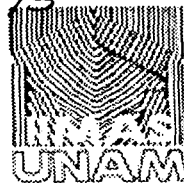




Universidad Nacional Autónoma de México

01130
15



Facultad de Ingeniería

TESIS
FALLA DE ORIGEN

DISEÑO DE UN DECODIFICADOR DIGITAL
DE ATSC A NTSC.

TESIS
QUE PARA OBTENER EL TÍTULO DE:
INGENIERO EN TELECOMUNICACIONES

PRESENTAN:

FLORES HUERTA HUGO
GASCA ACEVEDO ALBERTO
JAIME FRÍAS FRANCISCO J.

Director de Tesis:

M en I. Arturo González-Hermosillo M.

México, D.F.

2003



17



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

TESIS CON FALLA DE ORIGEN

A g r a d e c i m i e n t o s

A mi papá: Trinidad Flores L. y a mi mamá: Hilda Huerta E., por brindarme su apoyo y confianza, sin los cuales no habría logrado concluir mis estudios. A la Universidad Nacional Autónoma de México y Facultad de Ingeniería, por darme la oportunidad, por medio de las herramientas y conocimientos que imparte en sus aulas, laboratorios, bibliotecas y centros culturales, de superarme como persona y desempeñarme en una área laboral. Al M. en I. Arturo González-Hermosillo M., por invertir parte de su tiempo en la dirección de esta tesis. Y a mis compañeros de tesis Alberto y Javier.

Flores Huerta Hugo

AGRADECIMIENTOS

Este trabajo lo dedico a todos aquellos que contribuyeron para que realizara mi meta de ser Ingeniero en Telecomunicaciones.

Le doy las gracias a...

Jehová mi Dios, el Creador de todas las cosas, por haberme dado la vida así como la oportunidad de conocerlo. Sin Él no sería posible nada.

Mi papá Alberto Juan Gasca Gutiérrez, y mi mamá Carolina Acevedo Dionisio. Siempre han sido mi inspiración. Me han dado su amor, apoyo, confianza y cariño de forma incondicional. Gracias por los consejos, regaños y tiempo que me han dedicado. Gracias por animarme a seguir adelante a pesar de los problemas. Gracias por haber estado cerca de mi.

Mi hermano y mejor amigo, Arturo Gasca Acevedo, quien ha estado junto a mi en las buenas y en las malas.

Mis hermanos y hermanas espirituales, amigos y amigas, que siempre me han demostrado su amor fraternal.

Mis compañeros de tesis, *Hugo y Javier*, con los que compartí la ardua tarea de realizar este trabajo. A todos los compañeros de estudios que tuve, juntos tuvimos muchas experiencias durante la mejor etapa de la vida.

Nuestro director de tesis, *M. I. Arturo González-Hermosillo M.*

Todos mis profesores -en especial al *maestro Gualterio-*, quienes me enseñaron la importancia de trabajar para conseguir las metas que uno se traza.

La *Universidad Nacional Autónoma de México* por haberme permitido realizar estudios profesionales.

Alberto Gasca Acevedo.

AGRADECIMIENTOS

En primer lugar a mis padres, Naty y Javier, por brindarme esta oportunidad que no desperdiciare.

A mis hermanas, Griselda, Nelly y Carolina, debido a que son un gran soporte en mi vida.

A mis familias, por su apoyo incondicional.

A todos y cada uno de mis amigos(as), ya que me brindaron su amistad en todo momento, dedicándola especialmente y con gran cariño a Mónica, Manolo, Hugo y Andrés, por todo lo que me enseñaron a lo largo de estos años.

A mi niña, por fortalecerme cuando más lo necesitaba.

A la Facultad de Ingeniería, por brindarme las herramientas para alcanzar mis sueños.

A la Universidad Nacional Autónoma de México, por ser parte de mi historia, mi segundo hogar y uno de los lugares que más amo.

Fco. Javier Jaime Frias

ÍNDICE

Contenido	Página
Prólogo	IX
1. INTRODUCCIÓN A LA TELEVISIÓN DIGITAL.	1
1.1. ¿Qué es la TV digital?	3
1.2. Ventajas de la DTV sobre la TV analógica	5
1.3. ¿Que implica la transición de TV analógica a TV digital para el usuario?	6
1.3.1 Cambios que debe hacer la estación transmisora	7
1.4. Beneficios que genera el cambio de TV analógica a TV digital	7
1.5. Consecuencias socio-culturales de la DTV	10
1.6. Situación actual de la TV Digital	11
Referencias capítulo # 1	14
2. NORMAS INTERNACIONALES PARA TV DIGITAL	15
2.1. Estándares ATSC, DVB e ISDB	17
2.1.1 El estándar Americano ATSC	17
2.1.2 El estándar Europeo DVB	25
2.1.3 El estándar Japonés ISDB	28
2.2. Comparación entre estándares	32
2.2.1 Sistemas de modulación utilizados	32
2.2.2 Sistema de compresión utilizado	34
2.2.3 Comparación de parámetros de transmisión	37
2.3. ¿Qué estándar adoptara México?	39
2.3.1 ¿Por qué ATSC?	39
2.4 Descripción general de ATSC	40
Referencias capítulo # 2	42
3. COMPRESIÓN DE VÍDEO	43
3.1 Compresión de Datos	45
3.2 Introducción a la Compresión de Video	46
3.3 El sistema visual humano	47
3.4 Fundamentos del video digital	49
3.4.1 Video analógico	49
3.4.2 Video digital	51
3.5 Compresión de video bajo MPEG	56
3.6 Compresión de Video con MPEG-2	56
3.6.1 Compresión Intraframe	57

Contenido	Página
3.6.2 Compresión Interframe	64
3.6.3 Diagrama de Bloques de un Codificador de Video MPEG-2	73
3.6.4 Decodificador de video MPEG-2	76
3.7 Perfiles y niveles MPEG-2	78
Referencias capitulo # 3	80
4. COMPRESIÓN DE AUDIO	81
4.1 Introducción a la Compresión de Audio	83
4.2 El Sistema Auditivo Humano	84
4.3 Procesos de Compresión de Audio	89
4.3.1 Codificación de la sub-banda	89
4.3.2 Codificación por Transformación	91
4.4 Introducción a la técnica de compresión de audio AC-3	93
4.4.1 Desarrollo de AC-3	93
4.4.2 AC-3, sus 5.1 Canales de Audio y los Servicios de Audio	95
4.5 Compresión de Audio Digital mediante AC-3	98
4.6 Bloques principales del codificador AC-3	101
4.6.1 Señal digital de entrada	102
4.6.2 Buffer de entrada	102
4.6.3 Filtrado en la entrada	102
4.6.4 Detección de transitorios	102
4.6.5 Banco de filtros TDAC y Transformada directa	103
4.6.6 Conversión a punto flotante	105
4.6.7 Estrategia de Acoplamiento	105
4.6.8 Codificación exponencial y Asignación de bits	106
4.6.9 Cuantización	112
4.6.10 Empaquetado de datos	112
4.7 La trama de bits de AC-3	112
4.7.1 Características de la trama de bits de AC-3	112
4.7.2 Resumen de la trama de bits	114
4.8 El proceso de decodificación de AC-3	117
4.9 Bloques principales del decodificador AC-3	118
4.9.1 Buffer de entrada	119
4.9.2 Verificación de errores	119
4.9.3 Desempacado de los datos de formato fijo	119
4.9.4 Decodificación de la asignación de bits	119
4.9.5 Desempacado de datos de formato variable	119
4.9.6 Conversión a punto fijo	120
4.9.7 Reinserción de la portadora de alta frecuencia (Desacoplamiento)	120
4.9.8 Transformada Inversa	120
4.9.9 Buffer de salida PCM	120

Contenido	Página
4.9.10 Salida PCM	120
4.10 Algunas características adicionales de AC-3	120
Referencias capítulo # 4	122
5. TRANSPORTE Y MULTIPLEXADO.	125
5.1 Multiplexado de las señales en MPEG-1	127
5.2 Generación de los flujos de la señal MPEG-2.	127
5.3 Procesamiento de la señal	130
5.4 Codificación MPEG.	131
5.5 Empaquetado P.E.S	132
5.5.1. Cabecera del paquete PES	133
5.5.2 PES-Packet start code	134
5.5.3 Flags 1 y 2	134
5.5.4 Time Stamps	134
5.5.5 PES_header_length	135
5.6 Flujo de programa	136
5.7 Flujo de transporte (Transport stream)	136
5.7.1 Formación del "Flujo de transporte"	137
5.7.2 Paquetes de transporte	138
5.7.3 Cabecera del paquete de transporte	138
5.8 Información Especifica de los Programas (PSI)	142
5.8.1 Tabla de asociación de programa (PAT)	143
5.8.2 Tabla de acceso condicional (CAT)	143
5.8.3 Tabla del mapa del programa (PMT)	144
5.9 Inserción de las secciones en los paquetes de transporte	148
5.9.1 Composición de las secciones	149
5.10 Cómo se realiza el demultiplexado MPEG-2	150
5.11 Errores	150
5.11.1 Distancia Hamming y distancia de código	151
5.11.2 Códigos lineales y corrección de errores mediante el método del síndrome	151
5.11.3 Bloque FEC de ATSC	154
5.11.4 Interfase en banda base	155
5.11.5 Generador de datos aleatorios	155
5.11.6 Codificador Reed-Solomon	156
5.11.7 Reordenamiento de datos	165
5.11.8 Codificador Trellis	166
Referencias capítulo # 5	171

