Macrobloques por cuadro	396 macrobloques

Macrobloques por segundo	396x25 = 330x30 macrobloques
Frecuencia de cuadro	30 cuadros/s
Tasa binaria	1,86 Mbps
Buffer del descodificador	376.832 bits

Para lograr una mayor compresión, MPEG-1 explota la redundancia temporal de las escenas. Cada frame en una secuencia de vídeo suele variar poco con respecto a los anteriores y siguientes, y es precisamente este parecido lo que se pretende emplear para transmitir una menor cantidad de información.

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

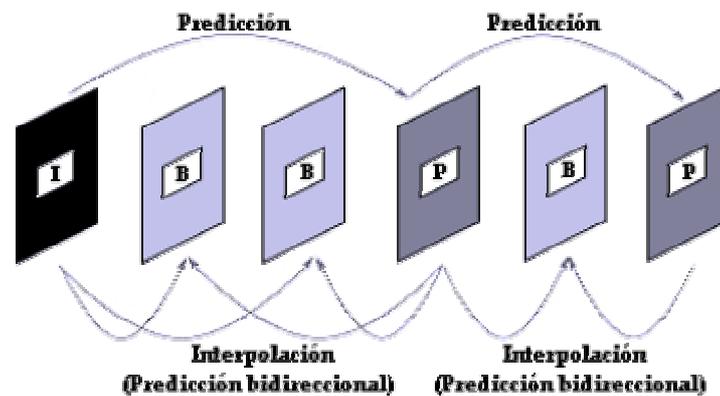


Ilustración [23] Imágenes MPEG

14.1.1. Las imágenes I (Intra).

Son imágenes que no requieren información adicional para su decodificación. Son codificadas sin ninguna referencia a otras imágenes (codificación intracuadro), de la misma manera que 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.

14.1.2. Las imágenes P (previstas).

Se codifican con respecto a las imágenes de tipo I o P anteriores (codificación intercuadro), 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.

14.1.3. Las imágenes B (Bidireccionales)

Se codifican por interpolación entre dos imágenes de tipo I o P precedentes y siguiente que las enmarcan (codificación intercuadro). 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 alta, 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.

Se definen asimismo los parámetros M y N que especifican 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.

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.

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 siguiente.

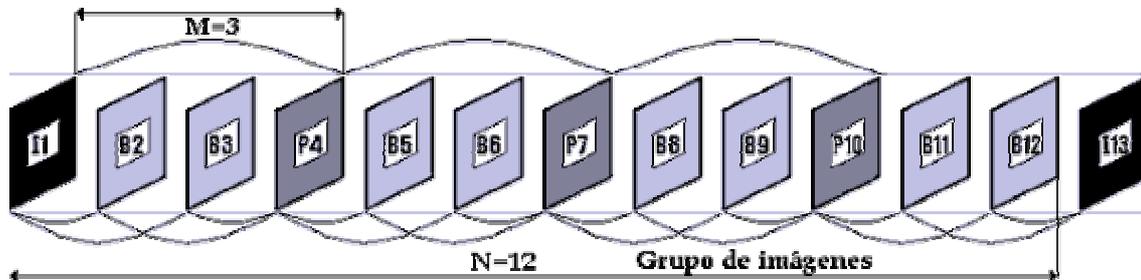


Ilustración [24] Grupo de Imágenes MPEG. $M=3$, $N=12$

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 captaron.

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

...**I** **B** **B** **P** **B** **B** **P** **B** **B** **P** **B** **B** **I** **B** **B** **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, como se ve en la figura:

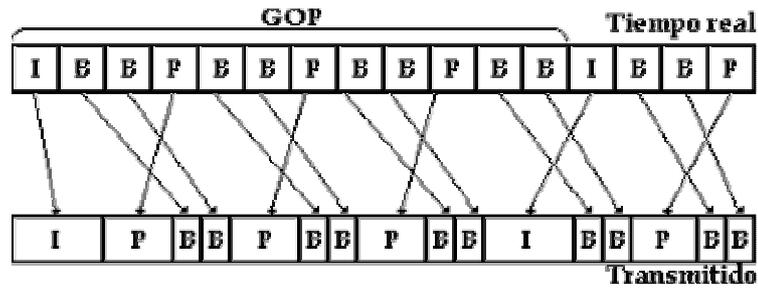


Ilustración [25] Comparación antes y después de la compresión

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 siguiente 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).

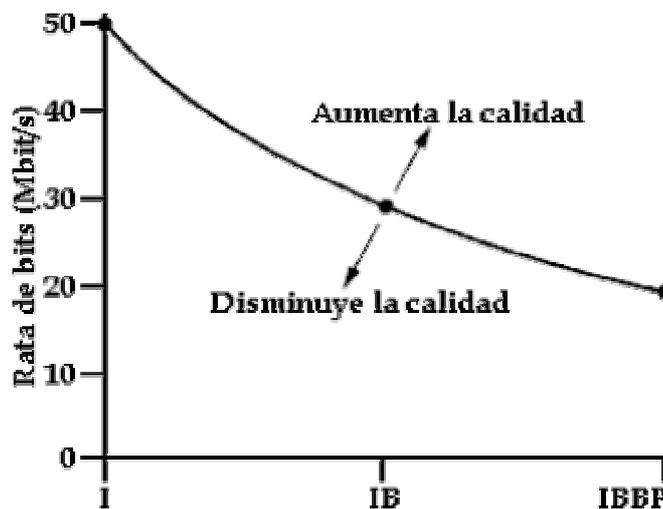


Ilustración [26] Curva de calidad constante

14.1.4. 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 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 siguiente figura muestra la construcción de una secuencia de video MPEG constituida por capas bien definidas.

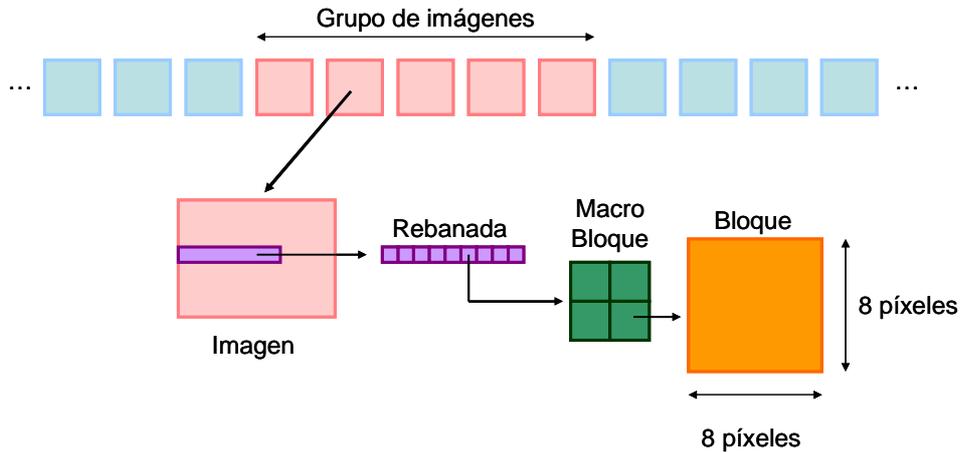


Ilustración [27] Descomposición de una secuencia MPEG

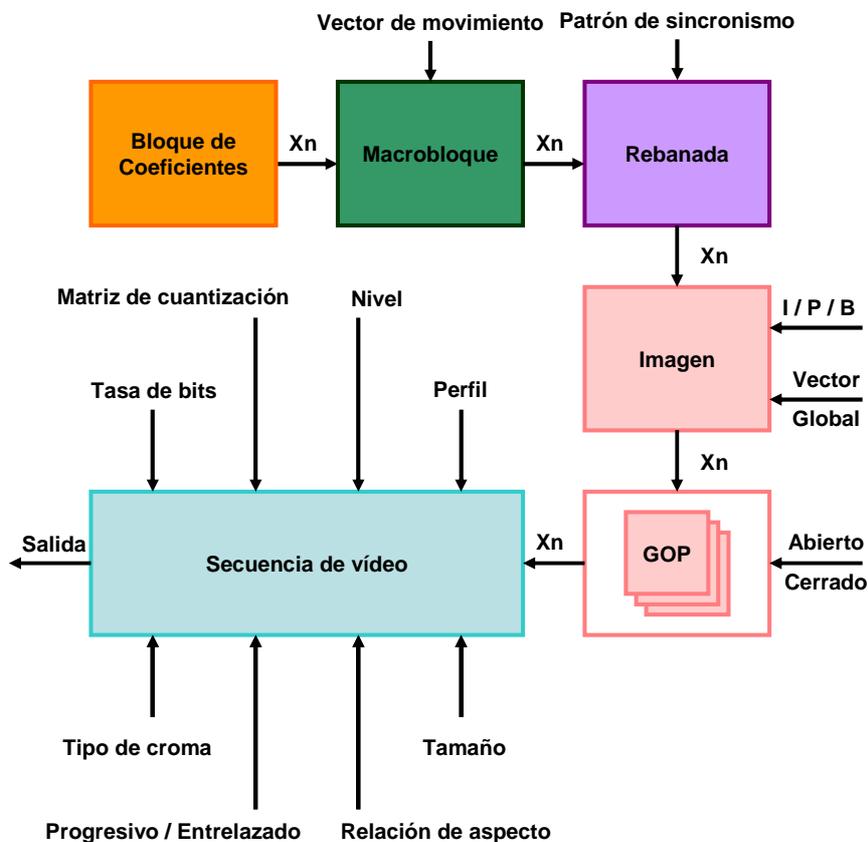


Ilustración [28] Descomposición de una secuencia MPEG

A continuación se presentan cada una de las unidades que componen una secuencia de vídeo en formato MPEG-1

14.1.4.1. Bloque (Block)

Es la unidad fundamental de la información de la imagen y esta representada por un bloque de coeficientes transformados DCT, que tienen un tamaño de 8x8 píxeles, los cuales representan valores de luminancia y croma.

Aquí, el coeficiente DC es enviado el primero, ya que este representa con mayor precisión la información de este bloque (corresponde con el nivel de gris de la imagen). Los demás coeficientes se envían después de este.

14.1.4.2. 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, según esté indicado. También se indica la escala utilizada para la recuantificación de los coeficientes. 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 añadida 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 16x16 píxeles (referidos a la componente de luminancia, 4 bloques Y), y dos bloques con información de color.

14.1.4.3. Rebanada (Slice)

Los macrobloques son reunidos en rebanadas, y estas 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 el ATSC (Advance Televisión Systems Committee) establece que 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.

14.1.4.4. Imagen 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. Se puede enviar un vector global para toda la imagen, y luego se pueden enviar vectores individuales que lleguen a crear la diferencia en el vector global.

14.1.4.5. 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 se puede colocar un número variable de imágenes P y/o B 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.

14.1.4.6. 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 tasa de imágenes, si se usa un barrido progresivo o entrelazado, el perfil, nivel, velocidad de transferencia de bits, y qué 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 espectador está cambiando de canal.

Resumiendo todo lo expuesto, un codificador MPEG-1 podría definirse como "un sistema de codificación de una señal digital de vídeo por componentes, con codificación híbrida formada por codificación de transformada intracuadro (imágenes I) y codificación diferencial con compensación de movimiento intercuadro (imágenes P,B), que utiliza control de buffer para mantener una tasa binaria constante".

14.1.5. Codificación intracuadro

Para la codificación intracuadro (imágenes Intra), se implementa un algoritmo cuya función es reducir el tamaño total del archivo, comprimiendo todas las imágenes I que componen el video de forma independiente, sin tener en cuenta la información de otros fotogramas.

El algoritmo que se utiliza para comprimir las imágenes I es el que sigue el formato JPEG (Joint Photographic Experts Group, 1986) con pérdidas. Esto significa que al descomprimir la imagen no obtenemos exactamente la misma imagen que teníamos antes de la compresión. Una de las características que hacen muy flexible el JPEG es el poder ajustar el grado de compresión. Si especificamos una compresión muy alta se perderá una cantidad significativa de calidad, pero, obtendremos ficheros de pequeño tamaño. Con una tasa de

compresión baja obtenemos una calidad muy parecida a la del original, y un fichero mayor.

Esta pérdida de calidad se acumula, lo que quiere decir que si se comprime una imagen y la se descomprime, obtendremos una calidad de imagen, pero si se vuelve a comprimir y descomprimir una segunda vez obtendremos una pérdida mayor. Cada vez que esto suceda la imagen perderá algo de calidad.

El algoritmo de compresión JPEG se basa en dos defectos visuales del ojo humano. Uno es el hecho de que el ojo humano es mucho más sensible al cambio en la luminancia que en la crominancia, es decir, notamos más claramente los cambios de brillo que de color. El otro es que notamos con más facilidad pequeños cambios de brillo en zonas homogéneas que en zonas donde la variación es grande, por ejemplo en los bordes de los objetos.

El proceso que tiene lugar para la compresión de una imagen JPEG es:

1. Transformación del espacio de color
2. Submuestreo
3. Transformación discreta del coseno
4. Cuantización
5. Codificación entrópica
6. Decodificación

14.1.5.1. Transformación del espacio de color

El proceso comienza convirtiendo la imagen desde su modelo de color RGB al modelo YUV ó YCbCr. El modelo YUV define un espacio de color en términos de una componente de luminancia y dos componentes de crominancia, una para la saturación y otra para la tonalidad del color. Esto es así porque YUV está más próximo al modelo humano de percepción que el estándar RGB.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Los coeficientes de la matriz de paso se eligen de acuerdo a las características subjetivas del ojo humano. Son ecuaciones que están en constante cambio, pero la matriz arriba expuesta con esos coeficientes concretos es la más difundida.

14.1.5.2. Submuestreo

Una opción que se puede aplicar al guardar la imagen es reducir la información del color respecto a la de brillo (debido al defecto en el ojo humano comentado anteriormente). Hay varios métodos para reducir la cantidad de información del color. Si, por ejemplo, este paso no se aplica, la imagen sigue en su espacio de color YUV, (esto se entiende como un submuestreo 4:4:4, 4 componentes de luminancia por cada 4 componentes de saturación y hue), con lo que la imagen no sufre pérdidas. Puede reducirse la información cromática a la mitad, por ejemplo 4:2:2, mientras el brillo (luminancia) sigue intacto. Otro método muy utilizado en el submuestreo es el 4:2:0. Si la imagen de partida estaba en escala de grises (blanco y negro), puede eliminarse por completo la información de color, quedando como 4:0:0.

14.1.5.3. Transformación discreta del coseno

Después del submuestreo, cada componente de la imagen (YUV) se divide en pequeños bloques de 8×8 píxeles, que se procesan de forma independiente (de esto resulta la formación de los bloques, que se hace notable en imágenes guardadas con altas compresiones).

Hay que tener en cuenta que no todas las imágenes van a tener dimensiones divisibles entre 8. Para solventar esto, tendremos que realizar dos operaciones, una para la dimensión "x" y otra para la dimensión "y". Para el caso de que la anchura de la imagen no sea divisible entre 8, debemos completar la imagen con las columnas más a la derecha de la imagen original. Cuando la anchura ya es divisible entre 8, debemos hacer la misma operación de antes para las filas, usando la filas inferiores de la imagen original. De este modo ya tenemos una imagen cuyas dimensiones horizontal y vertical son múltiplos de 8.