Contenido	Página
6. MODULACION 8-VSB.	173
6.1. ¿Qué es 8-VSB?	175
6.2. Descripción de un transmisor VSB.	175
6.3. Procesamiento de datos.	176
6.3.1. Aleatorizador de datos.	176
6.3.2. Codificador Reed-Solomon.	177
6.3.3. Entrelazado de datos.	177
6.3.4. Codificador Trellis.	178
6.4. Procesamiento de la señal.	178
6.4.1. Sincronía e inserción de la señal piloto.	178
6.4.2. Modulador VSB.	180
6.4.3. Conversor ascendente de RF y el resto de los procesos de 8-VSB.	185
6.5. Descripción de un receptor VSB	185
Referencias capítulo # 6	188
7. RECEPTOR DE ATSC.	189
7.1. Diagrama general de un receptor de TV digital.	191
7.2. Antena.	192
7.3. Sintonizador (tuner).	193
7.4. Demodulador de 8-VSB.	196
7.4.1. Filtro IF y detector síncrono.	196
7.4.2. Sincronía y reloj	199
7.4.3. Ecuador.	200
7.4.4. Seguidor de fase (Phase Tracker).	202
7.4.5. Decodificador Trellis.	203
7.4.6. Des-entrelazado de datos.	208
7.4.7. Reed-Solomon FEC.	210
7.4.8. Des-aleatorización de datos.	210
7.5. Transporte y demultiplexaje.	211
7.6. Decodificador de Video	214
7.6.1. Buffer	215
7.6.2. Demultiplexaje	215
7.6.3. Decodificación espacial	218
7.6.4. Reconstrucción de la señal de video 4:2:0	225
7.6.5. Compensación de movimiento y predicción de imágenes	226
7.6.6. Reordenación de imágenes.	228
7.7. Decodificador de Audio	229
7.7.1. Tiempo de sincronización del decodificador	230
7.7.2. Procesamiento de datos en el decodificador	230
7.7.3. Resumen de la memoria del decodificador	233

Contenido	Página
7.7.4 Decodificando la Trama de bits AC-3	234
Referencias capítulo # 7	247
8. DISEÑO DE UN RECEPTOR PARA DTV	249
8.1. Planteamiento teórico.	251
8.2. Componentes necesarios para el diseño	251
8.2.1 Antena	251
8.2.2 Sintonizador (tuner)	252
8.2.3 Demodulador VSB	255
8.2.4 Transporte, demultiplexaje y decodificación de vídeo.	261
8.2.5 Decodificación de audio.	267
8.2.6 Convertidor Digital a Analógico para audio y vídeo	269
8.2.7 Procesador Maestro (Host)	274
8.3. Fuente de alimentación.	276
8.4. Tarjeta decodificadora de DTV.	276
8.4.1 Lista de componentes	277
8.4.2 Características del decodificador	282
Referencias capítulo # 8	283
Conclusiones	285
Bibliografía	289

PRÓLOGO

En la actualidad, México necesita en cuestión de tecnología una oportunidad para demostrar sus capacidades, pensamos que dicha oportunidad dentro de la radiodifusión, es la televisión digital, ya que no existen muchos especialistas en el tema a nivel mundial, e incluso las empresas más poderosas de televisión están en incertidumbre acerca del concepto. Nosotros hemos visualizado que es inevitable la conversión y que tarde que temprano llegará a nuestro país. Tomando en cuenta este último punto, queremos estar preparados para ese momento y aprovechar la oportunidad de desarrollar el concepto para mejorarlo con calidad. Sabemos que la infraestructura nacional no permite la construcción de componentes, pero podemos aprovechar los que se encuentran en el mercado para la propuesta del diseño de un receptor de TV digital, con lo que se aprovecharían mejor los recursos destinados para este fin, dado que el desarrollo desde los circuitos integrados es un camino azaroso y pudiera ser pérdida de recursos y tiempo, debido a que existen poderosas compañías, las cuales invierten grandes cantidades de dinero en investigación y desarrollo de estos circuitos integrados.

Este trabajo ha sido el resultado de varios meses de investigación en el tema y esperamos que sirva para seguir el proceso de aprendizaje en México sobre la televisión digital y que se llegue a uno de los objetivos primordiales del Grupo Universitario de Televisión Digital, al cual pertenecemos, la construcción de equipo en torno a la televisión digital que pueda competir con los que existen actualmente, siendo el objetivo para nuestro equipo de trabajo en específico, y a nivel del grupo, la puesta en marcha de un canal propio de televisión digital.

La visión que se tiene al realizar esta tesis es la introducción en la televisión digital, para entender todos los procesos, y con base en ello poder realizar un diseño basado en los circuitos integrados que se tienen en el mercado, con el fin de igualar en calidad y precio a los decodificadores existentes a nivel comercial. Siendo esto un punto de partida para próximas investigaciones en el país relacionadas al tema, ya que estamos conscientes que no abarcamos todo lo concerniente al tema debido a que es muy extenso y que no disponemos de una base muy bien definida.

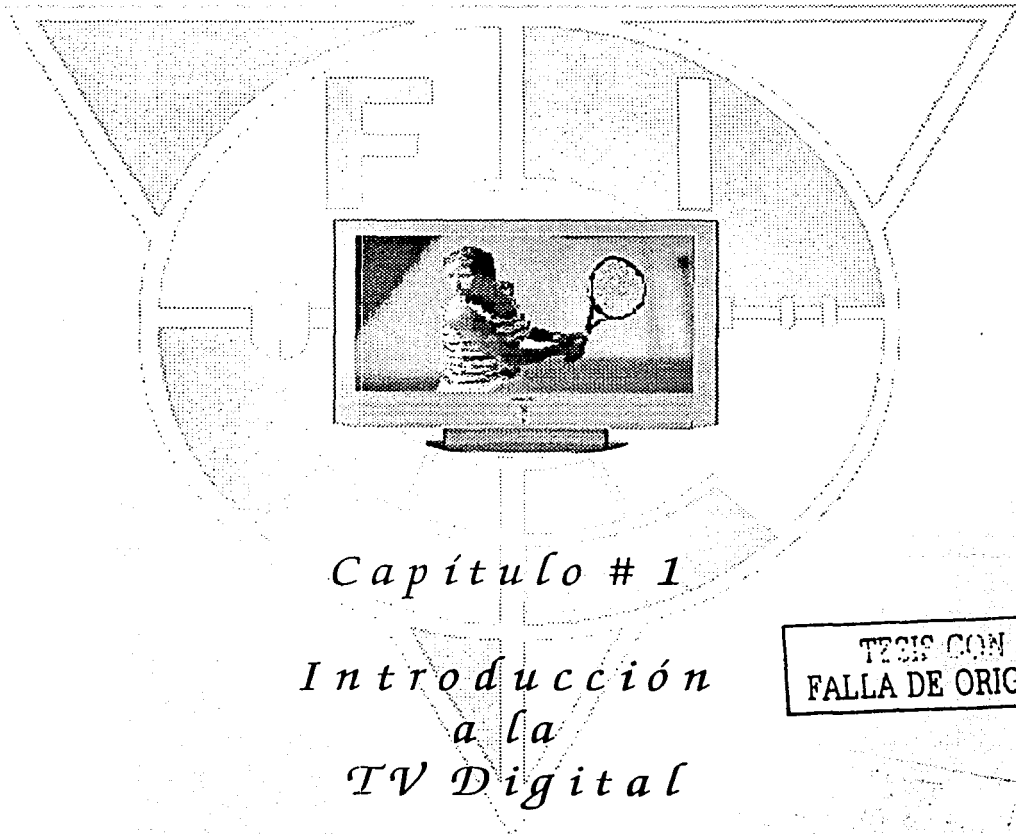
El objetivo de esta tesis es al final llegar al diseño de un aparato decodificador para TV digital, el cual opera transformando una señal en estándar ATSC de TV digital a una señal en estándar NTSC de TV analógica, para lo cual se invirtieron varios meses para recopilar, estudiar y entender, los distintos procesos que con lleva esta tecnología. La información obtenida aprendida e incluida en esta tesis proviene de varias fuentes como son: internet, libros, artículos, exposiciones, conferencias, hojas de especificaciones de circuitos integrados y sobre todo los documentos oficiales de la norma ATSC.

Prólogo

Cabe mencionar que también se llevo acabo un análisis comparativo entre los estándares existentes, así como de circuitos integrados que realizan las operaciones necesarias para conformar un decodificador de las características que apunta esta tesis. El diseño incluye desde la propuesta de la antena a utilizar para la recepción, así como los elementos que conformarn la fuente de alimentación, sintonizador de canales, etc., es decir cada una de las etapas que conforman al decodificador digital.

Sin pecar de vanidad, consideramos que este trabajo es de gran ayuda para toda aquella persona con elementos básicos en ingeniería que desee conocer el funcionamiento, y operación de el nuevo sistema de televisión digital ATSC, que seguramente se implantara en nuestro país.