La aplicación del algoritmo de la transformada bidimensional DCT tiene como objetivo la reducción de las componentes de alta frecuencia, para compactar más la energía.

$$DCT_{UV} = \frac{1}{\sqrt{2N}} C_U C_V \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p_{xy} \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right], \text{ donde}$$

$$C_U C_V = \begin{cases} 1/2 & \text{si } u, v = 0. \\ 0 & \text{eoc} \end{cases}$$

A modo de ejemplo, se aplica la transformación DCT al siguiente bloque:

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

El resultado de esta transformación es:

1260	-1	-12	-5	2	-2	-3	1
-23	-17	-6	-3	-3	0	0	-1
-11	-9	-2	2	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	1	2	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

14.1.5.4. Cuantización

Como ya habíamos comentado, el ojo humano es muy bueno detectando pequeños cambios de brillo en áreas relativamente grandes, pero no cuando el brillo cambia rápidamente en pequeñas áreas (variación de alta frecuencia), esto permite eliminar las altas frecuencias, sin perder excesiva calidad visual.

Esto se realiza normalizando cada componente en el dominio de la frecuencia mediante la aplicación de una matriz de cuantización definida por el usuario, dividiendo cada elemento del bloque transformado por el mismo elemento en la matriz de cuantización, y posteriormente redondeando al número entero más cercano.

Si elegimos como matriz de cuantización:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

El ejemplo anterior resulta:

79	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Este es el proceso en el que se pierde la mayor parte de la información (y calidad) cuando una imagen es procesada por este algoritmo. El resultado de esto

Es ahora cuando el vector resultante se codifica mediante codificación de Huffman o codificación aritmética (esta última con mejores prestaciones que la de Huffman, pero que rara vez se usa, ya que está cubierta por patentes). Esta compresión produce archivos un 5% menores a costa de un mayor tiempo de codificación y decodificación. Puede emplearse también en aplicar un menor grado de compresión a la imagen, y obtener más calidad para un tamaño parecido.

El resultado tras el proceso de codificación entrópica puede variar en función de la agresividad de los divisores de la matriz de cuantización. A mayor valor de esos divisores, más coeficientes se convierten en ceros, y más se comprime la imagen. Pero mayores compresiones producen mayor ruido, empeorando la calidad de la imagen. Una imagen con una fuerte compresión (1%-15%) puede tener un tamaño de archivo mucho menor, pero tendrá tantas imperfecciones que no resultará interesante.

14.1.5.6. Decodificación

El proceso es similar al seguido hasta ahora, sólo que de forma inversa. En este caso, al haber perdido información, los valores no coincidirán.

Se coge la información de la matriz, se decodifica, y se pone cada valor en su casilla correspondiente. Después se multiplica cada uno de estos valores por el valor correspondiente de la matriz de cuantización usada, como muchos valores son ceros, sólo se recuperan (y de forma aproximada) los valores de la esquina superior izquierda.

Después se deshace la transformación DCT.

14.1.6. Codificación diferencial con compensación del movimiento intercuadro.

La compensación del 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 siguiente 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.

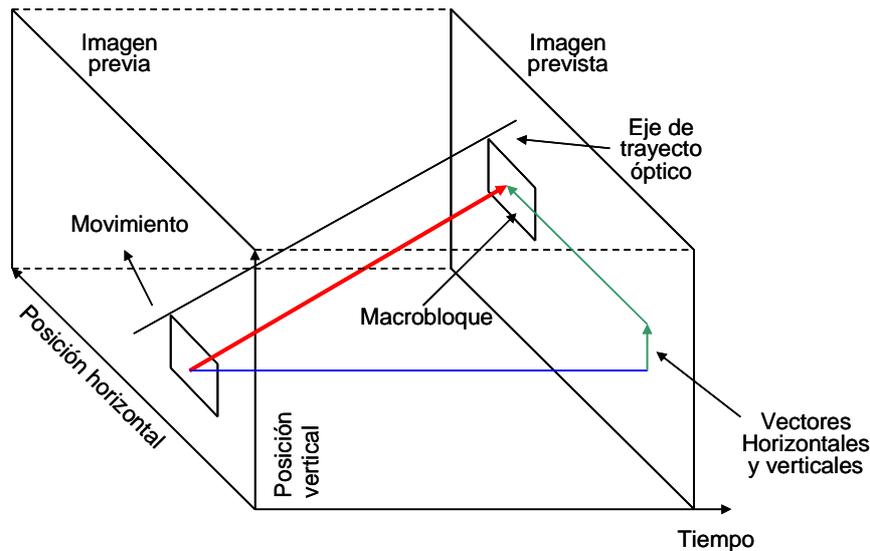


Ilustración [30] Movimiento en imágenes MPEG

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, 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.

La predicción de movimiento significa predecir un bloque de valores de los píxeles en el cuadro siguiente usando un bloque en cuadro actual. La diferencia de la localización entre estos bloques se llama Vector Movimiento, y la diferencia entre dos bloques se llama error de la predicción. En MPEG-1, el codificador debe calcular el vector del movimiento y el error de la predicción. Cuando el decodificador obtiene esta información, puede utilizarla junto cuadro actual para reconstruir el cuadro siguiente. Llamamos generalmente este proceso como

Compensación de Movimiento. En general, la compensación de movimiento es el proceso inverso de la estimación de movimiento.

• Motion Estimation

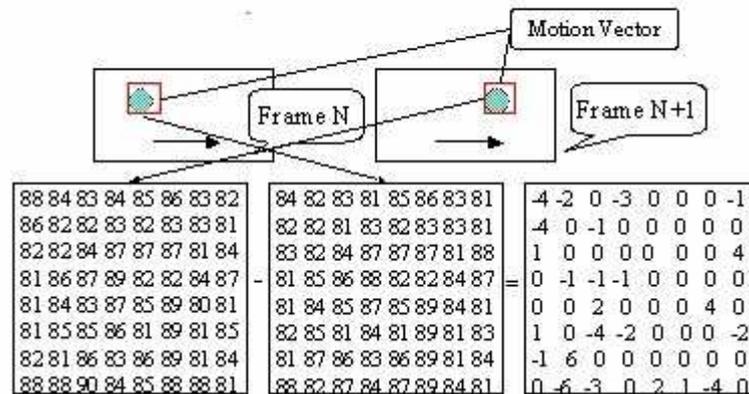


Ilustración [31] Estimación del movimiento

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. Esta es la manera de codificar una imagen P.

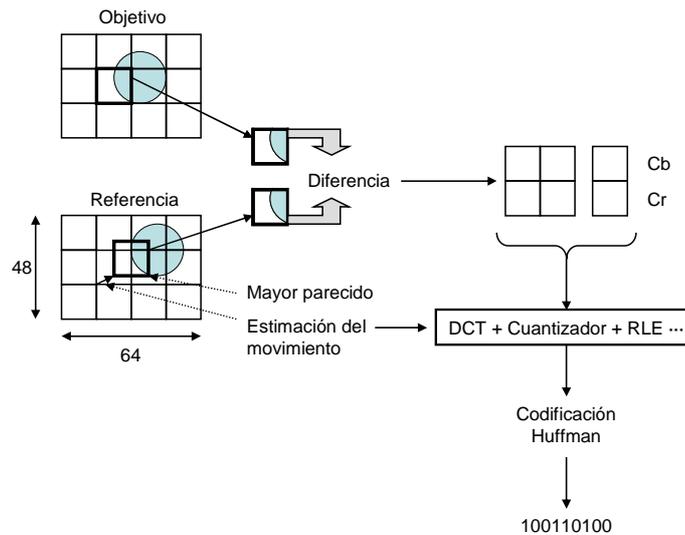


Ilustración [32] Estimación del movimiento. Codificación.

La predicción basada compensación de movimiento explota la redundancia temporal. Gracias a que los marcos se relacionan de cerca, es posible representar exactamente o "predecir" los datos de un bloque basado en los datos de la imagen de referencia, con tal que se estime el desplazamiento. El proceso de la predicción ayuda a una reducción de importante de bits. En los cuadros P, cada macrobloque de 16x16 es predicho por un macrobloque de un cuadro previamente codificado I. En un objeto móvil, los macrobloques en los dos cuadros sucesivos generalmente corresponden a una misma localización espacial. Por lo tanto, se realiza una búsqueda en el cuadro I para encontrar el macrobloque que empareje de cerca el macrobloque bajo consideración en el cuadro P. La diferencia entre los macrobloque es el error de la predicción. Este error se puede cifrar en el dominio de la DCT. La DCT del error da lugar a pocos coeficientes de alta frecuencia, que después del proceso del cuantización requieren un número pequeño de bits para la representación. Las matrices del cuantizador para los bloques de predicción de error de la son diferentes las usadas en el bloque Intra, debido a la naturaleza distinta de su espectro de frecuencia. A los desplazamientos en las direcciones horizontal y vertical del macrobloque que mejor se adapta se le llama vectores del movimiento. Se utiliza la codificación diferencial porque reduce el requisito total de bits, transmitiendo la diferencia entre los vectores del movimiento de los cuadros consecutivos. Finalmente se utiliza codificación RLE y codificación Huffman.

Ejemplo: El macrobloque oscurecido en marco actual no tiene un buen ajuste con el marco anterior, sino que encontrará un mejor ajuste en el cuadro siguiente.

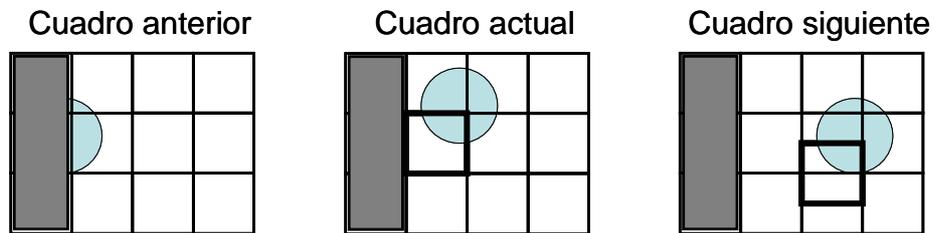


Ilustración [33] Justificación de las imágenes tipo B

Este es el caso en que se hace necesario los cuadros B o bidireccionales. Como se ve en la figura siguiente se muestra como se referencia la imagen predictiva con los vectores movimiento de los cuadros anterior y posterior, promediando el resultado.

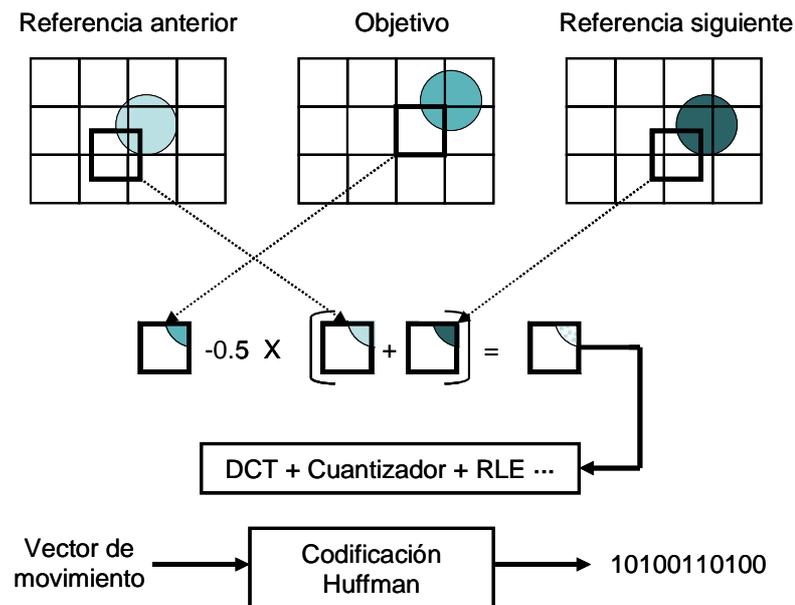


Ilustración [34] Codificación de las imágenes tipo B.

14.1.7. Control de buffering

El control de buffering trata de mantener una velocidad de reproducción constante. Las condiciones que debe cumplir la tasa binaria generada se establecen mediante un controlador del buffer de vídeo (VBV, Video Buffer Verifier). Este buffer tiene un tamaño determinado y es examinado a intervalos dados por la frecuencia de cuadro. En cada uno de esos momentos debe hacer al

menos un cuadro completo en el buffer. Ese cuadro se considera instantáneamente extraído del buffer, y justo después de la extracción debe haber siempre sitio para bits.

14.2. Sección de audio

El estándar MPEG-1 define tres capas (layers) o niveles de complejidad de la codificación de audio MPEG:

1. MP1 o MPEG-1 Parte 3 Capa 1 (MPEG-1 Audio Layer 1).
2. MP2 o MPEG-1 Parte 3 Capa 2 (MPEG-1 Audio Layer 2).
3. MP3 o MPEG-1 Parte 3 Capa 3 (MPEG-1 Audio Layer 3).

Las principales diferencias entre las tres capas son:

- El incremento en la complejidad del codificador y el decodificador (Capa I más simple, Capa III más compleja), especialmente cierto para el codificador de la Capa III.
- El mayor tiempo de respuesta al hacer la evaluación del codificador-decodificador (codec). La Capa III consume más tiempo para realizar el proceso de generar y/o leer los flujos de bits codificados.
- Y el incremento en el desempeño (Capa I más bajo, Capa III más alto desempeño; teniendo en cuenta que la tasa de bits permanece constante). Es decir, el audio codificado a 96 Kbps usando el esquema de la Capa III tiene más calidad que si se codifica a 96 Kbps con el esquema de la Capa II, o con el de la Capa I.

El algoritmo que realiza la codificación del audio en MPEG-1 sigue el siguiente esquema:

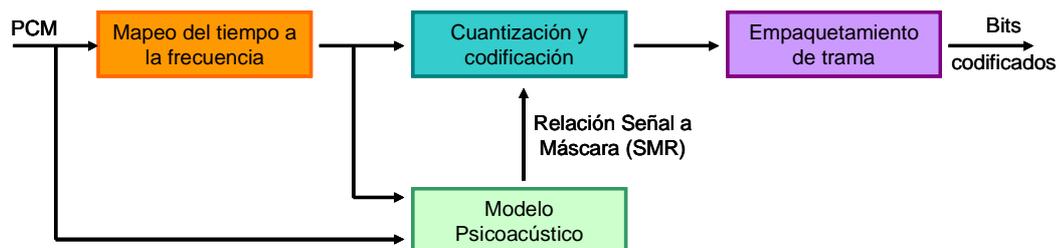


Ilustración [35] Codificación de audio en MPEG-1

14.2.1. MPEG-1 Capa I.

A continuación se detallan cada uno de los procesos para la codificación de audio en MPEG-1 Capa I.

14.2.1.1. Mapeo del tiempo a la frecuencia

Para la Capa I, el bloque de mapeo del tiempo a la frecuencia es un banco de filtros polifásico basado en la DCT, que divide el audio en 32 subbandas igualmente espaciadas en frecuencia (es equivalente a implementar filtros pasobanda), cada una aportando 12 muestras para un total de 384 muestras.