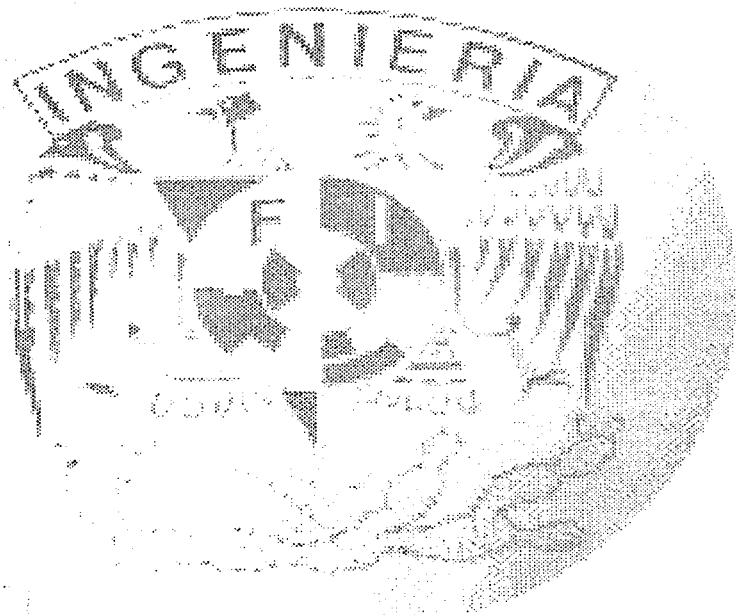
PAGINACIÓN DISCONTINUA



Capítulo # 1

*Introducción
a la
TV Digital*

TESIS CON
FALLA DE ORIGEN



TESTEADO
FALLA DE ORIGEN

1

En este capítulo se analizará el concepto de televisión digital en su forma global, tratando de dar un punto de vista del impacto de esta tecnología en la sociedad e intentando predecir el resultado que tendrá en ella, tanto en costos como en el desarrollo de esta nueva tecnología.

1.1 ¿Qué es la TV digital?

La TV Digital es la nueva tecnología desarrollada para mejorar el servicio de televisión actual en todos los aspectos, es el futuro de la televisión en todo el mundo ya que vendrá a sustituir a la antigua televisión analógica, no sólo en equipos, imagen y calidad, sino también con un nuevo concepto para el televidente, la interactividad. La televisión digital (también conocida por sus siglas en inglés como Digital TV o DTV) intenta reemplazar la televisión analógica transmitiendo una señal con mayor definición, menor ruido, formato de cine, audio surround, entre otras. Es decir, la DTV es la evolución de la TV como se conoce en la actualidad, una transición que ha tardado en darse, si tomamos en cuenta los avances y cambios en otras tecnologías como las PC's, telefonía móvil, la electrónica, las comunicaciones, etc.

La televisión digital envía una señal digital comprimida en el mismo espacio usado para transmitir la señal analógica de hoy en día. Los canales digitales de la DTV, pretenden reemplazar en algún momento los actuales canales analógicos que han venido operando por más de cincuenta años. Esta nueva tecnología digital permite transmitir una corriente de datos de 19 Mega bits por segundo, en la cual se pueden combinar múltiples señales digitales de audio, vídeo y datos. Esto es equivalente a juntar 400 módems de 56k. Con el objeto de incluir una inmensa cantidad de información en esa corriente, la señal es "comprimida" para ser almacenada y transmitida.

La llegada de la televisión digital supone un cambio tan radical como lo fue el cambio del televisor de blanco y negro al televisor de color. Se trata de conseguir imágenes mejores, pero no se queda ahí, sino que también se van a abrir las puertas a la futura introducción de servicios hasta ahora inimaginables, como la recepción móvil de televisión, la interactividad, la televisión a la carta o los servicios multimedia tan de moda hoy en día con la explosión de Internet. Existen tres formas de transmitir y recibir DTV hasta ahora, las cuales son: vía satélite, vía cable y la más reciente vía terrestre, es decir por medio de la radiodifusión como hasta ahora se hace con la TV analógica, usando como medio de transmisión el espacio, en esta tesis el análisis se enfocará al sistema de DTV terrestre.

Los servicios de TV Digital por satélite comenzaron en 1996, con un desarrollo desigual según la estrategia de implementación seguida y con los resultados más espectaculares en Francia, con tres plataformas y más de un millón de abonados en su conjunto. Un ejemplo de estos servicios son los ofrecidos por empresas como SKY y DirecTV.

Los primeros pasos para la introducción de la tecnología en la DTV por cable se comenzaron en 1996, ligados a la necesidad de introducir servicios interactivos avanzados como forma de dinamizar el mercado.

Actualmente existen tres estándares de televisión digital terrestre en el mundo:

- ATSC(Advanced Television Systems Committee) de Estados Unidos^[1]
- DVB-T(Digital Video Broadcasting - terrestre) de Europa^[2]
- ISDB-T(Integrated Services Digital Broadcasting - terrestre) de Japón^[3]

En esta tesis nos enfocaremos a las aplicaciones que ofrece el estándar de la ATSC debido a que consideramos será el estándar que adoptará México por razones que se explicarán durante el desarrollo de la misma. El estándar japonés se mencionara muy poco dentro de esta tesis debido a que su desarrollo aun es teórico y no se han implantado sistemas prácticos de transmisión para este sistema hasta el momento, ya que Japón considera que el mundo aun no está preparado para esta transición de TV analógica a TV digital, por que hasta el momento los equipos de transmisión y recepción son de costo muy alto.

Realizando un recorrido a través de la historia, en septiembre de 1995, ATSC aprobó el estándar de televisión digital A/53. Posteriormente, en diciembre de 1996 la Comisión Federal de Comunicaciones (FCC) de Estados Unidos adoptó los mejores elementos de este estándar para su uso en las radiodifusoras de la televisión digital terrestre; para el siguiente año la FCC comenzó a otorgar canales de 6 MHz a televisoras en dicho país para que pudieran ofrecer el servicio de televisión digital terrestre en paralelo con los servicios analógicos existentes. En enero de 1996, ATSC comenzó a trabajar conjuntamente con varios países de todo el mundo para que usaran su estándar en los servicios de televisión digital, de esta manera se fueron agregando países como Canadá, Corea del Sur, Taiwán y Argentina. Actualmente la ATSC tiene aproximadamente 200 miembros de una variedad de países de América, Europa, Asia y Oceanía.

El desarrollo de los sistemas de televisión digital en Europa tuvo su origen con la creación del grupo de trabajo DVB, en enero de 1995, el proyecto DVB publicó las especificaciones con las cuales se define el nuevo sistema de transmisión digital, este estándar constituye la base técnica para la transmisión digital en Europa, Asia, Australia y algunas otras regiones del mundo que comenzaron en 1996. La difusión terrestre arrancó antes de finales de 1998 en el Reino Unido y Suecia. El DVB está promoviendo fuera de Europa su sistema DVB-T(DVB funciona en las tres plataformas, T-terrestre, C-cable y S-satelital) para difusión terrenal como un sistema flexible, con capacidad para televisión de alta definición, adaptable a canales de diferente ancho de banda y susceptible de ser utilizado en red isofrecuencia_(que operan a la misma frecuencia) con cobertura de todo un país. El estándar DVB no contempla la televisión de alta definición, en lugar de ello le da prioridad a la cantidad de canales adicionales que se pueden manejar mediante la compresión de audio y vídeo, y el uso del formato SDTV(Televisión Digital Estándar), en contraste con ATSC que puede manejar tanto televisión de alta definición (HDTV) y también SDTV.

Para el sistema DVB, los tres sistemas de difusión (satélite, cable, terrestre) tienen en común la señal fuente MPEG-2, así como algunas partes de codificación, tales como la protección contra errores de código de bloque empleado Reed Solomon 204,188, algoritmo para realizar la dispersión de energía y entrelazado (interleaving). Sin embargo los sistemas de modulación empleados son dependientes del medio de transmisión.

1.2 Ventajas de la DTV sobre la TV analógica

La DTV permite que la frecuencia asignada del espacio radioeléctrico que se utiliza para transmitir una señal analógica pueda, mediante la digitalización y compresión de canales, difundir por la misma frecuencia varias señales. En simples palabras, una señal de DTV requiere aproximadamente 1/6 a 1/8 del ancho de banda ocupado por una transmisión de TV analógica de calidad equivalente.

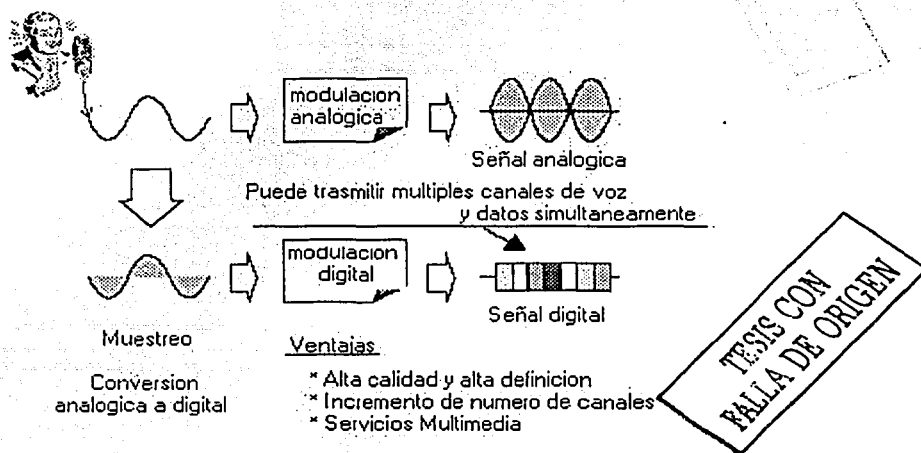
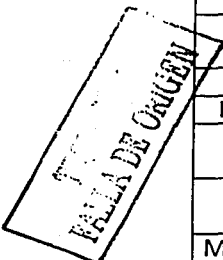


Figura 1.1 Señal analógica y señal digital

Al igual que en Estados Unidos, en México las concesiones de radiodifusión televisiva terrestre asignan a las emisoras de TV una frecuencia de 6MHz. Las transmisiones analógicas actuales requieren 3,8Mb/s, homologables a unos 5MHz. Con las técnicas actuales de compresión de señales digitales (MPEG-2), la DTV ofrecerá gran flexibilidad a las transmisoras en su trabajo de manejo del canal de 6 MHz. Las concesionarias podrán emitir un único canal digital con programación de alta definición (HDTV), o bien usar menor resolución para transmitir simultáneamente cuatro o cinco canales (multicasting) de definición digital estándar (SDTV). La calidad de vídeo de esta última es notablemente superior a la de las actuales transmisiones analógicas, y por si eso fuera poco, ofrece sonido digital.