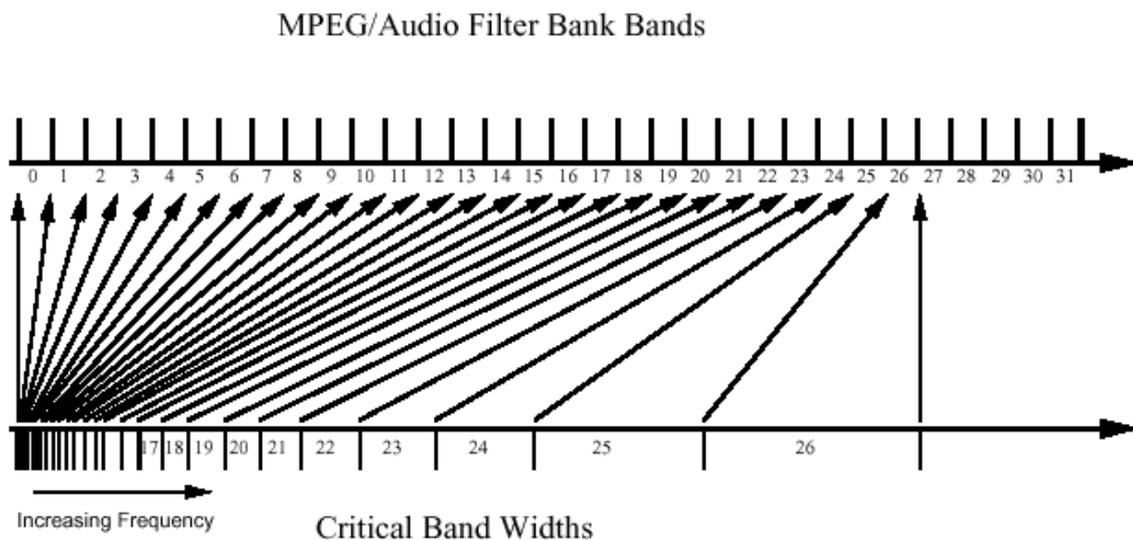


Ilustración [36] 32 Subbandas igualmente espaciadas en frecuencia

14.2.1.2. Modelo psicoacústico

El modelo psicoacústico determina qué frecuencias no serán perceptibles y así elimina la información referente a esas bandas.

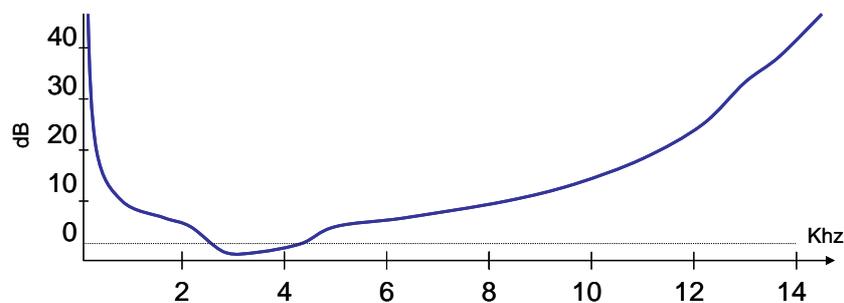


Ilustración [37] Respuesta en frecuencia del oído humano

El modelo psicoacústico sólo usa enmascaramiento en frecuencia, por medio de una FFT de 512 puntos. La salida de la FFT se usa para encontrar tanto el enmascaramiento tonal (sinusoidal) como el no-tonal (ruido) de la señal. Cada componente de enmascaramiento produce un umbral de enmascaramiento dependiente de su frecuencia, intensidad y tonalidad. Para cada subbanda, los umbrales de enmascaramiento individuales se combinan para formar el umbral de enmascaramiento global. El umbral de enmascaramiento se compara con el máximo nivel de señal para la subbanda, produciendo una relación señal a máscara (SMR, Signal-to-Mask Ratio), que es la entrada al siguiente bloque (cuantizador).

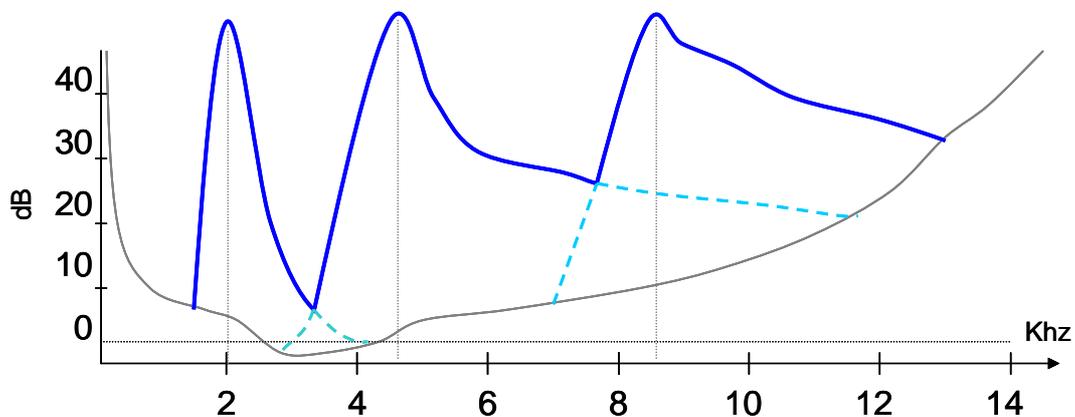


Ilustración [38] Enmascaramiento en frecuencia para tres tonos

14.2.1.3. Cuantizador / Codificador

El cuantizador/codificador primero examina las muestras de cada subbanda, encontrando el valor máximo absoluto de esas muestras, y realizando la cuantización con seis bits. Estos seis bits son el factor de escala para la subbanda. Luego se determina la repartición de bits para cada subbanda minimizando la SMR total con respecto a los bits repartidos para cada subbanda. Por último, las muestras subbanda son linealmente cuantizadas teniendo en cuenta la repartición de bits para esa subbanda.

14.2.1.4. Empaquetamiento de trama

El empaquetamiento de trama se encarga de formar un flujo MPEG válido. Cada trama comienza con información del encabezado para sincronización y

control, además de un CRC (Código de Redundancia Cíclica) para detección y corrección de errores. Cada una de las 32 subbandas usa cuatro bits para describir la repartición de bits (nivel de cuantización) y seis bits para el factor de escala. Los bits restantes en la trama se usan para las muestras subbanda, con un espacio opcional para información extra.

A 48 KHz., cada trama en la Capa I lleva información de 8 ms de audio. La calidad más alta se logra a 384 Kbps. Aplicaciones típicas de la Capa I incluyen grabación digital en cintas, discos duros o discos magneto-ópticos, que soportan la alta tasa de bits.

14.2.2. MPEG-1 Capa II.

El mapeo del tiempo a la frecuencia es el mismo de la Capa I, un banco de filtros polifásico tipo DCT con 32 subbandas; pero usa tres conjuntos de muestras en el filtro (anterior, actual, próximo) logrando 36 muestras por subbanda para un total de 1152 muestras subbanda.

El modelo psicoacústico es similar al de la Capa I, pero usa una FFT de 1024 puntos para una mejor resolución en frecuencia, y añade un poco de enmascaramiento temporal. El procedimiento es igual al de la capa anterior para producir las SMR para cada una de las 32 subbandas.

El cuantizador/codificador genera factores de escala de seis bits para cada subbanda, igual que en la Capa I. Sin embargo, las tramas de la Capa II son tres veces más largas que aquellas de la Capa I, así que la Capa II permite tres factores de escala sucesivos para cada subbanda, y el codificador usa 1, 2 o los 3 factores, dependiendo de cuánto difieran entre sí. De esta manera se logra, en promedio, reducir a la mitad la cantidad de bits que se usan para los factores de escala, en comparación con el consumo de bits para los factores de escala usado con el esquema de la Capa I. La repartición de bits se computa de manera similar a la repartición que se realiza en la Capa I.

El empaquetamiento de trama usa la misma estructura de encabezado y CRC de la Capa I. Sin embargo, el número de bits usados para describir la repartición de bits varía con la subbanda: 4 bits para las subbandas bajas, 3 bits para las subbandas medias y 2 bits para las subbandas altas (debido a que se tiene en

cuenta el ancho de las bandas críticas). Los factores de escala (1, 2 ó 3 dependiendo de los datos) se codifican junto con un código de 2 bits que describe cuál combinación de factores de escala se está usando. Las muestras subbanda son cuantizadas de acuerdo a la repartición de bits, y luego combinadas en grupos de tres (llamados gránulos). Cada gránulo se codifica con un código especial. Esto permite a la Capa II capturar mucha más información irrelevante en la señal, en comparación con la Capa I.

A 48 Khz., cada trama en la Capa II lleva 24 ms de audio. La calidad más alta se alcanza a tasas de 256 Kbps, aunque a 64 Kbps tiene un nivel aceptable. Las aplicaciones de esta capa incluyen radiodifusión de audio, grabación profesional y multimedia.

14.2.3. MPEG-1 Capa III.

La Capa III es mucho más compleja que la Capa II. Usa un banco de filtros híbrido conmutado, conformado por un filtro polifásico DCT similar al de la capa II y por una transformación discreta del coseno modificada, que ayuda a incrementar la resolución en frecuencia, permitiendo dividir el audio en bandas que se ajustan a las bandas críticas del oído (no hay igual espaciamiento entre subbandas).

$$X(m) = \sum_{k=0}^{n-1} f(k)x(k) \cos \left(\frac{\pi}{2n} \left(2k + 1 + \frac{n}{2} \right) (2m + 1) \right)$$

$$f(x) = \sin \left(\pi \frac{x}{n} \right)$$

Ilustración [39] Transformada Discreta del Coseno Modificada

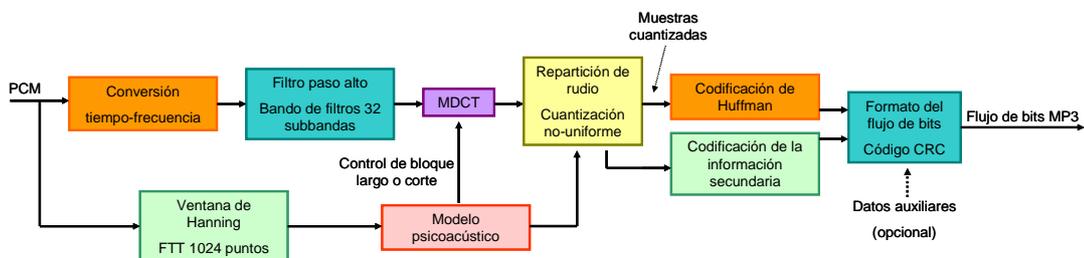


Ilustración [40] Codificación del audio MPEG-1 Capa III.

Emplea un modelo psicoacústico que incluye los efectos totales del enmascaramiento tanto en la frecuencia como en el tiempo. Utiliza un sofisticado esquema de codificación por entropía y cuantización no uniforme donde se involucran la redundancia estéreo y los códigos de Huffman, permitiendo crear tramas de longitud variable. El empaquetamiento de trama incluye el bit reservoir, que permite usar más bits en partes de la señal que lo necesiten. Además, permite alta calidad en el audio a tasas tan bajas como 64 Kbps.

14.3. Sección de sistema

Esta parte del estándar resuelve el problema de combinar uno o más flujos de datos de audio y vídeo, sincronizados para formar un único flujo. Una vez combinados, el flujo resultante es almacenado o transmitido.

La parte 1 de la norma ISO/IEC 11172 define así las reglas para la constitución de una capa de sistema (system layer) que agrupa el video, el audio y los datos privados en un solo tren, así como las condiciones sobre los trenes elementales que permitan asegurar esta combinación.

Las cinco funciones básicas de la capa de sistema que envuelve (o, mejor empaqueta) la capa de compresión son las siguientes:

- Sincronización de los trenes elementales comprimidos
- Combinación de estos trenes en un solo tren de datos.
- Inicialización de los buffers al comienzo de la reproducción.
- Gestión continua de estos buffers.
- Identificación del tiempo por marcadores (time stamps).

Un conjunto de codificación MPEG-1 de sistema debe, por tanto realizar la codificación de los datos de video y audio, y después el multiplexado de esta información y de los datos privados, añadiendo la información necesaria para su sincronización y otras relativas a los recursos necesarios para la decodificación del tren binario MPEG-1 (tales como el tamaño de las memorias intermedias necesarias para la decodificación de cada uno de los trenes elementales con un decodificador de frecuencia llamado System Target Decoder, STD).

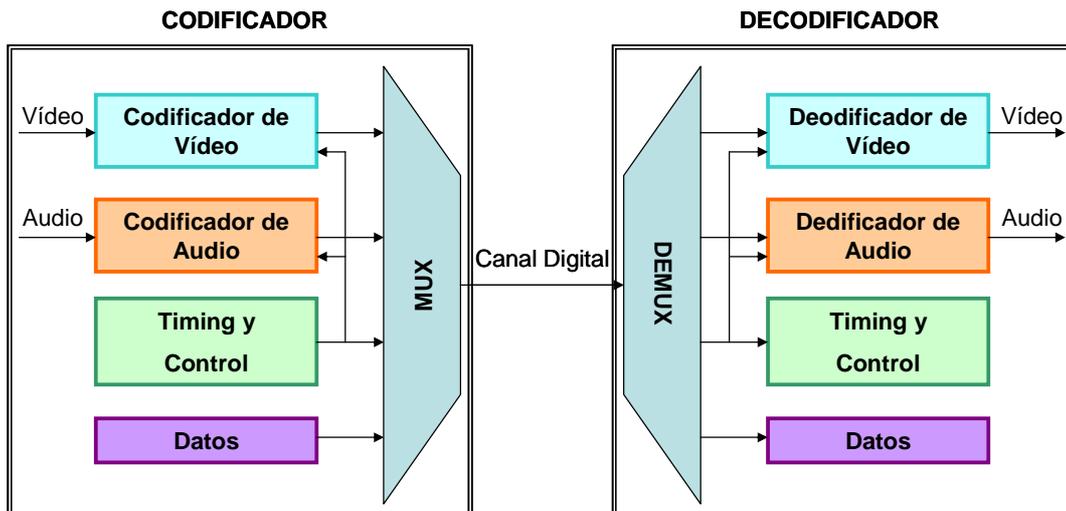


Ilustración [41] Codificación / Decodificación

14.3.1. Packetized elementary stream (PES)

Cada tren elemental se divide en paquetes que constituyen así un Packetized Elementary Stream (PES). Un paquete está constituido por una cabecera de paquete (packet header) seguido de los datos propiamente dichos.



Ilustración [42] Estructura de un paquete PES

La cabecera del paquete empieza por un código de inicio de 32 bits, donde los 8 últimos bits identifican el tipo (audio, video o datos privados) y el número de tren elemental al que el paquete de datos pertenece. La cabecera indica la longitud del paquete y el tamaño de la memoria intermedia de entrada necesaria para su decodificación por el STD.

Estructura de un paquete MPEG-1

Campo	Definición	Bits
start_code_prefix	prefijo del código de inicio (00 00 01 hex)	24
stream_id	tipo (4 MSB) y núm. (4 LSB) del PES	8

packet_length	longitud del PES (bytes restantes)	16
stuffing_bytes	bytes de relleno opcionales (valor FF hex)	0 a 16x8
bits "01"	principio del campo STD_buffer	2
STD_buffer_scale	factor de escala buffer (0=128 bytes, 1=1024 bytes)	1
STD_buffer_size	tamaño del buffer (múltiplos de 128 o 1024 bytes)	13
PTS (opcional)	marcador presentación (código de 4 bits + 33 bits + marcador de 3 bits)	40
DTS (opcional)	marcador decodificación (misma estructura que PTS)	40
packet_data_byte	datos (N=packet_length menos tamaño de los 6 campos sig.)	Nx8

También puede contener un marcador de decodificación (Decoding Time Stamp, DTS) que indique el momento de decodificación de la primera unidad de acceso (AU) del paquete y/o un marcador de presentación (Presentation Time Stamp, PTS) que indique el instante donde la unidad de presentación correspondiente (PU) debe ser "presentada" (visualizada o sonorizada, según proceda).

Por ejemplo, cuando una secuencia de video debe ser presentada en un orden de IBBP, estas son transmitidas con anterioridad en un orden de IPBB. Consecuentemente, deben usarse dos tipos de marcas de tiempo. El decodificador de marca de tiempo DTS indica el tiempo cuando una imagen debe ser codificada, mientras la presentación de marcas de tiempo PTS indica cuando esta debe ser presentada a la salida del decodificador.

Cuando una secuencia IPBB es recibida, las dos imágenes I y P deben ser decodificadas antes de presentar la primera imagen B. Como un decodificador sólo puede decodificar una imagen al tiempo; se debe decodificar una imagen I primero y luego debe ser almacenada, para también decodificar después una imagen P, con las cuales se pueden crear las imágenes B esperadas.

La figura siguiente muestra esto, primero se recibe una imagen I con sus respectivas marcas de tiempo DST y PTS, estas marcas de tiempo son separadas

un periodo de imagen. Si la codificación es bidireccional, una imagen P debe seguir y esta también vendrá con marcas de tiempo DTS y PTS, pero la separación entre las dos marcas de tiempo es de tres periodos de imagen, para permitir la inclusión de la imagen B. De este modo una secuencia IPBB es recibida. I es demorada un periodo de imagen, P es demorada tres periodos de imagen, las dos imágenes B no son demoradas y la presentación de la secuencia descodificada es IBBP. Claramente si la estructura de las GOPs es cambiada de tal forma que hayan mas imágenes B entre I y P las diferencias entre DTS y PTS de las imágenes P será muy grande.

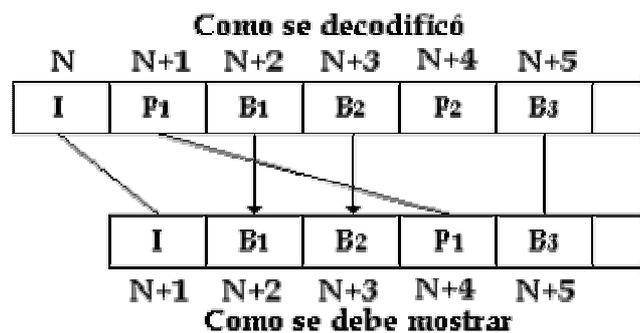


Ilustración [43] Uso de PTS/DTS para sincronizar una codificación bidireccional

En los paquetes de audio existen marcas de tiempo del tipo PTS, los paquetes de audio nunca son transmitidos fuera de secuencia, no hay marcas de tiempo del tipo DTS.

Estos marcadores sirven también para sincronizar las funciones de decodificación de audio y video, enviándose con relativa frecuencia (la norma especifica un intervalo máximo de 0.7 segundos).

Se codifican sobre 33 bits, que representan un tiempo absoluto expresado en ciclos de reloj de referencia a 90 KHz.

Un paquete puede contener a continuación un número variable de bytes de datos, en función de las características del soporte de grabación o de transmisión utilizado.

La norma prevé también la posibilidad de añadir un tren elemental de ajuste (padding stream) para obtener un flujo constante o alinearse sobre los sectores

físicos de un dispositivo de almacenamiento, así como los bytes de relleno (stuffing bytes) en el interior de un paquete (máximo 16 por paquete).

Los paquetes se agrupan en packs, cuya cabecera (pack header) proporciona la información de timing y de flujo por medio de los campos de referencia de reloj de sistema (System Clock Reference, SCR) sobre 33 bits y del flujo del múltiplex (mux_rate_field).

Los campos SCR se utilizan para sincronizar en el decodificador un reloj de sistema común a 90 KHz. (System Time Clock, STC) que sirve de base de tiempos y de unidad de medida para los campos DTS y PTS de los paquetes. Obsérvese esto en la figura siguiente.

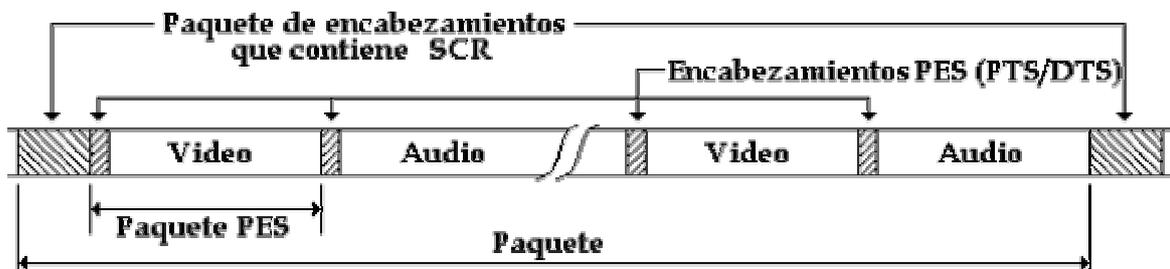


Ilustración [44] Un paquete es un conjunto de paquetes PES

La cabecera del paquete, cuyo formato se ilustra en el siguiente cuadro, comienza con un código de inicio de 32 bits.

Estructura de la cabecera de un paquete MPEG-1

Campo	Definición	Bits
pack_start_code	código de inicio del paquete (00 00 01 BA hex)	32
bits "0010"	principio del campo SCR	4
SCR [32...30]	reloj de referencia del sistema (3 MSB)	3
marker_bit	bit siempre a "1"	1
SCR [29...15]	reloj de referencia del sistema (15 bits de peso medio)	15
marker_bit	bit siempre a "1"	1
SCR [14...0]	reloj de referencia del sistema (15 LSB)	15

marker_bit	bit siempre a "1"	1
marker_bit	bit siempre a "1"	1
mux_rate	velocidad multiplex MPEG (múltiplo de 50 Bps)	22
marker_bit	bit siempre a "1"	1

El primer paquete de un tren MPEG-1 contiene obligatoriamente una cabecera de sistema (system header).

Se trata de un paquete especial que resume el conjunto de parámetros del sistema utilizados para el transcurso de este tren (flujo máximo, identificación de los trenes elementales de audio, video y datos, tamaño mínimo del buffer de entrada, etc.).

La cabecera de sistema opcionalmente puede ser respetada durante cualquier nuevo paquete en curso del tren MPEG-1 para facilitar el acceso a un punto cualquiera de este tren. Esta cabecera de sistema empieza también por un código de inicio de 32 bits. Ver Cuadro 16.

Estructura de la cabecera del sistema MPEG-1

Campo	Definición	Bits
system_header_start_code	código de inicio de cabecera de sistema	32
header_length	longitud de cabecera (Nº de bytes que quedan por seguir)	16
marker_bit	bit siempre a "1"	1
rate_bound	velocidad máx (mux_rate) de la secuencia	22
marker_bit	bit siempre a "1"	1
audio_bound	Nº de PES de audio del bitstream (0 a 32)	6
fixed_flag	indica velocidad fija ("1") o variable ("0")	1
CSPS_flag	"1" si bitstream utiliza el "constrained parameter set"	1
system_audio_lock_flag	"1" indica relación armónica entre STC y el audio	1

system_video_lock_flag	"1" indica relación armónica entre STC y F.imagen	1
marker_bit	bit siempre a "1"	1
video_bound	Nº de PES de video del bitstream (0 a 16)	5
reserved_byte	reservado para futuras ampliaciones	8
stream_ID1	identificación del primer PES (naturaleza y número)	8
bits "11"	bit de comienzo del campo STD_buffer	2
STD_buffer_bound_scale1	"0" =128 bytes (audio), "1"=1024 bytes (video)	1
STD_buffer_size_bound1	tamaño máx. del buffer (múltiplo de 128 o 1024 bytes)	13
stream_ID2	identificación del segundo PES (naturaleza y número)	8
bits "11"	bit de comienzo del campo STD_buffer	2
STD_buffer_bound_scale2	"0" =128 bytes (audio), "1"=1024 bytes (video)	1
STD_buffer_size_bound2	tamaño máx. del buffer (múltiplo de 128 o 1024 bytes)	13
etc. para (n-2) otros PES	ídem para todos los PES (máx. 16 video, 32 audio, 2 private)	(n-2)x24

El número de trenes elementales que pueden componer un tren MPEG-1 se especifica del siguiente modo:

- o video: 0 a 16
- o audio: 0 a 32
- o datos privados: 0 a 2

El tren MPEG-1 termina con un código final de 32 bits (00 00 01 B9).

La figura siguiente muestra de forma esquemática el contenido de un tren MPEG-1 completo.

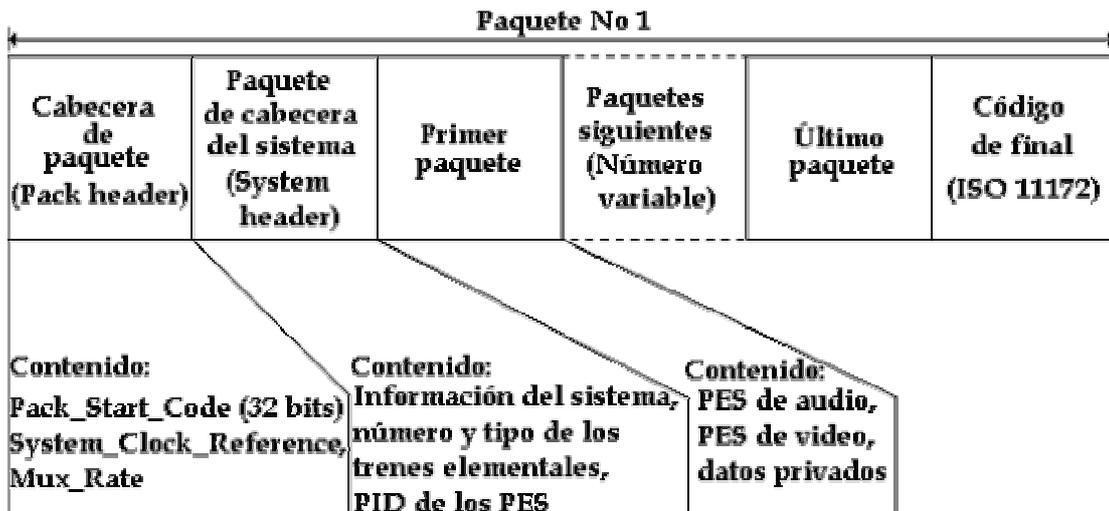


Ilustración [45] Contenido de un paquete MPEG-1

14.4. Sección de pruebas de la conformidad.

Especifica cómo se deben realizar las pruebas para comprobar que los bitstreams y los decodificadores se adaptan a los requerimientos de las partes 1,2 y 3 del estándar MPEG-1. Estas pruebas las pueden usar:

1. Fabricantes de codificadores y sus clientes, para comprobar que el codificador consigue bitstreams válidos.
2. Fabricantes de decodificadores y sus clientes, para verificar que el decodificador cumple con las especificaciones relatadas en las partes 1,2 y 3 del estándar.
3. Aplicaciones, para verificar que las características del bitstream cumplen las necesidades de la misma, por ejemplo que el tamaño de la imagen codificada no excede el valor máximo permitido por la aplicación.

14.5. Software de referencia.

Técnicamente no es un estándar, sino un informe técnico. Proporciona una completa implementación del software de las tres primeras partes del estándar MPEG-1. El código fuente no es de libre distribución.

15. ESTANDAR MPEG-2

La segunda fase de MPEG, llamada MPEG 2, consta de tres partes o estándares, cubiertas por

1. ISO/IEC 13818-1 Sistemas MPEG-2 (ITU-T Rec. H.222)
2. ISO/IEC 13818-2 Vídeo MPEG-2 (ITU-T Rec. H.262)
3. ISO/IEC 13818-3 Audio MPEG-2.

Estas fueron aprobadas finalmente como Estándar Internacional (IS) por la asamblea N° 29 de la ISO/IEC JTC1/SC29/WG11 (MPEG) hecha en Singapur en Noviembre de 1994. Después se fueron añadiendo más partes al estándar:

4. ISO/IEC 13818-4 Describe maniobras de pruebas de cumplimiento.
5. ISO/IEC 13818-5 Describe sistemas para simulación por Software.
6. ISO/IEC 13818-6 Describe extensiones para DSM-CC (Comando Digital de herramientas de almacenamiento y control)
7. ISO/IEC 13818-7 codificación avanzada de audio. (AAC)
8. ISO/IEC 13818-9 Extensión para interfaces en tiempo real.
9. ISO/IEC 13818-10 conformidad con extensiones para DSM-CC.

15.1. Sección de vídeo

El registro ITU-T H.262 trata con codificación de vídeo de alta calidad con posible vídeo entrelazado de NTSC, PAL o Televisión de Alta Definición (HDTV). Esto es un intento para operar en un rango de 2 a 15 Mbit/s. Sin embargo puede funcionar a velocidades superiores de 100 Mbit/s. Un amplio rango de aplicaciones, velocidades, resolución calidades de las señales y servicios son direccionados, incluyendo todas las formas de medios de almacenamiento digital, televisión (incluyendo HDTV), broadcasting y comunicaciones.

MPEG-2 puede describirse como una evolución de las 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.

Entre las varias mejoras o extensiones introducidas en los codificadores MPEG 2, tenemos:

- Nuevos modos de predicción de campos y tramas para scanning entrelazado.
- Cuantización mejorada.
- Nuevos códigos Intra-trama de longitud variable (VLC).
- Extensión escalada de resoluciones para compatibilidad, servicios jerárquicos y robustos, y
- Dos nuevas capas de sistema para multiplexión y transporte que provee celdas/paquetes de vídeo de alta o baja prioridad, cuando son llevados a través de una red conmutada.
- Incrementos soportados por accesos aleatorios.
- Soporte resistente para incremento de errores.
- Múltiples programas con un multiplexor (MPEG 1 no puede hacer esto, y esto fue un driver principal para el MPEG 2).

Al igual que el H.261 y JPEG, el estándar MPEG-2 es un esquema híbrido de compresión para imágenes en pleno movimiento que usa codificación inter-trama y codificación Intra-cuadro, y combina la codificación predictiva con la codificación con la transformada DCT 8x8. Al igual que en MPEG-1, la DCT elimina redundancia en la imagen a través de la compresión de la información contenida en 64 píxel. El cuantizador otorga los bits para los coeficientes DCT más importantes, los cuales son transmitidos

El concepto de MPEG-2 es similar al MPEG-1, pero incluye extensiones para cubrir un amplio rango de aplicaciones. La principal aplicación destinada durante el proceso de definición de MPEG 2 fue todas las transmisiones de vídeo con calidad de TV codificadas a velocidades entre 5 y 10 Mbit/s.

Sin embargo, la sintaxis del MPEG-2 ha sido descubierta para ser eficiente para otras aplicaciones como las de altas velocidades binarias y velocidades de muestreo (HDTV). La característica más relevante con respecto a MPEG-1 es la sintaxis para codificación eficiente de vídeo entrelazado.

Otras características más específicas (precisión 10 bit DCT DC, cuantización no-lineal, tablas VLC) son incluidas, y otorgan una mejoría notable en la eficiencia

de la codificación. Otra característica clave de MPEG-2 son las extensiones escalables, las cuales permiten la división de continuas señales de vídeo dentro de dos o más cadenas binarias codificadas, representando el vídeo en diferentes resoluciones, calidades (por ejemplo SNR), o velocidades

15.1.1. Perfiles y niveles MPEG-2

MPEG-2 es una recomendación muy compleja, tiene una larga variedad de combinaciones (sobre 10^6). Sin embargo, un reducido conjunto de combinaciones son definidas bajo "perfiles" y "niveles".

Dentro de los perfiles, son posibles una gran variación de. Por otra parte los niveles son un conjunto de derivaciones impuestas para los perfiles. La combinación de un perfil y un nivel produce una arquitectura muy bien definida para una cadena particular de bits. Los perfiles limitan la sintaxis (por ejemplo los algoritmos), mientras los niveles limitan los parámetros (velocidad de muestreo, dimensiones de las tramas, velocidad binaria codificada, etc.).