Capítulo # 1

El principal problema de la televisión analógica es que no saca partido del hecho de que en la mayoría de los casos, las señales de video varían muy poco al pasar de un elemento de imagen (pixel) a los contiguos, o por lo menos existe una dependencia entre ellos. En pocas palabras, se derrocha espectro electromagnético. Además al crecer el número de estaciones transmisoras, la interferencia pasa a convertirse en un grave problema. En la tabla 1.1 se muestra una comparación técnica entre ambos sistemas.



	TV analógica	DTV
Parámetro	NTSC	HDTV-High Definition TV
Total Líneas	525	1125
Líneas Activas	486	1080
Sonido	2 Canales	5.1 Canales (Surround)
Relación de Aspecto	4 x 3	16 x 9
Max. Resolución	720 x 486	1920 x 1080

Tabla 1.1 Comparación entre TV analógica y DTV

Las ventajas que ofrece la TV digital frente a la analógica convencional son:

- Mayor calidad.
- Mantenimiento de la calidad extremo a extremo.
- Mayor capacidad y flexibilidad de envío de información.
- Servicios interactivos adaptados a cualquier entorno (empresa, doméstico, público,...)
- Integración de servicios: video + audio + datos.
- Nuevos servicios (por ejemplo: información adicional en texto relacionada con los programas emitidos).

Como se puede observar la DTV supera en todos los aspectos a la TV analógica ya que en resumen ofrece mejor calidad de video, audio y un mejor manejo del espectro radioeléctrico, debido a que es capaz de transmitir ocupando menor ancho de banda. La única desventaja hasta ahora de la DTV con respecto a la analógica es el precio, dado que por el momento los equipos son muy costosos pero con el tiempo y el aumento de la demanda los precios irán disminuyendo como sucede con todos los aparatos electrónicos.

1.3 ¿Qué implica la transición de TV analógica a TV digital para el usuario?

Los televisores actuales no permiten la recepción de la nueva señal digital. Para lograr decodificar la señal digital y poder desplegar las imágenes en el aparato de televisión, existen dos soluciones:

- La solución lógica es comprarse un televisor digital, pero hasta el momento este sistema no está completamente desarrollado, por lo cual los televisores aptos para recepción digital son muy caros.
- La solución más económica es añadir al receptor de televisión común un aparato decodificador digital (Set-Top-Box) que convierta la señal digital en una señal analógica y así seguir utilizando el antiguo televisor analógico. Aunque el espectador no percibirá la calidad propia de la televisión digital, la calidad de la imagen superará la que tendría el mismo programa transmitido por un canal analógico.

Como ejemplo de lo mencionado anteriormente, en los Estados Unidos donde ya se transmite en DTV de alta definición desde el año de 1998, pocos consumidores parecen haberse interesado en el nuevo medio, ya que para ver Televisión de Alta Definición según el sistema ATSC se tiene que realizar una inversión extra de dinero por parte del televidente, pues hay que agregar un decodificador o Set-Top-Box para adaptar correctamente la señal de DTV a un televisor analógico. Lo anterior se repite para el estándar DVB implementado en Europa. Con esto se observa que la aceptación que el usuario tenga para el nuevo sistema dependerá de la rapidez con la cual los costos de los decodificadores digitales puedan bajar.

Esta tesis estudia la factibilidad de diseñar un decodificador digital, analizando la posibilidad de que estos receptores, mejor conocidos como Set-Top-Box, sean desarrollados y fabricados en nuestro país, con la calidad y capacidad necesarias para competir en el mercado.

1.3.1 Cambios que debe hacer la estación transmisora.

Prácticamente se deben cambiar todos los aparatos utilizados para la transmisión. Hay que cambiar cámaras, equipos de grabación y sistemas de post-producción, además de capacitar a su personal para el manejo del nuevo equipo. La inversión necesaria es muy grande por lo que un canal público o de servicios tal vez no podría costear esta transición.

1.4 Beneficios que genera el cambio de TV analógica a TV Digital

En efecto, el canal de 6 MHz es capaz de transportar, mediante la modulación de tipo 8-VSB (banda lateral residual), un canal de datos de alta velocidad paquetizada de hasta 19.4 Mbps directo al hogar, en el cual no sólo es capaz de entregar en vídeo, sino también servicios de información a muy altas tasas, con posibles retornos mediante línea telefónica.

El empleo de la televisión digital terrestre como medio para la difusión de televisión proporciona una serie de beneficios frente a otras posibles opciones:

- Al utilizar como medio de difusión la red terrestre nos permite una recepción sencilla en el hogar y que será poco costosa conforme se desarrolle esta

tecnología, ya que emplea el mismo sistema de recepción de la televisión analógica, e incluso con la antena anterior, sin merma de calidad.

- Permite la recepción portátil y en movimiento.
- Puede emplear redes de frecuencia única lo que conlleva el uso de un menor número de frecuencias.
- Requiere menor potencia de transmisión.
- Incrementa el número de programas con respecto a la televisión analógica actual, permitiendo múltiples programas y servicios multimedia en cada canal radioeléctrico.
- Mejora de la calidad de la imagen y del sonido (se evitan los efectos de nieve y de doble imagen de la televisión analógica) en la zona de cobertura, consecuencia de la robustez de la señal digital frente al ruido, las interferencias y la propagación multitrayecto.
- La elevada resolución espacial de un sistema de televisión digital permite un realismo mayor, que se puede apreciar en una pantalla más grande.
- Permite el aumento de la relación de aspecto. El formato convencional es de 4:3, mientras que con la televisión digital se permite el formato panorámico de 16:9.
- Se puede ofrecer un sonido multicanal, con calidad de disco compacto. Además la multiplicidad de canales de audio permite conseguir el efecto de sonido perimétrico empleado en las salas de cine. Aparte, estos canales podrían emplearse para transmitir diferentes idiomas con el mismo programa de vídeo.
- Abre las puertas del hogar a la Sociedad de la Información, debido a que permite la convergencia TV-PC. El televisor pasará a convertirse en un terminal multimedia que podrá admitir datos procedentes de los servicios de telecomunicaciones, suministrando servicios de valor añadido como correo electrónico, cotizaciones de bolsa, videoteléfono, guías electrónicas de programas (EPG), vídeo bajo demanda, pago por evento, teletexto avanzado, banco en casa, tienda en casa, etc.
- Facilita los servicios de ámbito nacional, regional y local.
- Permite el desarrollo equilibrado entre servicios de TV abierta (Servicio Universal) y servicios de pago.

Estas características técnicas traerán consigo un gran número de ventajas no sólo a los usuarios, sino también al resto de los participantes que están estableciendo la TV digital.

- Fabricantes: aparecerán nuevas oportunidades de negocio.
- Radiodifusores: menor costo de alquiler de transpondedores por programa y mayor oferta a los usuarios.

- Operadores: mayor oferta de canales debido al aprovechamiento del ancho de banda.
- Instaladores: reformas y actualizaciones de instalaciones.

Finalmente cabe destacar la flexibilidad de las técnicas empleadas y su capacidad evolutiva hacia nuevos horizontes como la HD (alta definición), televisión interactiva o la fusión con el universo multimedia. Ver figura 1.2.

Las aplicaciones interactivas que se pueden tener son:

- Guía de programación
- Miniguía
- Guía Radio/Música
- Concursos, participación en juegos
- Canal meteorológico
- Canal de tráfico
- Descarga software por abono del canal
- Descarga software con pago por paquete
- Telebanca
- Telecompra
- Transacciones o pagos bancarios
- Acceso Web
- Publicidad sincronizada
- Publicidad personalizada
- Juegos
- Mensajería
- Pago por evento (PPV, IPPV)

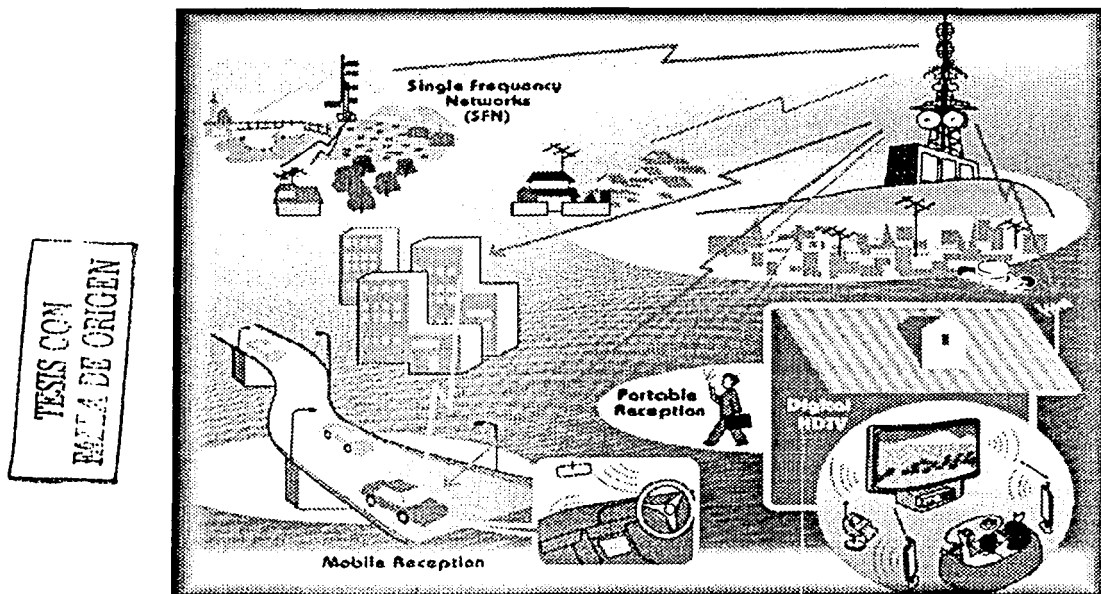


Figura 1.2 Servicios de TV digital.