15.1.1.1. Niveles

Proveen un rango de calidades potenciales, definen los máximos y mínimos para la resolución de la imagen, muestras Y por segundo (luminancia), el número de capas de audio y vídeo soportados por los perfiles escalados, y la máxima velocidad binaria por perfil. A continuación una explicación resumida de cada uno de ellos:

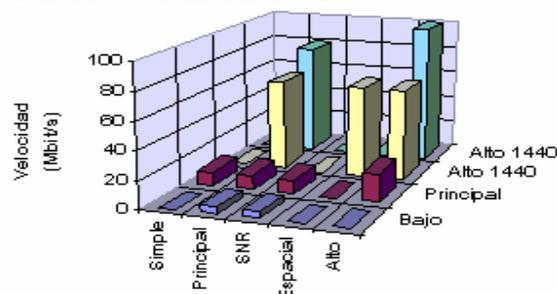
- Nivel Bajo: tiene un formato de entrada que es un cuarto de la imagen definida en el registro ITU-R 601.
- Nivel Principal: tiene una trama de entrada completa definida en el registro ITU-R 601.
- Nivel Alto 1440: tiene un formato de alta definición con 1440 muestras por línea.
- Nivel Alto: tiene un formato de alta definición con 1920 muestras por línea (para aplicaciones sin cualquier limitación en velocidades de datos).

15.1.1.2. Perfiles

Son definidos subconjuntos con características de sintaxis (por ejemplo: algoritmos), usados para converger la información. Hay cinco diferentes perfiles y cada uno es progresivamente más sofisticado y agrega herramientas adicionales (y por supuesto más costoso para el cliente) con la característica adicional de ser compatible con el anterior. Esto significa que un decodificador equipado con un alto perfil descodificará perfiles simples. A continuación una pequeña explicación de los perfiles:

- Perfil Simple: es el que ofrece pocas herramientas.
- Perfil Principal: tiene herramientas extendidas o mejoradas del perfil simple y predicción bidireccional. Tendrá mejor calidad para la misma velocidad binaria que el perfil simple.
- Perfil Escalable SNR y Perfil Escalable Espacial: son los próximos pasos. Estos dos niveles son llamados escalables porque ellos permitirán codificar datos de vídeo que sean particionados dentro de una capa base y una o más señales "Top-up". La señal Top-up puede tanto tratar la proporción S/N (SNR escalable) o la resolución (escalable espacial).
- Perfil Alto: este incluye todas las herramientas de las versiones anteriores y mejoradas. Tiene la habilidad de codificar diferencias de color entre líneas simultáneamente. Este es un súper sistema diseñado para aplicaciones donde no están contraídas sobre las velocidades de los bits.

Para muchas de las aplicaciones (incluyendo transmisión de satélites) el Perfil Principal, Nivel Principal (MP@ML) provee una buena relación entre calidad de imagen y la complejidad VLSI. El siguiente gráfico es un resumen o representa de todas las combinaciones entre perfiles y niveles. También muestra las máximas velocidades de datos para cada combinación:



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

	Simple	Principal	4:2:2	SNR	Espacial	Alto
Alto		4:2:0 1920x1152 80 Mbps				4:2:0 ó 4:2:2 1192x1152 100 Mbps
Alto 1440		4:2:0 1440x1152 60 Mbps			4:2:0 1440x1152 60 Mbps	4:2:0 ó 4:2:2 1440x1152 80 Mbps
Principal	4 : 2 : 0 750x576 15 Mbps Sin B	4 : 2 : 0 720x576 15 Mbps	4 : 2 : 2 720x608 50 Mbps	4 : 2 : 0 720x576 15 Mbps		4 : 2 : 0 ó 4 : 2 : 2 720x576 20 Mbps
Bajo		4:2:0 352x288 4 Mbps		4:2:0 352x288 4 Mbps		

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.

15.1.2. Modos escalables

Actualmente hay cuatro modos escalables en MPEG 2. Estos modos fragmentan el vídeo MPEG-2 en diferentes capas (base, media, y alta) para propósitos de priorización de datos de vídeo. Otro propósito de la escalabilidad es para divisiones complejas. Por ejemplo, en HDTV, la alta prioridad de la cadena binaria (720x480) puede ser descodificada bajo condiciones de ruido donde la baja prioridad (1440x960) no pueda.

A continuación una breve explicación de los modos escalables:

- Escalabilidad espacial: Este método de dominio espacial codifica la capa base a una dimensión de muestro bajo (por ejemplo: resolución) que las capas superiores. Las capas bajas (base) reconstruidas del muestro son usadas como predicción de las capas superiores.
- Particionamiento de datos: es un método de dominio de frecuencia que rompe los bloques de 64 coeficientes cuantizados de la transformada dentro de dos cadenas binarias. La primera, cadena de alta prioridad contiene los coeficientes más críticos de las frecuencias bajas e información (tales como valores DC, vectores, etc.), la segunda, cadena binaria de baja prioridad lleva datos AC de las altas frecuencias.
- Escalabilidad SNR: es un método de dominio espacial donde los canales son codificados a velocidades de muestreo idénticas, pero con diferentes calidades de imágenes. La cadena binaria de alta prioridad contiene datos de la capa base que pueden ser añadidos a la capa de refinamiento de baja prioridad para construir una imagen de alta calidad.
- Escalabilidad temporal: Un método de dominio temporal usado por ejemplo en vídeo estereoscópico. La primera, la cadena binaria de alta prioridad codifica vídeo a baja velocidad de tramas, y las tramas intermedias pueden ser codificadas en una segunda cadena binaria usando la reconstrucción de la primera cadena binaria como predicción. Por ejemplo en una visión estereoscópica, el canal de vídeo izquierdo puede ser predicho a partir del canal derecho.

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 siguiente muestra un codificador MPEG convencional, con coeficientes de cuantificación de elevado peso, que codifica una imagen con una moderada

SNR. Después esta imagen, al ser decodificada y sustraída de la imagen original píxel a píxel, 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 bits con ruido de cuantificación, pero un decodificador más complejo puede decodificar ambas imágenes con diferentes flujos de bits y combinarlos para producir una imagen de bajo ruido. Este es el principio del perfil SNR escalable.

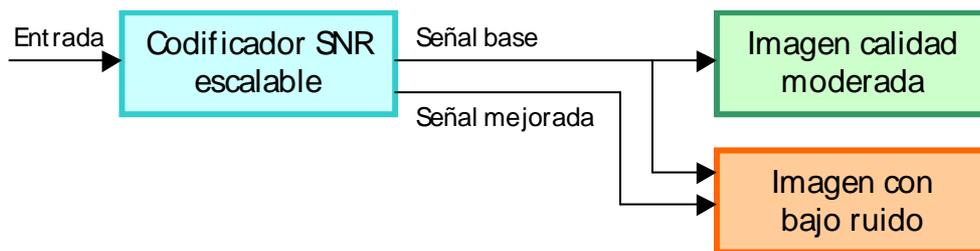


Ilustración [46] Codificador MPEG-2

Como otra alternativa, la figura inferior 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.

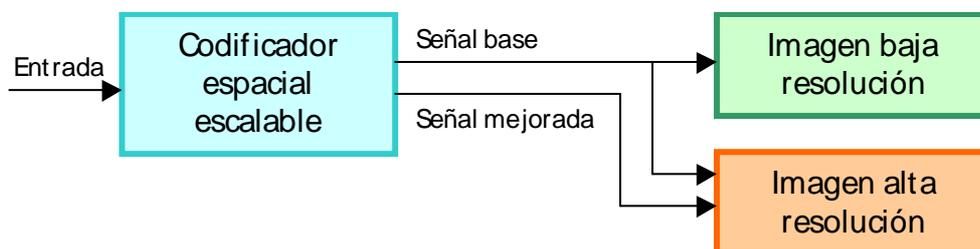


Ilustración [47] Codificador MPEG-2

El perfil high 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 píxeles x 576 líneas).
- El nivel high-1440 (alto-1440) está destinado a la HDTV (de hasta 1440 píxeles x 1152 líneas).
- El nivel high (alto) está optimizado para la HDTV (de hasta 1920 píxeles 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 siguiente).

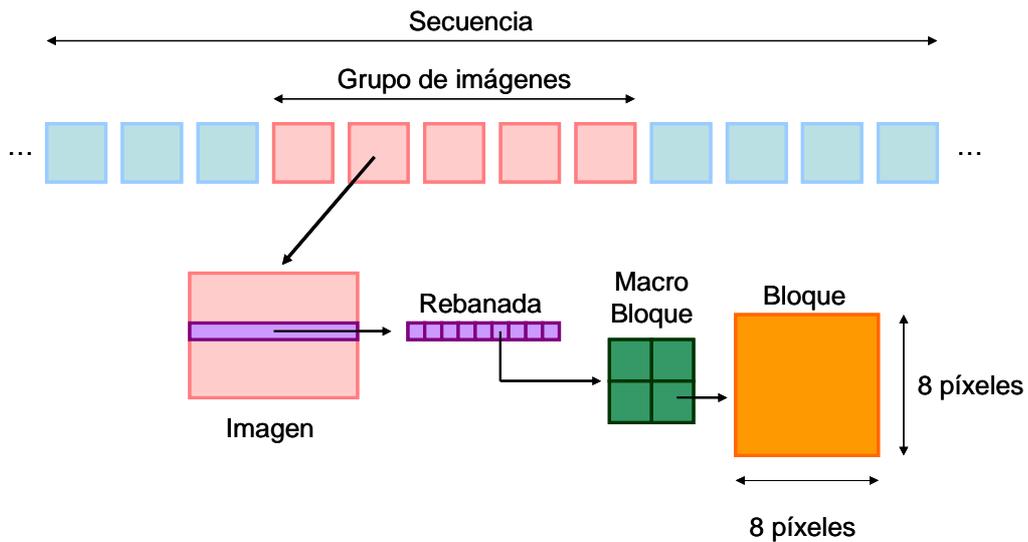


Ilustración [48] Capas MPEG-2

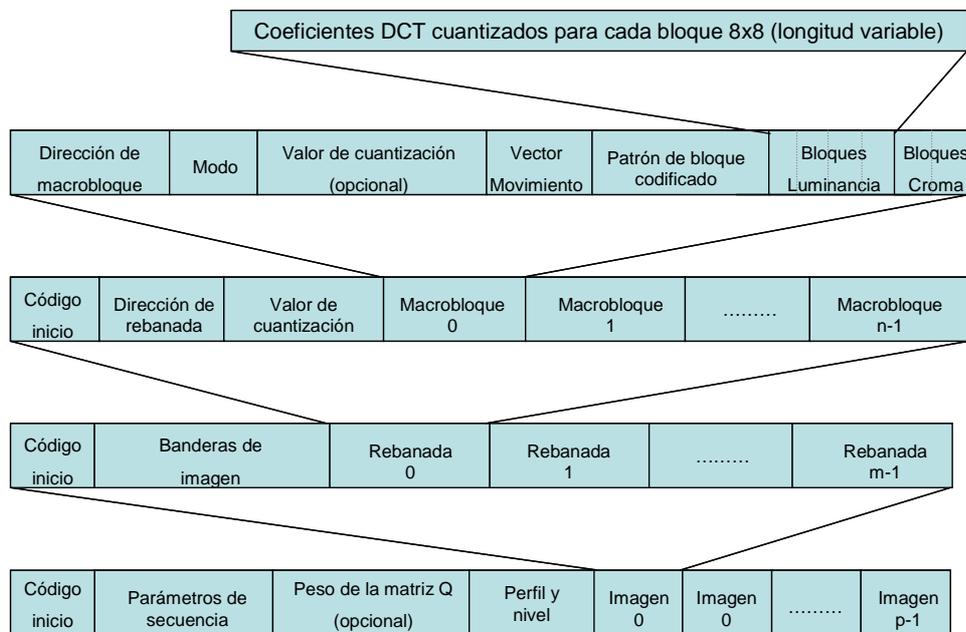


Ilustración [49] Secuencias MPEG-2

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

15.1.3. Modos de predicción específicos 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 siguiente representa la secuencia temporal de la posición vertical de las líneas de los campos sucesivos en un sistema entrelazado.

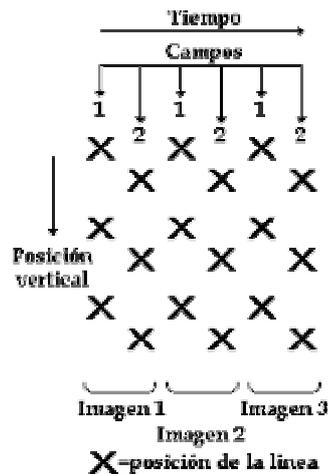


Ilustración [50] Entrelazado MPEG-2

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").

15.1.4. 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 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.

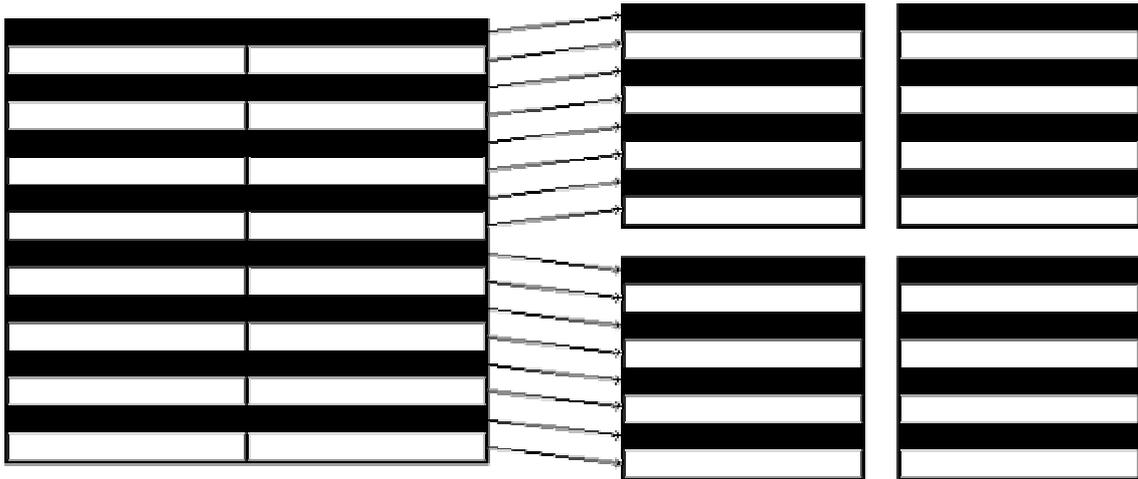


Ilustración [51] La estructura Frame en MPEG-2

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

15.1.5. 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. Verla figura siguiente

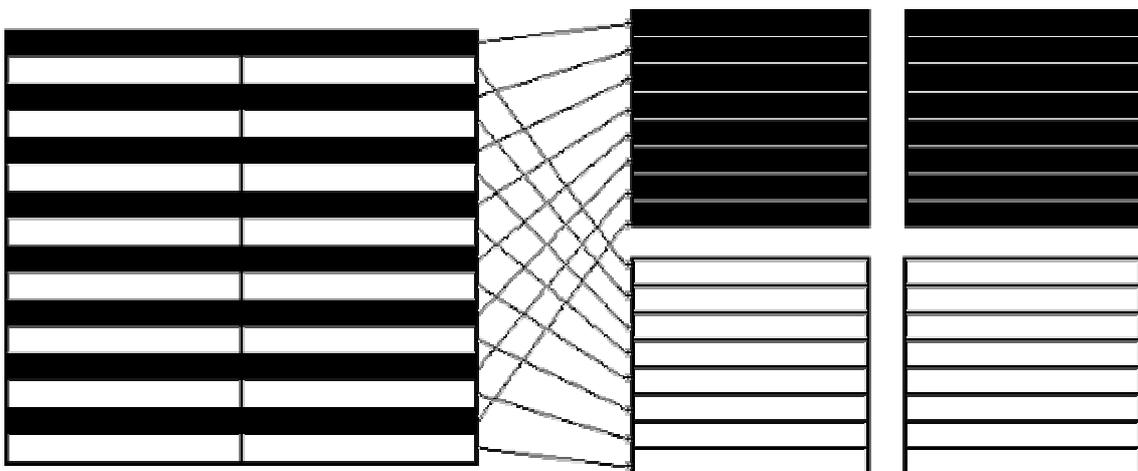


Ilustración [52] La estructura FIELD en MPEG-2

En cuanto a la estimación de movimiento, también hay varios modos previstos. Un macrobloque puede predecirse en modo "imagen", "campo" o "mixto".