1.5 Consecuencias socio-culturales de la DTV

Como es conocido por todos, la televisión es un medio de comunicaciones que junto con la radio tienen el mayor número de audiencia en el mundo, y nuestro país no es la excepción, en la mayoría de los hogares se cuenta con un aparato de televisión con el cual millones de personas pasan mucho tiempo frente a él. Este medio tiene gran influencia en el comportamiento y manera de pensar de los individuos, por todo lo que muestra (sus personajes, programas, etc.). En pocas palabras es un sistema de comunicaciones que puede influir en la cultura de la gente tanto para bien, como para mal, ya que es capaz de imponer modas, comportamientos, prototipos, héroes, pero también puede transmitir conocimientos.

Por esto el crear un nuevo sistema de televisión, o mejor dicho un nuevo concepto de televisión, tendrá un gran impacto cultural en las sociedades del mundo, debido a que ahora además de una mejor calidad de audio y video, que atraparán aun más a su televidente, el usuario dispondrá de más canales y lo más innovador es que podrá interactuar con su televisor. Ahora no sólo será un espectador si no que contará con un sistema interactivo, en el cual podrá realizar compras en el momento de ver el producto, participar en concursos, y muchos servicios interactivos más.

Este sistema puede enfocarse al consumismo, ya que el gran negocio de las marcas publicitarias se incrementará con este sistema, al poder "comprar el producto que estás

viendo, en el momento que lo ves". Además los canales culturales, como por ejemplo el canal 22 y 11 en México, estarán en gran peligro, puesto que su capacidad económica de respuesta para el cambio de sus instalaciones, de analógicas a digitales se encontraría fuera de sus alcances por el momento. De ahí la importancia del apoyo que los gobiernos de los países puedan proporcionar a estos canales. La televisión puede ser un arma de dos filos por que también se puede aprovechar para proporcionar educación a toda la población, y en poblados alejados se pueden impartir mediante el televisor programas educativos, y por que no, clases interactivas en tiempo real.

1.6 Situación actual de la TV digital

Considerando la aceptación del cambio de televisión en blanco y negro a la de color, de discos de vinilo a discos compactos, y de videocintas a discos ópticos DVD, queda claro que la tecnología de DTV será igualmente adoptada. Lo anterior ha sido corroborado por las reacciones de consumidores que han tenido la oportunidad de asistir a una demostración de televisores digitales de alta resolución. En México las empresas *TV Azteca* y *Televisa*, organizaron demostraciones de esta nueva tecnología durante los meses de diciembre de 1997 y febrero de 1998, respectivamente. Este cambio no sólo representa la oportunidad de revolucionar la calidad de la imagen y del sonido, sino que presenta un gran potencial de interactividad.

En el caso de DVB a continuación se muestran las tablas 1.2 y 1.3 donde aparece un resumen detallado por país de la forma en que se está implantando dicho sistema en Europa.












Pais	 UK	 Suecia	 España	 Alemania	 Dinamarca
Fecha de lanzamiento	15/11/98	1/4/99	1999 / 2000	?	?
Cobertura para el lanzamiento	70% / 90%	50%	20% / 50%	?	?
Fin analog.	2015?	2008 / 2012	1/1/2012	2010 / 2015	?
Transmitiendo en pruebas*	On-air	3	3	3	3
Servicios					
Subscripción*	24	8	15		
PPV*	3	7	3		
16:9*	3	3	3	?	3
8k/2k	2k	8k	8k	8k	8k
SFN/MFN	MFN	MFN & SFN	SFN & MFN	MFN & small SFN	SFN

Tabla 1.2 Servicio Europeo para la Televisión Digital Terrestre.

TESIS CON FALLA DE ORIGEN

País	 Irlanda	 Finlandia	 Francia	 Italia	 Holanda	 Noruega
Fecha de lanzamiento	9/2000	10/2000	End 2001?	?	2000/2002	2000
Cobertura para el lanzamiento	95%	50%	65%/ 85%	?	18%/ 52%	65%
Fin analog.	2009/2015	2015?	2015?	?	2010?	?
Transmitiendo en pruebas*	3	3	3	3	3	3
Servicios						
Subscripción*	25	3			2	?
PPV*	3	3			?	?
16:9*	3	3	?	Algunos	3	?
8k/2k	8k	8k	8k	8k	8k	8k
SFN/MFN	MFN Gap SFN	MFN Gap SFN	MFN	SFN & MFN	SFN	MFN Gap SFN

* Cantidades en número de televisoras

Tabla 1.3 Servicio Europeo para la Televisión Digital Terrestre.

En Estados Unidos opera sobre todo en la banda de 470 a 890 MHz (canales 14 a 83) y en frecuencias UHF. El sistema nuevo y el antiguo deberán coexistir hasta el 2006, en que deben cesar las emisiones de señales NTSC, tanto en la banda de 54 a 216 MHz (canales 2 a 13), como en frecuencias UHF y VHF. La FCC reasignará entonces dichos canales a la televisión digital. Se estima que para el año 2006, 30% de los 100 millones de hogares en los EEUU, tendrán un receptor de televisión digital.

Los televisores tradicionales continuarán recibiendo señales analógicas transmitidas al espacio por lo menos hasta el año 2006 en los EU, y mucho tiempo después, en nuestro país.

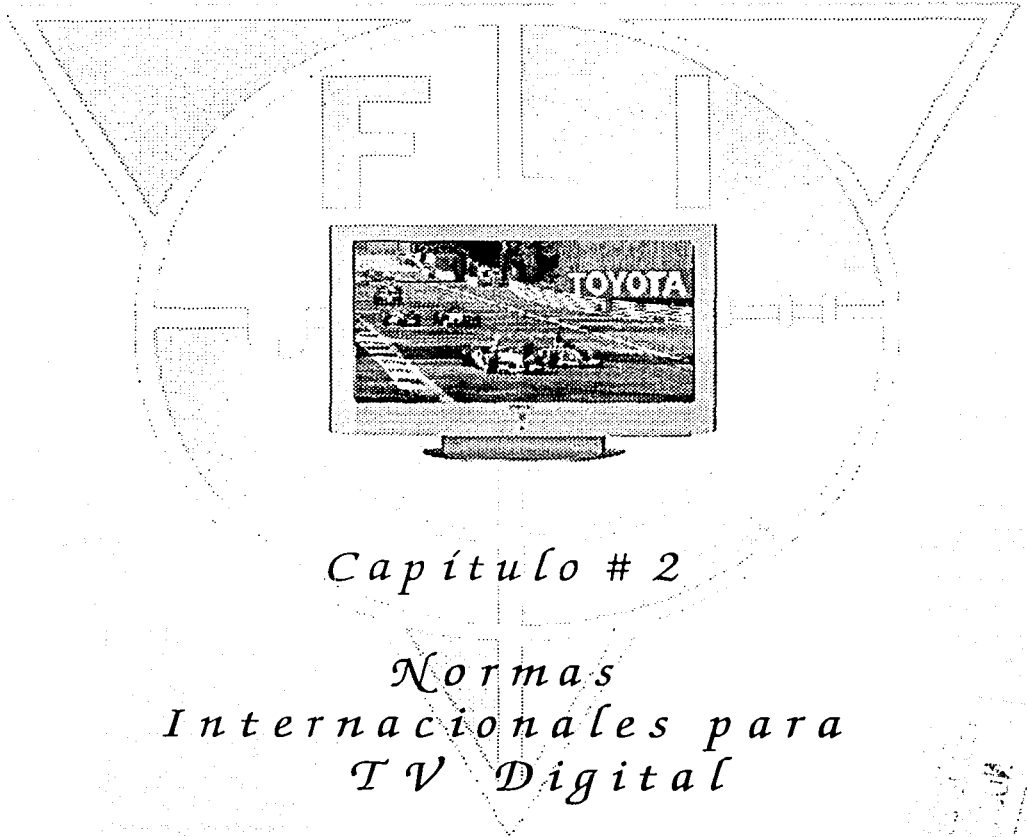
Una vez terminado el cambio a transmisiones digitales, los televisores de hoy serán capaces de recibir las nuevas señales digitales sólo mediante el uso de una caja decodificadora adicional (Set-Top-Box). Sin embargo, otras tecnologías actuales, tales como el cable, las videograbadoras, los discos ópticos DVD y los sistemas de recepción satelital directo al hogar, continuarán funcionando con nuestro "viejo" televisor NTSC por algunos años más.