15.1.5.1. 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).

15.1.5.2. 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.

15.1.5.3. El Modo Mixto

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

15.1.6. Scanning (Exploración)

Después de realizar la DTC a un bloque de 8x8 píxeles, se nota que generalmente los coeficientes más significativos de la DTC se encuentran en la parte superior izquierda de la matriz. Una vez evaluados, los coeficientes de menor valor pueden ser redondeados a cero. Permitiendo de este modo, una transmisión de datos más eficiente, debido a que los coeficientes no-cero son enviados primero, seguido de un código el cual indica que todos los demás números son ceros.

La exploración es una técnica que aumenta la probabilidad de alcanzar este resultado, porque ella envía los coeficientes en orden descendente según su probabilidad. La figura muestra que en un sistema no-entrelazado, la probabilidad de hallar coeficientes de mayor peso es más alta en la parte superior izquierda que en la parte inferior derecha. Aquí una exploración en forma diagonal a 45° es

la que se denomina una exploración en zig-zag, la cual es la mejor secuencia para emplear en este caso.

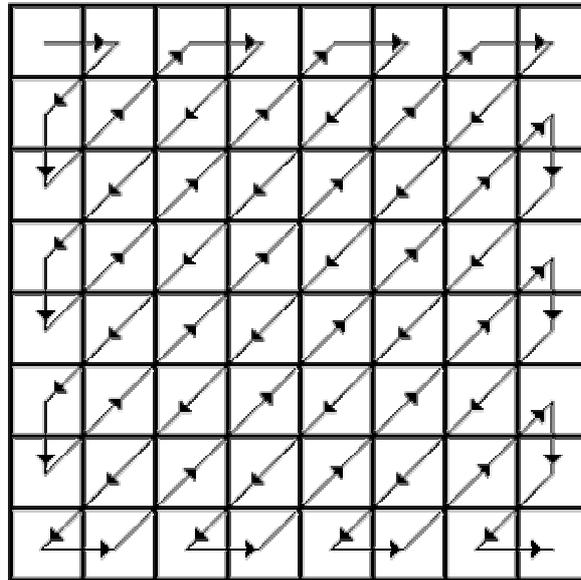


Ilustración [53] Exploración de un bloque 8x8

En la figura siguiente muestra una exploración para una fuente entrelazada, se observa que la exploración se extiende dos veces más por encima del área vertical, de este modo se pueden conseguir más detalles acerca de la imagen. Las frecuencias verticales aparecen dos veces más que las frecuencias horizontales. Por tanto, la exploración ideal para una imagen entrelazada será sobre una diagonal de 67.5° . La figura muestra que esta forma de exploración, entrega primero las frecuencias espaciales verticales y luego las frecuencias espaciales horizontales.

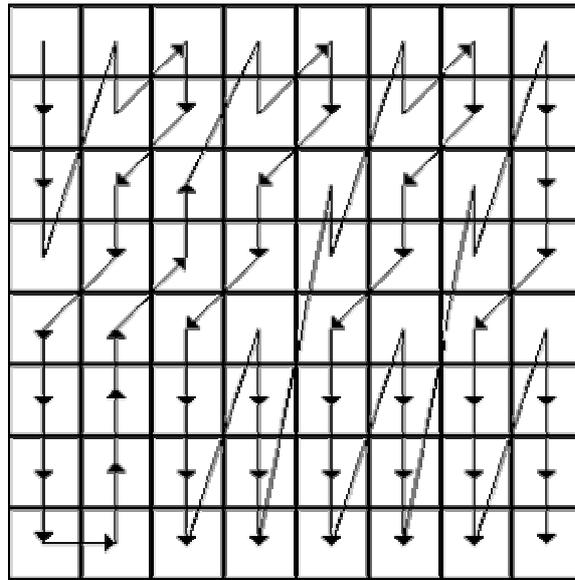


Ilustración [54] Exploración de un bloque 8x8 entrelazado.

15.1.7. 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.

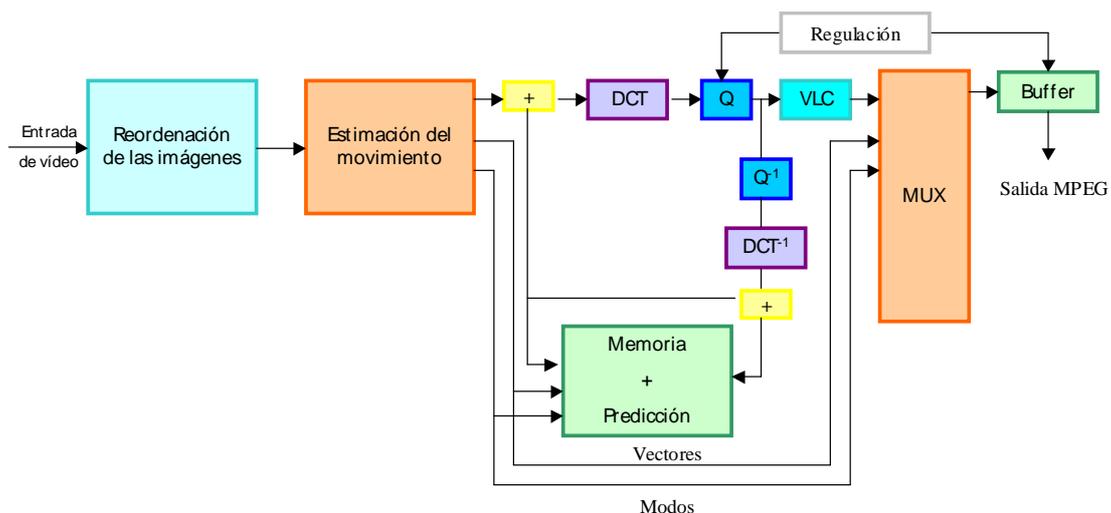


Ilustración [55] Esquema simplificado codificador MPEG-2

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.

15.1.8. 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.

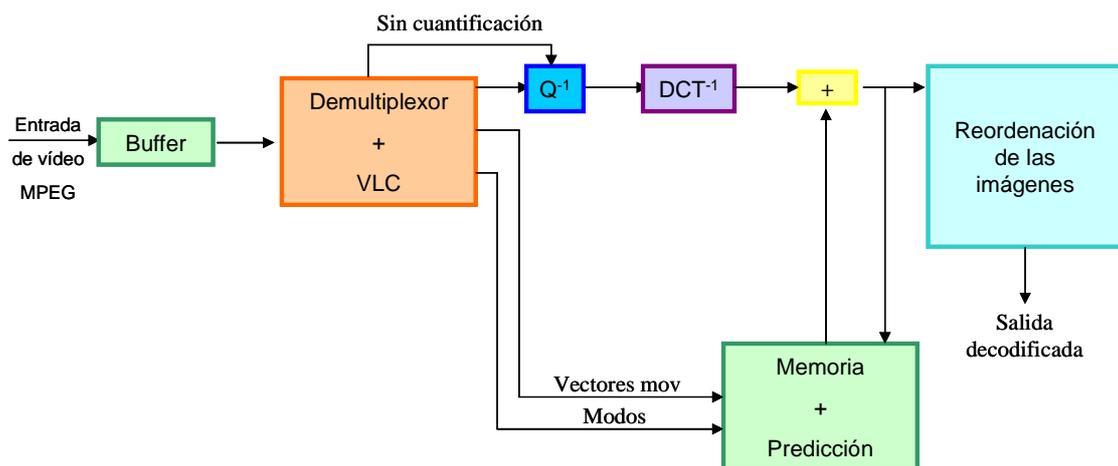


Ilustración [56] Esquema simplificado decodificador MPEG-2

El esquema de bloques del decodificador de la figura anteriores el que se va a analizar para 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 8x8 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.

15.2. La sección de Audio MPEG-2

El sistema de multiplexión MPEG-2, soporta cualquier número de canales de entrada de audio tan largos que la velocidad de transporte seleccionada pueda soportar la suma de datos. Los usuarios tienen la flexibilidad para seleccionar su propio algoritmo de compresión de audio, tales como: Audio MPEG 2, MUSICAM, DOLBY AC-2 o AC-3. Los canales pueden ser configurados independientemente o en pares estéreo. Diferentes velocidades de audio es otra de las características del sistema. Una vez más, la velocidad también será asociada con la calidad.

La compresión de audio MPEG-2 es un algoritmo que, como el vídeo MPEG-2, explota las limitaciones del sistema humano, en este caso el oído. Como en la compresión de vídeo, el algoritmo de compresión de audio también elimina la información irrelevante dentro de la señal de audio. La información irrelevante es cualquier señal imperceptible. Por ejemplo, en presencia de una señal fuerte, todas las señales vecinas flojas son enmascaradas y aunque son parte del

espectro, no son percibidas por el oído. El algoritmo MPEG 2 es del tipo "lossy" o con pérdidas pero la distorsión insertada por la señal será inaudible.

La configuración básica del audio MPEG 2 ofrece seis canales de audio. Esta característica debe ser usada para distribuir tres pares de estéreos (o seis canales mono) para aplicaciones multilinguajes o para crear un sistema estereofónico multicanal. Lo anterior crea una realidad como a la de un campo de audio. La recomendación de cornetas configuradas para sistemas estereofónicos multicanales es conocido como estéreo -p/q , donde p es el número de cornetas en el frente y q es el número de cornetas en el fondo. Por ejemplo un estéreo-3/2 proveerá un sistema con canales al frente en la derecha, centro y a la izquierda más canales posteriores que rodean el área y ofrecen un mejor e impresionante realismo a la audiencia.

Las configuraciones típicas para sistemas estereofónicos multicanales son:

- 1 Canal modo 1/0: Mono
- 2 Canales estéreo 2/0: izquierda y derecha
- 3 Canales estéreo 3/0: izquierda, derecha y centro
- 4 Canales estéreo 3/1: izquierda, derecha, centro y posterior
- 5 Canales estéreo 3/2: izquierda, derecha, centro, posterior izquierda y derecha
- 5.1 Canales estéreo 3/2: izquierda, derecha, centro, posterior izquierda y derecha y un canal de efectos especiales de 100 Hz LFE (Low Frequency Enhancement).

Audio MPEG 2 ofrece tres diferentes capas de compresión (capa I, II, y III). Cada capa usa un esquema de reducción incremental de la velocidad binaria, con la ayuda del incremento de la velocidad de compresión mientras se mantiene la calidad.

Para la capa II, la técnica de reducción de bit corresponde para el algoritmo MUSICAM, el permite varias combinaciones de velocidades de bit (32 a 224 Kbit/s por canal), y calidad de audio sin comprometer la complejidad del hardware. Por ejemplo una velocidad de muestreo de 56 Kbit/s y 16 Khz. en la capa II ofrece mejor calidad que la definida en el registro ITU-G.722.

15.3. Parte de Sistema MPEG 2 para Multiplexión y Transporte

La multiplexión y transporte definidos bajo MPEG 2 especifica el formato de codificación para multiplexar audio, vídeo y datos dentro de una forma manejable para almacenar o transmitir. Hay dos formatos de cadenas de datos definidos:

1. Cadena de transporte (TS), la cual porta o lleva uno o más programas simultáneamente, es optimo para aplicaciones donde la pérdida de datos puede ser requerida (como los enlaces satelitales). Tales errores pueden ser manifestados como errores de bit o pérdida de paquetes. Una TS es una sucesión de paquetes de 188 bytes de longitud llamados "paquetes de transporte". Es posible construir una TS a partir de:
 - o Cadenas elementalmente paquetizadas (PES), vienen de múltiples programas. Cada programa debe tener velocidad variable, sin embargo, la TS será de velocidad fija. Cada programa tiene un Reloj Referencial Primario (PCR) asociado con el para indicar la actual velocidad del programa.
 - o Cadenas de Programa u otra TS las cuales puedan contener uno o más programas.

La cadena de transporte es óptima para transmisiones de satélites, cable, ISDN, redes ATM y videotelefonía.

2. Cadena de programa. Este tipo de tren esta destinado a aplicaciones donde el dispositivo de almacenamiento o de transmisión sea susceptible de introducir muy pocos errores (error free medium), como es el caso, por ejemplo, de las aplicaciones multimedia en CD-ROM o disco duro. Aquí, estos paquetes pueden ser relativamente largos (por ejemplo, 2.048 bytes) y dado que está organizada de manera similar a un tren MPEG-1, no se entrará en detalles.

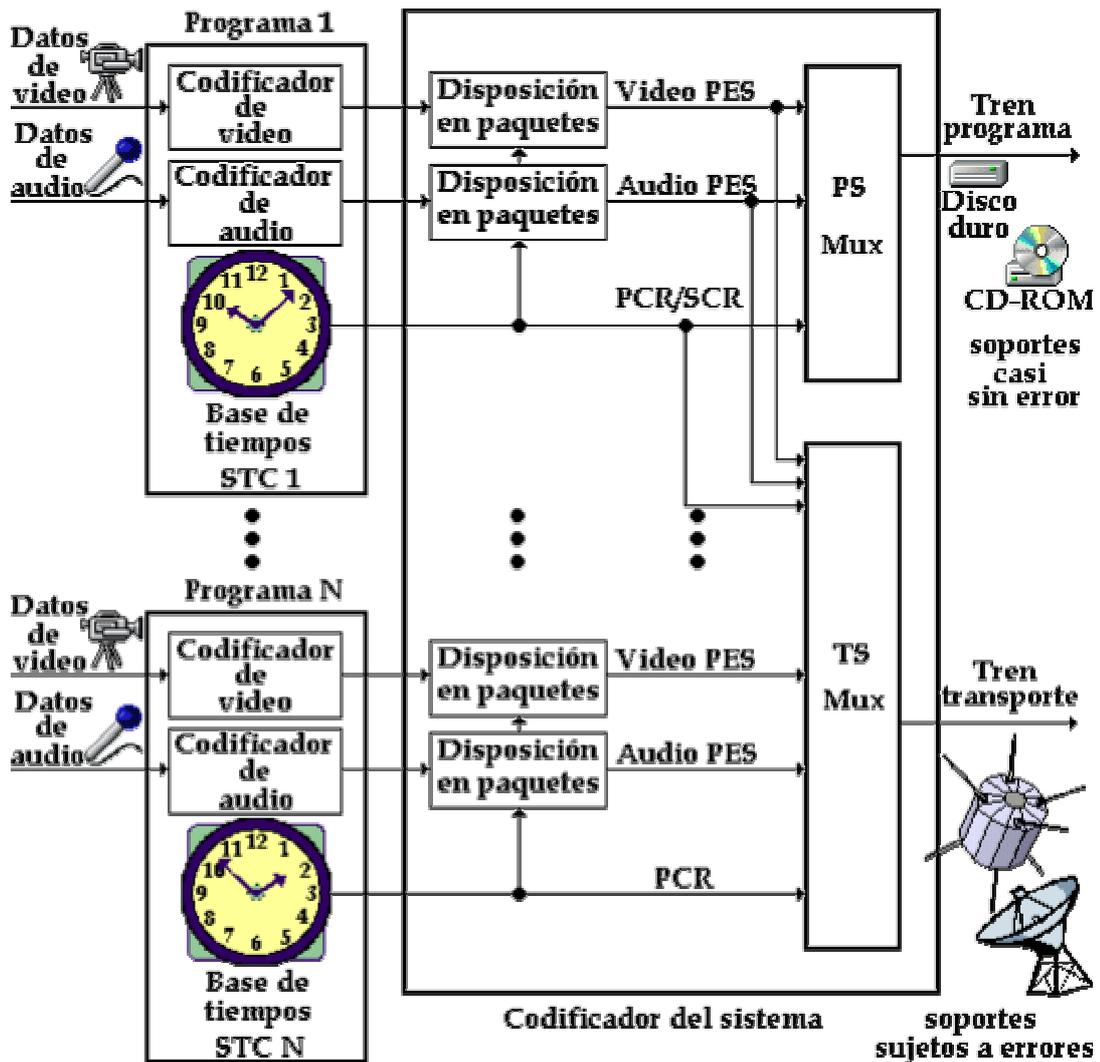


Ilustración [57] Trenes de programa y transporte

El tren de transporte de MPEG-2 está principalmente destinado al transporte de programas de televisión a larga distancia sobre soportes o en medios susceptibles de introducir un índice de errores bastante elevado (error prone medium); la longitud de los paquetes debe ser relativamente corta para permitir la introducción de los dispositivos de corrección de errores eficaces.