TESIS CON FALLA DE ORIGEN

El sistema Japonés ISDB tiene muy poco desarrollo al momento, los países que implantaran o adoptarán este sistema, hasta el momento conocidos son Japón y China, aunque se cree que este sistema reemplazará al DVB Europeo ya que ISDB es básicamente una mejora del DVB, se puede decir que ISDB es el DVB mejorado.

TESIS CON
FALLA DE ORIGEN

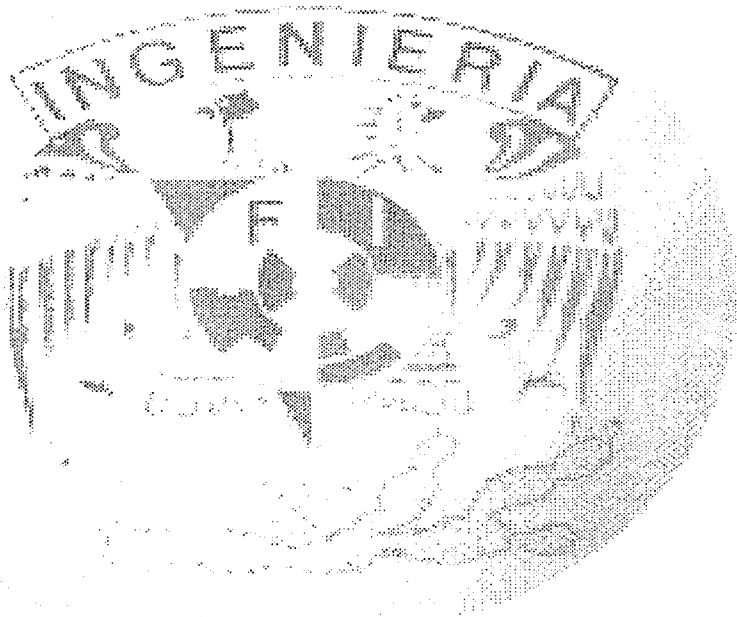
TESIS CON
FALLA DE ORIGEN



Capítulo # 2

*Normas
Internacionales para
TV Digital*

1900 21297
FALLA DE ORIGEN



TECIS CON
FALLA DE ORIGEN

2

2.1 Estándares ATSC, DVB e ISDB.

Actualmente existen tres estándares comerciales de televisión digital en el mundo, el ATSC(Advanced Television Systems Committee)^[1] desarrollado en Estados Unidos, el DVB (Digital Video Broadcasting)^[2] en Europa y el ISDB (Integrated Services Digital Broadcasting)^[3] en Japón. En el caso de los Estados Unidos lo primordial del estándar es darle mayor calidad a la imagen, lo que lleva al sistema HDTV (High Definition TV), debido a que el costo de esta tecnología es elevado, se espera que su transición sea más lenta que la de los otros estándares.

El estándar DVB no contempla la televisión de alta definición HDTV, éste estándar le da prioridad a obtener un mayor provecho del ancho de banda, es decir obtener el mayor número de canales posibles adicionales, que se pueda de uno solo canal analógico, mediante el manejo digital del audio y video, es decir su compresión, esto nos lleva al sistema SDTV(Standard Definition TV)

El estándar ISDB de Japón contempla de igual manera la televisión de alta definición y la adquisición de más canales. Su formato de compresión de video es el MPEG-2^[4] (Moving Picture Expert Group) y usa una modificación del COFDM (Codec Orthogonal Frequency Division Multiplex). Para el audio se usa el AAC (Advanced Audio Coding).

2.1.1 El estándar americano ATSC.

❖ Introducción

La FCC (Federal Communications Commission) de Estados Unidos decidió en 1987 crear la comisión ATSC para que la asesorara en televisión avanzada: formato HDTV y transmisión digital. En 1993 se revisaron los resultados y la industria forma la GA (Grand Alliance) que en 1994 selecciona los mejores elementos de las distintas propuestas presentadas en ATSC y los resultados se presentan en 1995 a la FCC. En 1997 la FCC tomó la decisión de pasar del sistema NTSC (National Television Systems Committee) al ATSC-DTV (Digital TeleVision) que incorpora el formato HDTV digitalizado basándose en el MPEG-2. En el año 2006 el sistema NTSC desaparecerá de acuerdo con las normas actuales en ese país.

❖ Procesamiento de la información en ATSC.

A continuación se enumeran las distintas etapas de procesamiento de la información de imagen y audio hasta llegar a la señal compuesta de video de 6 MHz de ancho de banda. La secuencia de procesos se observa en la figura 2.1.

TESIS CON FALLA DE ORIGEN

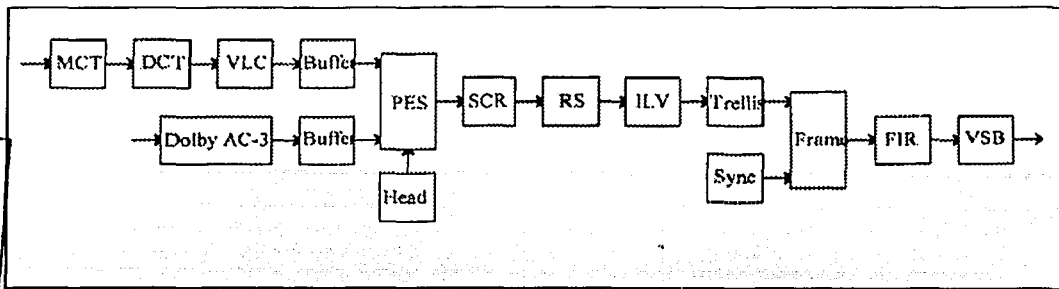


Figura 2.1 Diagrama General del codificador ATSC.

Frecuencia de barrido. Los estudios llevados a término para la industria del cine habían llevado la velocidad de cuadro a 24 Hz. Para la televisión de Europa se tomó 25 Hz derivado de la frecuencia de línea eléctrica de 50 Hz. En USA este valor es de 30 Hz derivado de la frecuencia de línea de 60 Hz. Par evitar causas de disturbios en el scanner y el procesamiento de la señal, el valor de 30 Hz se puede reducir en el 0,1% a 29,97 Hz.

Formato de imagen. Se disponen de varios tipos de formatos de imagen, entre ellos el 1920x1080. Esto genera una estructura del tipo 16:9. Otros formatos son el 1280x720 y el 704x480. En la estructura de la televisión actual 4:3 se tiene el formato 640x480. El sistema de codificación de video requiere que el número de líneas en el área de imagen sea un múltiplo de 32 para el formato entrelazado y de 16 para el no entrelazado. Por ello el caso de 1080 líneas verticales el resultado es 1088 líneas de las cuales 8 líneas no se utilizan. Ver tabla 2.1.

Tamaño de imagen	Relación de imagen	Cuadros/seg
1920H x 1080V	16:9	30-24
1280H x 720V	16:9	30-24
704H x 480V	16:9	30-24
640H x 480V	4:3	30-24

Tabla 2.1. Formatos de imagen para DTV de ATSC.

Componentes de luminancia y crominancia. Las componentes de luminancia Y y crominancia Cb y Cr se soportan sobre los colores básicos G -R -B . Se trata de:

$$\begin{aligned}
 Y &= 0.59G + 0.30R + 0.11B \\
 Cb &= 0.89B - 0.59G - 0.30R \\
 Cr &= 0.70R - 0.59G - 0.11B
 \end{aligned}$$

Muestreo. El tipo de muestreo se escribe como 22:11:11 (corresponde al valor múltiplo de 3,375 MHz) para la producción de imágenes de HDTV. Esto corresponde a 74,25 MHz para luminancia Y y 37,625 MHz para cada componente de crominancia R-Y y B-Y. Cada 4 muestras de luminancia se tienen dos de crominancia, una de R-Y y otra de B-Y. Las muestras restantes de crominancia son interpoladas.

El resultado final para la emisión en broadcasting a 6 MHz de ancho de banda, utiliza la codificación MPEG-2, que agrupa las muestras en bloques de 8x8. Los bloques son organizados en macrobloques de 2x2 bloques. Esto permite obtener 16x16 muestras de luminancia y 8x8 de crominancia.

Compresión de imagen MPEG-2. Se trabaja con macrobloques (2x2 bloques). La tasa de compresión de imágenes es del orden de 50:1; desde 1 Gb/s a 20 Mb/s. La codificación utiliza las técnicas de MPEG-2. En MPEG solo se establecen la sintaxis y la semántica dejando la implementación a cada aplicación. Se trata de un modelo de referencia que utiliza las siguientes técnicas:

- Definición de imágenes de tipo interframe I, predictiva P e interpolada B.
- Codificación de redundancia espacial mediante la transformada **DCT**^[5] (Transformada Discreta del Coseno).
- Codificación de redundancia temporal. Compensación de movimiento **MCT** (*Motion Compensation Transform*).

Frame I-P-B. La interpolación de imágenes permite reducir la cantidad de imágenes que son codificadas. Algunas de las imágenes son usadas de referencia y otras son el resultado de interpolación.

Intraframe (cuadro I). Se trata de cuadros con compresión moderada para usarlas como referencia:

- No poseen predicción temporal de movimiento MCT como en el caso de las imágenes intermedias.
- Posee solo compresión espacial mediante codificación DCT.

Imagen Predictiva (cuadro P). Se codifica basándose en la I anterior (compensación MCT) y sirve para el P siguiente.

Imagen interpolada (cuadro B). Calculada basándose en las imágenes I y P anteriores y posteriores.

La secuencia de imágenes es: I-B-P-B-P-B-P-B-P-B-P-B. Periódica cada 12 imágenes.