La longitud de los paquetes transporte de MPEG-2 ha sido fijada, por tanto, en 188 bytes, valor reservado especialmente para las emisiones vía satélite, cable o terrestres de la norma europea DVB.

Este tipo de tren está destinado a combinar varios programas que no compartan forzosamente el mismo reloj de sistema (STC) en el interior de un mismo multiplexor.

Los diferentes PES (video, audio, etc.) que forman un programa dado, deben sin embargo compartir el mismo reloj con el fin de poder ser sincronizados por el decodificador.

La figura siguiente ilustra el proceso de creación de un tren de transporte MPEG-2, del que se va a detallar ahora su constitución.

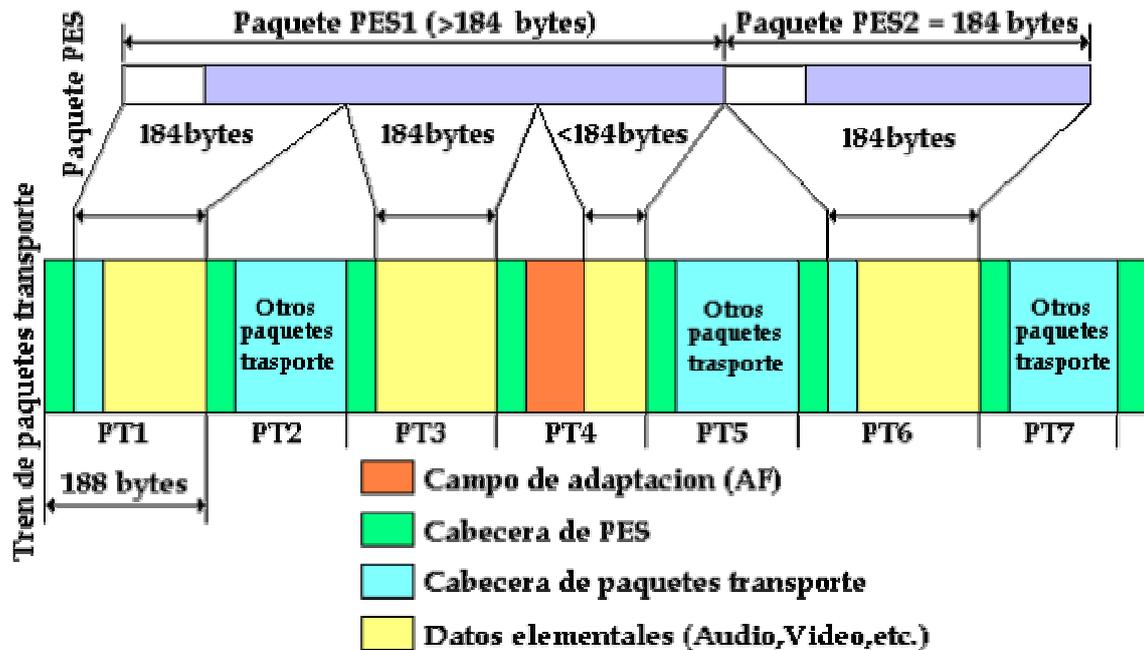


Ilustración [58] Tren de transporte MPEG-2

La multiplexión basada en paquete es muy flexible. Paquetes pertenecientes a diferentes programas pueden ser distribuidos de varias maneras. Un buffer es usado en el decodificador para asegurarse que todos los datos son descodificados y presentados a tiempo. Ya que la naturaleza inherente de la PES es variable en velocidad, el multiplexaje basado en paquetes es de gran aprovechamiento porque puede distribuir PES de diferentes fuentes con diferentes velocidades instantáneas dentro de una salida común con velocidad fija. Explotando la no coincidencia de picos de velocidad, se puede transportar altas velocidades instantáneamente en una salida de baja velocidad. Este tipo de multiplexor es algunas veces referido como multiplexor estadístico.

Las capas del sistema ofrecen diferentes modos de operación:

1. Multiplexaje estadístico, descrito anteriormente.

2. Multiplexaje estadístico limitado en el cual la velocidad de cualquier cadena de vídeo no podrá caer bajo un umbral programable.
3. Multiplexaje con velocidad fija, las velocidades de vídeo son fijas.
4. Modo mixto. Todos los modos previos pueden ser mezclados dentro de un simple TS.

Después del multiplexaje, la señal sufrirá el proceso final, cada paquete de transporte puede ser encriptado bajo control de acceso condicional. La salida de un multiplexor portando múltiples canales siempre será a velocidad fija.

16. Estándar MPEG4

MPEG-4, introducido a finales de 1998, es el nombre de un grupo de estándares de codificación de audio y video así como su tecnología relacionada normalizada por el grupo MPEG (Moving Picture Experts Group) de ISO/IEC. Los usos principales del estándar MPEG-4 son los flujos de medios audiovisuales, la distribución en CD, la transmisión bidireccional en videoconferencia emisión de televisión.

MPEG-4 toma muchas de las características de MPEG-1 y MPEG-2 así como de otros estándares relacionados, tales como soporte de VRML (Virtual Reality Modeling Language) extendido para Visualización 3D, archivos compuestos en orientación a objetos (incluyendo objetos audio, video y VRML), soporte para la gestión de Derechos Digitales externos y variados tipos de interactividad.

La mayoría de las características que conforman el estándar MPEG-4 no tienen que estar disponibles en todas las implementaciones, al punto que es posible que no existan implementaciones completas del estándar MPEG-4. Para manejar esta variedad, el estándar incluye el concepto de perfil (profile) y nivel, lo que permite definir conjuntos específicos de capacidades que pueden ser implementados para cumplir con objetivos particulares

16.1. Partes de MPEG-4

MPEG-4 está formado por varios estándares, llamados "partes", que incluyen:

1. Parte 1 (ISO/IEC 14496-1): Sistemas: Describe la sincronización y la transmisión simultánea de audio y video.
2. Parte 2 (ISO/IEC 14496-2): Visual: Un codec de compresión para elementos visuales (video, texturas, imágenes sintéticas, etc.). Uno de los muchos perfiles definidos en la Parte 2 es el Advanced Simple Profile (ASP).
3. Parte 3 (ISO/IEC 14496-3): Audio: Un conjunto de codecs de compresión para la codificación de flujos de audio; incluyen variantes de Advanced Audio Coding (AAC) así como herramientas de codificación de audio y habla.
4. Parte 4 (ISO/IEC 14496-4): Conformidad: Describe procedimientos para verificar la conformidad de otras partes del estándar.

5. Parte 5 (ISO/IEC 14496-5): Software de referencia: Formado por elementos de software que demuestran y clarifican las otras partes del estándar.
6. Parte 6 (ISO/IEC 14496-6): Delivery Multimedia Integration Framework (DMIF).
7. Parte 7 (ISO/IEC 14496-7): Software optimizado de referencia: Contiene ejemplos sobre como realizar implementaciones optimizadas (por ejemplo, en relación con la Parte 5).
8. Parte 8 (ISO/IEC 14496-8): Transporte sobre redes IP: Especifica un método para transportar contenido MPEG-4 sobre redes IP.
9. Parte 9 (ISO/IEC 14496-9): Hardware de referencia: Provee diseños de hardware que demuestran implementaciones de otras partes del estándar.
10. Parte 10 (ISO/IEC 14496-10): Advanced Video Coding (AVC): Un codec de señales de video técnicamente idéntico al estándar ITU-T H.264.
11. Parte 12 (ISO/IEC 14496-12): Formato para medios audiovisuales basado en ISO: Un formato de archivos para almacenar contenido multimedia.
12. Parte 13 (ISO/IEC 14496-13): Extensiones para el manejo y protección de Propiedad Intelectual (IPMP).
13. Parte 14 (ISO/IEC 14496-14): Formato de archivo MPEG-4: El formato de archivo de contenedor designado para contenidos MPEG-4; basado en la Parte 12.
14. Parte 15 (ISO/IEC 14496-15): Formato de archivo AVC: Para el almacenamiento de video Parte 10, basado en la Parte 12.
15. Parte 16 (ISO/IEC 14496-16): Animation Framework eXtension (AFX).
16. Parte 17 (ISO/IEC 14496-17): Formato de subtítulos (en elaboración - el último avance en su revisión data de enero de 2005).
17. Parte 18 (ISO/IEC 14496-18): Compresión y transmisión como flujo de fuentes tipográficas (para fuentes OpenType).
18. Parte 19 (ISO/IEC 14496-19): Flujos de texturas sintetizadas.
19. Parte 20 (ISO/IEC 14496-20): Representación liviana de escenas (LAsER).
20. Parte 21 (ISO/IEC 14496-21): Extensión de MPEG-J para rendering (en elaboración - el último avance en su revisión data de enero de 2005).

También es posible definir perfiles a nivel de las partes, dado que una implementación de una parte no necesariamente contiene toda esa parte.

La concepción MPEG-4 comenzó a definirse formalmente en la reunión de MPEG que se celebró en Bruselas en Septiembre de 1993. Se trataba, en un principio, de elaborar una iniciativa orientada al desarrollo de procedimientos de codificación para sistemas con una velocidad de transmisión muy baja. Sin embargo, este objetivo con el tiempo se ha visto ampliado, de forma que, en la actualidad, el proyecto MPEG-4 está orientado a establecer mecanismos de codificación, eficientes y universales, para los distintos tipos de información audiovisual, denominados objetos audiovisuales. Dichos objetos pueden ser de origen natural y/o sintético. Para alcanzar este objetivo han de definirse dos elementos básicos:

- Un conjunto de herramientas de codificación para objetos audiovisuales, capaz de dar soporte a diferentes funcionalidades, tales como la interactividad y la escalabilidad orientadas a objetos, la robustez frente a errores y la eficiencia de compresión.
- Una descripción sintáctica de los objetos audiovisuales codificados, que proporcione un método formal para describir la representación codificada de dichos objetos, así como de los procedimientos empleados para su codificación.

Se pone así de manifiesto, desde el primer momento, la importancia que en MPEG-4 van a tener los conceptos de objeto audiovisual, funcionalidad y lenguaje de descripción sintáctica (MSDL).

16.2. Nuevas funcionalidades

Existen ocho funcionalidades básicas relacionadas con la interactividad basada en el contenido, la compresión y el acceso universal y, por tanto, dentro del ámbito de interés de MPEG-4, cuyo soporte por otros estándares ya existentes o emergentes no es totalmente satisfactorio. El conjunto de herramientas de codificación y el MSDL son precisamente los elementos que soportan dichas funcionalidades y que a se describen continuación:

16.2.1. Acceso a datos multimedia basado en contenido

Se desea habilitar el acceso a datos basado en contenido, mediante diversos procedimientos tales como indexado e hiperenlaces.

16.2.2. Edición y manipulación del flujo de datos codificados basados en contenido

Se va a proporcionar un MSDL y unos esquemas de codificación, a fin de soportar la manipulación y edición de flujos de datos codificados sin precisar. El MSDL debe ser suficientemente flexible para admitir extensiones en usos futuros.

16.2.3. Codificación de datos híbridos naturales y sintéticos

Se pretende incluir métodos eficientes para combinar escenas u objetos sintéticos con escenas u objetos naturales, posibilitar la codificación y manipulación de datos de audio y vídeo sintéticos, y dotar al descodificador de procedimientos controlables para componer datos sintéticos con otros ordinarios, de forma que se admita la interactividad.

16.2.4. Acceso temporal aleatorio mejorado

Se va a proveer a MPEG-4 de procedimientos que permitan acceder de forma aleatoria a partes de una secuencia audiovisual, con la limitación de tiempo correspondiente y con buena resolución temporal.

16.2.5. Eficiencia de compresión mejorada

MPEG-4 debe proporcionar, a las aplicaciones que lo precisen, la mejor calidad audiovisual posible, a una velocidad de transmisión determinada.

16.2.6. Codificación de múltiples flujos de datos concurrentes

Se pretende facilitar la codificación eficiente de vistas o bandas sonoras múltiples y proporcionar mecanismos de sincronización para los datos codificados resultantes. En particular, se van a incorporar procedimientos para aplicaciones estereoscópicas, compatibles con audio y vídeo normales, que permitan explotar la redundancia existente entre diversos puntos de vista o audición, así como otros en los que no sea necesaria dicha restricción.

16.2.7. Robustez en entornos con propensión a error

Se va a dotar al estándar de una robustez frente a errores (que le permita proporcionar acceso a aplicaciones que utilicen redes de acceso, tanto móviles como cableadas) y de medios de almacenamiento de datos. En concreto,

suministrará una suficiente protección contra errores, en aplicaciones de baja velocidad con condiciones de error importantes (ráfagas largas de errores).

16.2.8. Escalabilidad basada en contenido

Va a proporcionarse mecanismos para conseguir codificaciones escalables con granularidad en contenido, calidad y complejidad. Todo ello debe estar orientado a proporcionar escalabilidad basada en contenido a la información audiovisual.

Existen, además, otras funcionalidades nuevas o mejoradas relacionadas con la sincronización, el multiplexado de los datos, la seguridad, el formato y la calidad.

16.3. Codificación orientada a objetos

Una diferencia fundamental de MPEG-4, con respecto a otros sistemas de codificación, es el empleo de objetos audiovisuales. Un objeto audiovisual (AV) es la representación de un objeto real o virtual que se puede manifestar de forma visual o auditiva. Los objetos AV tienen frecuentemente carácter jerárquico, en tanto que pueden ser definidos como composiciones de otros objetos AV que se denominan, entonces, subobjetos. Los objetos AV que constan de diversos subobjetos son llamados objetos compuestos, se reserva la denominación de objetos AV primitivos para los que no cumplen dicha condición.

Dentro del ámbito de la codificación de vídeo, en MPEG-4 se usan Planos de Objetos de Vídeo (VOPs), entendiéndose por tales las instancias de un objeto de vídeo en un instante determinado de tiempo. De este procedimiento, debe destacarse la flexibilidad que supone el hecho de admitir procedimientos de codificación diversos para los distintos VOPs que componen la escena. En la práctica, un VOP consta de la información que se incluye mediante un plano alfa (análogo a los que se emplean para soportar composiciones con transparencias parciales en infografía) y de la información codificada de los píxeles que componen el VOP. La información sobre la forma de los VOPs se obtiene en el proceso de edición, en el caso de imágenes sintéticas, y mediante conmutación de croma (Chroma Key) u otros procesos de segmentación, en el caso de imágenes naturales.

16.4. El modelo de verificación

El grupo MPEG ha definido un modelo de codificador, conocido con el nombre de modelo de verificación (VM), con objeto de facilitar los trabajos para el desarrollo del estándar. El VM consta, en la actualidad, de descripción del codificador y el descodificador, y una especificación de la sintaxis del flujo de datos codificados.

El codificador está basado en esquemas convencionales adaptados al caso particular de VOPs. Para ello incluye elementos que le permiten codificar la información sobre la forma del VOP contenida en el correspondiente plano alfa. Está prevista la posibilidad de disminuir la redundancia temporal, mediante un proceso de compensación de movimiento. Seguidamente, se efectúa una codificación de la textura del VOP que en la actualidad hace uso de la DCT-2D y admite el empleo de cuantificadores tipo H.263 y MPEG. Por último, se multiplexan todos los elementos recabados en el análisis, empleando procedimientos de representación diferencial y códigos de longitud variable para reducir al mínimo la redundancia en el flujo de datos codificados. Es importante destacar que en el codificador se ha previsto el empleo de procedimientos mejorados para la representación y compensación de movimiento, similares a los que se emplean en el estándar H.263. Asimismo, se pueden codificar VOPs empleando predicción bidireccional según los modelos H.263 y MPEG. Está prevista también la posibilidad de producir flujos de datos codificados con escalabilidad espacial o temporal empleando, para ello, dos codificadores básicos MPEG-4 y un postprocesador específico. La idea básica consiste siempre en que el primer codificador MPEG-4 actúe sobre versiones de baja resolución (espacial o temporal) de la secuencia de entrada. Su salida se transforma, en caso de ser necesario, en el postprocesador y se emplea como información adicional por el segundo codificador MPEG-4 que debe dar lugar a un refinamiento de la secuencia codificada. La mezcla de ambos flujos de datos codificados se efectúa mediante un multiplexor MSDL.

La sintaxis del flujo de datos codificados es similar a la de los estándares anteriores y tiene una representación jerárquica de cuatro capas.

El descodificador propuesto para MPEG-4 incorpora elementos de recuperación de forma de los VOPs, así como elementos para combinar los distintos VOPs en

una imagen de salida. Por lo demás, es análogo a los descodificadores correspondientes a los estándares anteriores.

16.5. El lenguaje de descripción de algoritmos

En MPEG-4 se ha otorgado siempre una importancia crucial a la flexibilidad del estándar. De hecho, se acepta que existirán tres niveles de capacidad de programación del descodificador. En el nivel más bajo o nivel 0, el descodificador no es programable y es capaz de proporcionar un algoritmo de descodificación seleccionable perteneciente a un abanico de posibilidades preestablecido. En el nivel 1, denominado flexible, el descodificador ofrece también una colección de herramientas estandarizadas que pueden configurarse remotamente de forma flexible en un algoritmo de descodificación. En el nivel superior, o nivel 2, denominado extensible, el descodificador suministra adicionalmente mecanismos para permitir la definición remota de nuevas herramientas y algoritmos de descodificación. En este contexto, se concibe inicialmente el MSDL (MPEG-4 Syntax Definition Language) como un instrumento para la descripción de algoritmos y de su sintaxis. Sin embargo, con el transcurso del tiempo MSDL ha pasado a desempeñar el papel de trama de sistema de MPEG-4. Define, para ello, todos los elementos necesarios para soportar la flexibilidad, la extensibilidad y las diversas funcionalidades MPEG-4 anteriormente revisadas. Con esta finalidad, en MSDL se otorga una importancia primordial tanto a los objetos audiovisuales representables, como a sus procedimientos de representación.

En MSDL se distinguen distintas facetas de interés:

- Una descripción global de la arquitectura del sistema MPEG-4. Abarca el papel que tal sistema puede desempeñar en una aplicación audiovisual y los objetos conceptuales que es posible intercambiar entre el codificador y el descodificador. Esta faceta se denomina MSDL-A (MSDL-Architecture).
- Una especificación de las clases de objetos que van a ser útiles para aplicaciones audiovisuales específicas, por lo que deben aparecer en el estándar. Esta área de trabajo recibe el nombre de MSDL-O (MSDL-Objects).
- Una descripción de un formato legible, apto para la transmisión de directrices de descodificación, que se denomina MSDL-R (MSDL-Readable).

- Una especificación de un formato ejecutable binario, para descripciones o directrices de decodificación, que constituirá el lenguaje binario ejecutable por el decodificador y se denominará MSDL-B (MSDL-Binary).
- Una especificación del lenguaje de descripción sintáctico que se va a emplear como un primer paso para describir la especificación de la sintaxis del flujo de datos codificado. Este ámbito de definición se llama MSDL-S (MSDL-Syntax).
- Una descripción del procedimiento de multiplexado para la información codificada conocida como MSDL-M (MSDL-Multiplex).

16.6. Codificación de audio.

Existe una gran variedad de técnicas de codificación de audio, muchas de ellas no normalizadas, que se emplean en muy diversos ámbitos. Aquí sólo se revisan brevemente algunas de las técnicas empleadas con más frecuencia, junto con las técnicas de transmisión de imágenes, que forman parte de estándares internacionales.

Las técnicas de codificación de audio pueden dividirse en dos grandes grupos, según estén orientadas a sistemas de telefonía (donde generalmente se transmiten señales vocales) o a la reproducción o transmisión de señales más genéricas que abarquen un espectro más amplio. En el primer caso, se acepta cierta limitación en las prestaciones del codificador para obtener una calidad que permita oír con comodidad la voz humana. En el segundo, se trata de optimizar los codificadores, de manera que cualquier distorsión o ruido añadido quede por debajo del umbral de percepción del oído humano.

Al igual que en las técnicas de codificación de vídeo, una de las funcionalidades que históricamente se han buscado con más interés ha sido la compresión de las señales, con objeto de reducir la cantidad de información que es necesario almacenar o enviar. Sin embargo, los factores de compresión que se consiguen son más limitados, debido posiblemente a que, en este caso, no existe tanta redundancia en la información.

Los sistemas de codificación de audio con calidad musical parten generalmente de señales digitales obtenidas mediante el muestreo a 44,1 Khz. , con resolución de 16 bits por muestra de las señales originales, con lo que se consigue un

margen de frecuencias de 20Hz a 20 Khz. y un margen dinámico de 90 dB, que se adapta bien a la capacidad de percepción del oído humano. Este formato, sin otra codificación adicional, es el empleado en los CD's de audio y requiere un flujo de datos de 705 Kbps por canal. Es posible reducir este flujo disminuyendo sólo la resolución de las muestras, con lo que se aumenta el ruido, o la frecuencia de muestreo, con lo que se reduce el margen de frecuencias.

Los sistemas de codificación de audio utilizan técnicas para reducir la redundancia de la señal y emplean técnicas psicoacústicas para eliminar los sonidos que el oído humano no puede percibir. En el estado del arte actual se consigue un buen nivel de calidad con alrededor de 1 bit por muestra para señales vocales y 2 bits por muestra para señales musicales. Probablemente estos valores se reduzcan a la mitad en un futuro no muy lejano.

16.6.1. Codificación orientada a la telefonía

La recomendación más antigua es la G.711 que especifica cuáles son las claves de 8 bits que se obtienen de la codificación logarítmica (siguiendo la ley A en Europa y ley μ en EEUU) de canales telefónicos digitalizados, con una frecuencia de muestreo de 8 Khz. y una resolución de 12 bits por muestra. Da lugar a una tasa binaria de datos de 64 kbit/s.

Más adelante aparecieron normas de codificación basadas en Adaptive Differential PCM (ADPCM) en las que se codifica la diferencia entre el valor de cada muestra y el de un predictor que se calcula a partir de las muestras anteriores. La más empleada en aplicaciones audiovisuales es la G.722 que hace posible codificar señales con ancho de banda de hasta 7 Khz. a velocidades de 48 a 64 kbit/s.

Por último, cabe citar la familia de codificadores de tipo Code Excited Linear Prediction (CELP). En este tipo de codificador, se modela la señal mediante la respuesta de un filtro lineal a una excitación que se escoge convenientemente de entre una familia contenida en un libro de códigos. Tanto el índice del código como los coeficientes del filtro se obtienen reduciendo al mínimo la diferencia ponderada entre la señal de entrada y la señal sintetizada, procedimiento éste denominado análisis mediante síntesis. Existen tres recomendaciones del ITU en vigor que especifican codificadores basados en estos principios para señales con anchura de banda telefónica:

16.6.1.1. G.728

Esta recomendación contiene las especificaciones para un codificador de bajo retardo que da lugar a una tasa binaria de datos de 16 kbit/s.

16.6.1.2. G.729

Especifica un codificador con una estructura bastante sofisticada que produce una tasa binaria de 8 kbit/s.

16.6.1.3. G.723.1

Siguiendo esta recomendación se precisa tan sólo una capacidad de transmisión de 5,3 ó 6,4 Kbps. Como se puede apreciar, estos últimos codificadores tienen ya prestaciones muy cercanas al valor de 1/2 bit por muestra anteriormente citado.

Tradicionalmente se han usado los codificadores de vídeo basados en H.261 empleando esquemas de codificación de audio del tipo G.711, que requieren al menos 56 Kbps. En 1996 se revisó el estándar H.320, con objeto de incluir todos los modos de codificación mencionados antes.

Los codificadores de vídeo del tipo H.263 emplearán normalmente el codificador G.723.1, lo que supone una mejora, de un orden de magnitud, en la eficiencia de compresión con respecto al G.711 empleado ahora por la mayor parte de los videoteléfonos.

16.6.2. Sistemas de codificación de audio orientados a contenidos de tipo genérico

Los principales sistemas de codificación de audio están orientados a obtener calidad musical y concebidos para integrarse con los sistemas de codificación de vídeo que han sido desarrollados por el grupo MPEG de la ISO. En este grupo, se han desarrollado codificadores que se engloban en los estándares MPEG-1 y MPEG-2, y en la actualidad continúan los trabajos sobre las nuevas normas que se incluirán en MPEG-4.

Como parte de MPEG-1, se ha desarrollado un sistema de codificación de audio que permite codificar señales de tipo genérico con calidad equivalente a la

obtenida en un disco compacto. Utilizan técnicas de codificación en 32 subbandas y técnicas psicoacústicas para comprimir las señales. Presenta tres modos de codificación (layer 1 a 3) que se diferencian, principalmente, en la complejidad del codificador. Con el modo 3 es posible codificar un canal a velocidades de hasta 56 Kbps. En este caso, al igual que en los codificadores de vídeo, el estándar deja abierto cómo realizar el codificador, por lo que puede haber variaciones de calidad en distintas implementaciones. Existe, además, un modo específico para codificar señales estéreo, el cual tiene en cuenta la correlación existente normalmente entre los dos canales para aumentar la compresión.

En el estándar MPEG-2 de audio, terminado a finales de 1994, se han incluido nuevas opciones de codificación que se pueden resumir en lo siguiente :

Aumento, hasta cinco, del número de canales para codificar e inclusión de un canal para sonido de baja frecuencia (sonido surround).

- Inclusión de hasta 7 canales para comentarios de locutores.
- Nuevos modos de compresión a baja velocidad hasta 8 kbit/s.
- Codificación de audio en MPEG-4

El estándar MPEG-4 pretende ser extraordinariamente flexible en lo concerniente a la codificación de audio. En este momento, se está trabajando sobre un modelo de codificador orientado a señales de entrada, que abarca todo el rango de calidades comprendido entre el estándar telefónico y el de alta fidelidad. Se están considerando las tasas binarias que se encuentran entre 2 y 64 Kbps por canal, por lo que van a ser aptas para múltiples canales de comunicación de interés práctico.

Al igual que en el grupo de MPEG-4 de vídeo, se pretende incluir en este estándar de codificación numerosas funcionalidades nuevas, adicionales a la tradicional de compresión. Las principales son:

- Mayor flexibilidad. Se hará uso del lenguaje MSDL, ya mencionado, en la parte de vídeo para dotar al sistema de gran flexibilidad. Está previsto, al igual que en vídeo, poder enviar a los descodificadores algoritmos de descodificación.
- Adaptación a diversos tipos de redes incluyendo las móviles

- Escalabilidad.
- Se podrá descodificar el contenido con diversos grados de calidad, dependiendo del medio de transmisión utilizado.
- Tratamiento específico de contenidos sintéticos

Incluirá mecanismos para el manejo de música sintética, síntesis de voz, modelado de entornos acústicos, etc.

16.7. La protección de los derechos de autor y el control de acceso a la información.

La digitalización de la información audiovisual y su codificación con las técnicas descritas en este artículo, hace que sea posible su transmisión y duplicación tantas veces como se quiera, sin que por ello se produzca una merma de la calidad. Esto facilita, en consecuencia, la reproducción ilegal de contenidos y dificulta la protección de los derechos de autor.

En la actualidad, se están investigando técnicas que proporcionan diversos grados de protección. Una de ellas consiste en la inclusión de etiquetas en la trama digital que contiene la información relacionada con los derechos de autor. Por ejemplo, la trama de sistema MPEG-2 contiene un campo en el que se almacena este tipo de información. Los equipos digitales para la reproducción y grabación de la información deben leer esta etiqueta y atender las limitaciones que en ella se muestren (se pueden indicar el número de máximo copias que es posible realizar, los periodos de tiempo en que se puede reproducir los contenidos, etc.). La objeción a esta técnica es que puede ser fácilmente manipulable, sin más que utilizar un equipo programado para no atender las indicaciones de la etiqueta que contiene la información de los derechos de autor.

Otro procedimiento, que actualmente se está considerando, consiste en el marcado de la información audiovisual mediante etiquetas que se introducen en el sonido o las imágenes. La particularidad de estas etiquetas es que es muy difícil averiguar su contenido o borrarlas sin conocer una clave usada en el proceso de marcado. Por otra parte, pueden resistir procesos de copiado analógico de los contenidos, que mantengan cierto nivel de calidad. Telefónica I+D está trabajando para el desarrollo de este tipo de técnicas.

16.8. EI MPEG-4 SP y ASP

La primera versión de MPEG-4 fue el MPEG-4SP (protocolo simple) que fue pensado para los usos de bajo ancho de banda tales como acceso a Internet por línea telefónica.

MPEG-4 SP utiliza tanto Intra-cuadro como Intra-campo para barrido progresivo o entrelazado y emplea compresión y DPCM (modulación de código de pulso diferencial) para asegurar la compresión óptima para cada tipo de vídeo.

MPEG-4 ASP (protocolo simple avanzado) incluye un número de características adicionales incluyendo:

- Compensación de movimiento con 1/4 de exactitud de pel.
- Predicción bidireccional.
- Compensación de movimiento en la totalidad del área del cuadro.
- Compensación de movimiento por múltiples imágenes de referencia, comparada con solamente un cuadro anterior para codecs previos.

16.9. H.264 o MPEG-4 AVC MPEG-4

La parte 10, también conocida como MPEG-4 AVC (codificación video avanzada) o H.264, ofrece una compresión perceptiblemente mayor que sus precursores. Puede proporcionar vídeo de calidad DVD con un 40% inferior de la tasa binaria de MPEG-2 y se considera prometedor para vídeo a imagen completa terrestre, satélite, y conexiones ADSL de Internet. Es también uno de los codecs de video que se han elegido provisionalmente para el láser azul HD DVD.

H.264 hace uso las redundancias espaciales, temporales, y psico-visuales para mejorar la eficiencia de la codificación de video.

16.10. Comparación de formatos

La siguiente tabla compara los formatos MPEG-4 H.264 con MPEG-4 ASP y MPEG-2.

	MPEG-2	MPEG-4 ASP	MPEG-4 H.264
Cuadros I, P, B	X	X	X
Codificación	Huffmann	Huffmann	Huffmann o aritmética
Tamaño bloque	16x16	16x16	Hasta 4x4
Tasa de bit	100%	60%	40%

DATA-SHEETS DE SET-TOP-BOXES