Motion Compensation Transform. MCT. Es un método que permite compensar los cambios de posición de las imágenes con un corrimiento (offset) pequeño. Consiste en una comparación del macrobloque actual con el de la imagen anterior en el tiempo (predicción

interframe). El offset es un corrimiento sobre los ejes x-y para estimar los movimientos de las imágenes. Ver figura 2.2 a).

- MCT se aplica en la predicción diferencial mediante un proceso multi-etapas.
- Para seleccionar el macrobloque con menor diferencia se utiliza un algoritmo de aproximaciones sucesivas simple.
- Se obtiene entonces el vector de movimiento que identifica al macrobloque más adaptado.
- El vector se codifica y se transmite junto con la matriz 8x8.
- MPEG dispone de 1 o 2 vectores de acuerdo con el tipo de MCT aplicado.

Transformada discreta del coseno DCT. Es un método de codificación que convierte una matriz 8x8 puntos X_{ij} en otra T_{mn} . Cada punto de la matriz T_{mn} , se calcula con base a toda la matriz de puntos X_{ij} . En la matriz transformada T_{mn} se encuentran los valores mucho más correlacionados que en X_{ij} . Esto permite una codificación final con menor número de bits. La matriz 8x8 resulta ser el valor óptimo como velocidad de cálculo y de correlación de coeficientes. De esta forma, un error en la transmisión de los datos T_{mn} se difunde en toda la matriz X_{ij} de recepción.

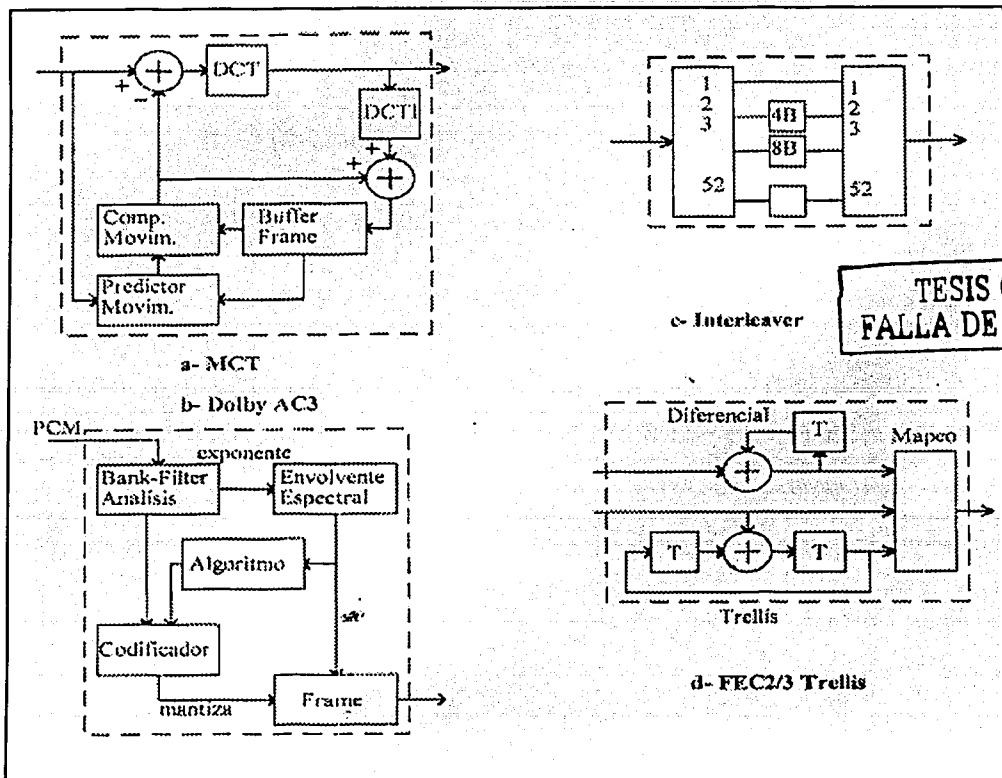
En la codificación ADPCM (modulación por codificación de impulsos diferencial adaptativa) los errores se observan como un granulado de puntos en la pantalla. En la DCT los errores se distribuyen en la matriz de 8x8 píxel y son prácticamente imperceptibles a la visión humana. En transmisiones de baja calidad (exceso de reducción de velocidad o mucho BER) se observan distorsiones por rectángulos (el rectángulo es el resultado de la matriz 8x8, o múltiplos de ella, distribuida con la diferencia de definición entre puntos horizontales y líneas verticales).

El algoritmo DCT se fundamenta en una representación similar a la frecuencia. Se denominan frecuencias normalizadas o coeficientes de transformada. El ojo humano es sensible en distinta forma a los coeficientes y por ello pueden ser tratados de distinta manera. Los coeficientes se encuentran des-correlacionados y solo unos pocos son importantes. Los coeficientes de la matriz se leen en zig-zag oblicuo desde el ángulo superior-izquierdo. Solo se transmiten los coeficientes que superan cierto umbral. El buffer de salida selecciona el proceso de cuantificación para mantener una velocidad constante y reducir el retardo.

Codificación de Longitud Variable (VLC) y buffer. La salida del codificador DCT se lee en zigzag desde el coeficiente superior derecho. Se codifican los coeficientes distintos de cero y la cantidad de coeficientes iguales a cero (Run Length Coding). Se agrega un código de final de bloque EOB (End-Of-Block). El codificador Huffman colocado a continuación permite asignar a las secuencias de mayor probabilidad un menor número de bits. Debido a que la salida de este codificador no es una velocidad constante se requiere un buffer de datos que regule la salida a una velocidad constante. En DTV el tamaño del buffer es de 8 Mbits. Cuando el buffer está por superar su capacidad (Overflow o Underflow) se procede a cambiar los umbrales de cuantificación.

Codificación de audio. La Codificación de la señal de audio se denomina **AC-3** (Audio Coding). Las características generales del proceso son las siguientes (Figura 2.2 b)):

- Son 6 canales: izquierdo, central y derecho; izquierdo y derecho *Surround* y *LFE* (Low Frequency Effects).
- El ancho de banda es de 3Hz-20 kHz con excepción del LFE que tiene 120 Hz para efectos especiales de baja frecuencia.
- La frecuencia de muestreo es de 48 kHz. La codificación se realiza en forma uniforme a 16 bits.
- Trabaja en el dominio de la frecuencia debido a que permite una mejor compresión psicoacústica.



TESIS CON FALLA DE ORIGEN

Figura 2.2 Diversos codificadores para ATSC

Los pasos de codificación para la reducción de la velocidad son:

Capítulo # 2

Banco de filtro de análisis: Ingresan 512 muestras de audio y se convierte desde el dominio del tiempo al dominio de la frecuencia.

Se utiliza una variante de **FFT** (Fast Fourier Transform) que permite remover el 100% de la redundancia.

Se disponen de 512 coeficientes de frecuencia.

Cada uno de los coeficientes es representado en notación exponencial (con valor máximo 1): mantisa y exponente.

Codificación de envolvente espectral: Los exponentes de la notación anterior se codifican para obtener la envolvente espectral.

Esta información alimenta una rutina que determina la cantidad de bits necesarios en la codificación de la mantisa.

Esta rutina se basa en un modelo psicoacústico para determinar la S/N óptima para cada coeficiente de frecuencia.

Cuantificación de mantisa. Cada mantisa se codifica desde 0 a 16 bits de acuerdo con el algoritmo anterior.

Formación de la trama AC-3: La información de la envolvente espectral y de la mantisa de los 6 canales de audio se ingresa en una trama.

Esta trama contiene 1536 muestras de audio; son 6 canales de audio por 256 muestras.

Contiene además información de sincronismo para alineación **SI** (*Sync Information*); información de servicio indicando el tipo de codificación utilizado **BSI** (*Bit Stream Information*); y paridad para control de errores del tipo CRC-16.

Frame video+audio. Se trata de formar paquetes denominados **PES** (*Packetized Elementary Stream*). Este tipo de formato fue definido para aplicaciones de memoria digital. La longitud del paquete es variable con un máximo de 216 Bytes. Contienen la siguiente información:

- 3 Bytes de inicio de paquete.
- 1 Byte para identificar el tipo de paquete.
- 2 Bytes para indicar la longitud total del paquete.
- 2 Bytes para indicadores de campos adicionales.
- 1 Byte para indicar la longitud del encabezado del PES.
- N Bytes para encabezado del PES.
- M Bytes para el bloque de datos del paquete.

Scrambler (aparato para perturbar las emisiones radiofónicas). Permite la distribución de potencia en el espectro (Figura 2.3). Sus características son:

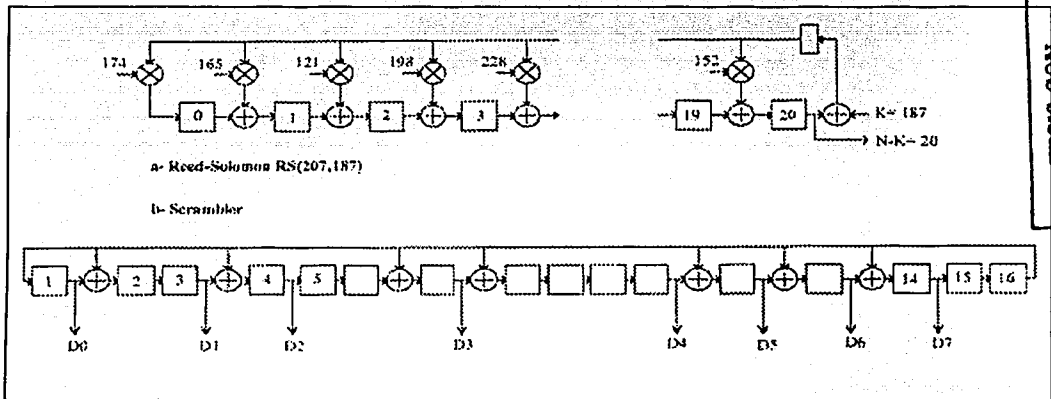
- Es del tipo sincrónico: se inicializa con el campo de datos de 1656 bits.
- No aleatoriza la información de sincronía y los campos de FEC.
- El polinomio generador es: $X^{16}+X^{13}+X^{12}+X^{11}+X^7+X^6+X^3+X+1$. Periodo de repetición de 216-1.
- Desde este generador se obtienen 8 bits en paralelo (1 Byte) desde los siguientes valores: X^{14} , X^{13} , X^{12} , X^{11} , X^7 , X^4 , X^3 , X .
- Estos se utilizan para la aleatorización de los datos (Byte en paralelo) a ser transmitidos.

Forward error correction (FEC), Reed Solomon (RS). Este sistema dispone de dos algoritmos de corrección de errores simultáneos: FEC convolucional asociado a la modulación Trellis y FEC a bloques del tipo Reed-Solomon. El FEC convolucional se encuentra luego del codificador Interleaver.

El FEC de tipo Reed-Solomon RS(208,188) cumple con la siguiente estructura (Figura 2.3):

- Cada 188 bytes de datos se calculan 20 bytes de FEC. Se obtienen 208 bytes en total.
- El procesamiento (producto y suma) es por byte en paralelo (módulo 256).
- Los bytes de paridad se calculan mediante la siguiente expresión:

$$X^{20} + 152.X^{19} + 185.X^{18} + 240.X^{17} + 5.X^{16} + 111.X^{15} + 199.X^{14} + 6.X^{13} + 220.X^{12} + 212.X^{11} + 150.X^{10} + 69.X^9 + 36.X^8 + 187.X^7 + 22.X^6 + 228.X^5 + 198.X^4 + 121.X^3 + 121.X^2 + 165.X + 174$$



TESIS CON FALLA DE ORIGEN

Figura 2.3 FEC y Scrambler para el codificador ATSC.

Interleaver (entrelazado). Este codificador permite reducir el efecto de las ráfagas de error y facilitar la corrección de errores mediante el FEC RS. Sus características principales son (Figura 2.2 c))

- Se trata de un interleaver del tipo convolucional sincronizado con inicio del campo de datos.
- Solo los datos son intercalados.
- Trabaja sobre la base de 52 segmentos paralelos y con retardo relativo entre segmentos de 4 Bytes. Total 208 Bytes.
- El interleaver introduce un retardo de 4 mseg.

Codificador Trellis. Este codificador incluye 3 etapas (Figura 2.2 d)) :

- Entrada de dos bits en paralelo. Se codifica en forma diferencial una de las líneas de datos.
- La otra línea de datos permite obtener la codificación convolucional. El resultado son 3 datos en paralelo: FEC-2/3.
- Los datos son mapeados en un símbolo de 8 estados (000= -7, 001= -5, -3, -1, +1, +3, +5, 111= +7).
- La secuencia de datos mapeados llegarán a un modulador de 8 estados.
- El codificador Trellis tiene incorporado un Interleaver para mejorar las prestaciones frente a ráfagas de errores.
- Opera con 12 codec Trellis idénticos en paralelo de forma que la salida extrae un símbolo de cada uno de ellos.
- El Codificador Trellis FEC-2/3 no se utiliza para la tasa de datos de 38,6 Mb/s (16-VSB) pero el FEC-RS es el mismo.

Datos de la trama. Se trata de formar una trama que contiene los datos codificados más la información de sincronía, la cual se utiliza para alineamiento y para el seguimiento del equalizador autoadaptativo de recepción. La trama se compone de:

- La trama de datos (Frame) puede imaginarse como una matriz con altura total de 2x313 segmentos (2x24,2 mseg).
- Cada segmento tiene un ancho de 832 símbolos (cada símbolo son 3 bits) y una duración de 77,3 μ seg.
- Los 4 primeros símbolos de cada segmento son de sincronía (Sync Segment).
- La secuencia es: 001 110 110 001 = +5, -5, -5, +5.
- Los segmentos 1 y 314 llevan 828 símbolos de campo de sincronía (Field Sync).
- Los restantes segmentos llevan 828 símbolos (total 207 Bytes) de información de datos.

SINCRONÍA. La señal de Field Sync es una secuencia que contiene los siguientes símbolos:

- 511 símbolos de secuencia Positivo-Negativo 511. Sirve para adaptar el equalizador.

- 3x63 símbolos con secuencia Positivo-Negativo N63. La secuencia central está invertida.
- 24 símbolos para el modo VSB.
- 104 símbolos reservados.
- 12 símbolos de precódigo.

Filtro FIR (Finite Impulse Response). Este filtro es del tipo coseno realzado y permite recortar el espectro a ser emitido. Se realiza mediante circuitos lógicos con el formato transversal lineal. La pendiente del filtro permite una caída de 3 dB a 310 kHz. La tasa de símbolos (3 bit/símbolo) de entrada al filtro es 10,76 MHz. La salida está recortada a 6 MHz.

Modulación VSB (Banda lateral Residual). Es la modulación responsable de colocar la señal digital dentro de la banda de 6 MHz asignada. Se puede acomodar 19,28 Mb/s mediante la modulación 8-VSB y 38,56 Mb/s mediante 16-VSB. El modulador recibe desde el filtro una señal de 6 MHz de ancho de banda y es modulada en 8 estados de amplitud. Entrega una IF (Intermediate Frequency) al equipamiento de broadcasting. Trabaja con una frecuencia piloto en 310 kHz para una mejor adquisición en recepción en condiciones de ruido.

Este tipo de modulación 8-VSB permite trabajar a 14,9 dB de S/N lo cual determina una SER (Segment Error Rate) de $1,93 \times 10^{-4}$. Esto equivale a 2,5 segmentos errados por segundo que es el umbral subjetivo de visibilidad de errores.

Señal compuesta. Se utiliza un pre-ecualizador en el transmisor para reducir la intermodulación del amplificador. Se trata de un circuito transversal lineal de 80 etapas del tipo feed-forward. Cada etapa está espaciada en 93 nseg. Como la etapa de referencia esta centrada permite una corrección en un rango de $\pm 3,7 \mu\text{seg}$. La muestra para esta corrección es tomada de la señal en el circuito de antena.

El espectro NTSC y ATSC ocupan en ambos casos 6 MHz. Sin embargo, mientras NTSC dispone de una portadora de luminancia, una portadora de crominancia y otra de sonido (Multiplexación por División de Frecuencias FDM), el ATSC solo posee una portadora suprimida y una señal piloto para referencia. Las informaciones de luminancia, crominancia y audio se multiplexan en TDM (Multiplexaje por división de tiempo). El espectro ATSC ocupa una banda efectiva desde 0.31 MHz hasta 5.69 MHz, con un total de 5.38 MHz.

Reloj del sistema. La base de temporización del sistema es un reloj de 27 MHz. Permite generar las frecuencias de video y de audio, así como las necesarias para procesamiento de la información.

2.1.2 El estándar Europeo DVB.

El DVB (Digital Video Broadcasting) es la norma de Europa para aplicaciones similares a ATSC. El proyecto se inició en 1993 y reemplazó al Eureka. Se dispone de formatos para

DVB-S(Satélite), DVB-C(Cable) y DVB-T(Terrestre). Utiliza anchos de banda de 6 y 8 MHz. Como se ha analizado la norma ATSC detalladamente, en el caso de DVB la descripción se limita a señalar las similitudes y las diferencias. Además en esta tesis solamente se analizara a fondo el ATSC por considerarse el posible estándar adoptado por México.

Las particularidades del diagrama a bloques son las siguientes:

- DVB posee las funciones a bloques, similares en muchos casos a ATSC:
- Interfaz de entrada con separación de componentes y digitalizador. La compresión de vídeo y audio es la MPEG-2. Un scrambler asegura la distribución espectral.
- Codificadores: FEC RS(204-188), etapa interleaver Outer, codificador FEC convolucional, interleaver Inner.
- Luego de la codificación FEC se procede al mapeado de símbolos Trellis y la trama de transmisión.
- Se utiliza una codificación OFDM (Orthogonal Frequency Division Multiplex) y una banda de guarda.
- Se filtra la señal mediante un FIR, se modula (quadrature phase shift keying QPSK, Quadrature Amplitude Modulation 16-QAM o 64-QAM) y se envía a la interfaz del transmisor.
- La modulación depende de la tasa del FEC convolucional (posibles 1/2, 2/3, 3/4, 5/6 o 7/8).

La trama de transmisión contiene la siguiente información (Figura 2.4):

- Los datos de vídeo y audio codificados en MPEG-2 con codificación SCR(System Clock Reference), FEC e interleaver.
- Dos tipos de señal piloto para sincronismo de tiempo (Scattered) y frecuencia (Continuous) del receptor.
- Señal con parámetros de transmisión TPS (Transmission Parameter Signaling): velocidad de guarda, tasa de FEC convolucional, tipo de modulación.
- La codificación de la señal de audio utiliza la compresión de MPEG-2 similar a Dolby AC3.
- La frecuencia de muestreo del audio es de 32, 44 y 48 kHz. La velocidad de codificación está entre 32 y 384 kb/s.

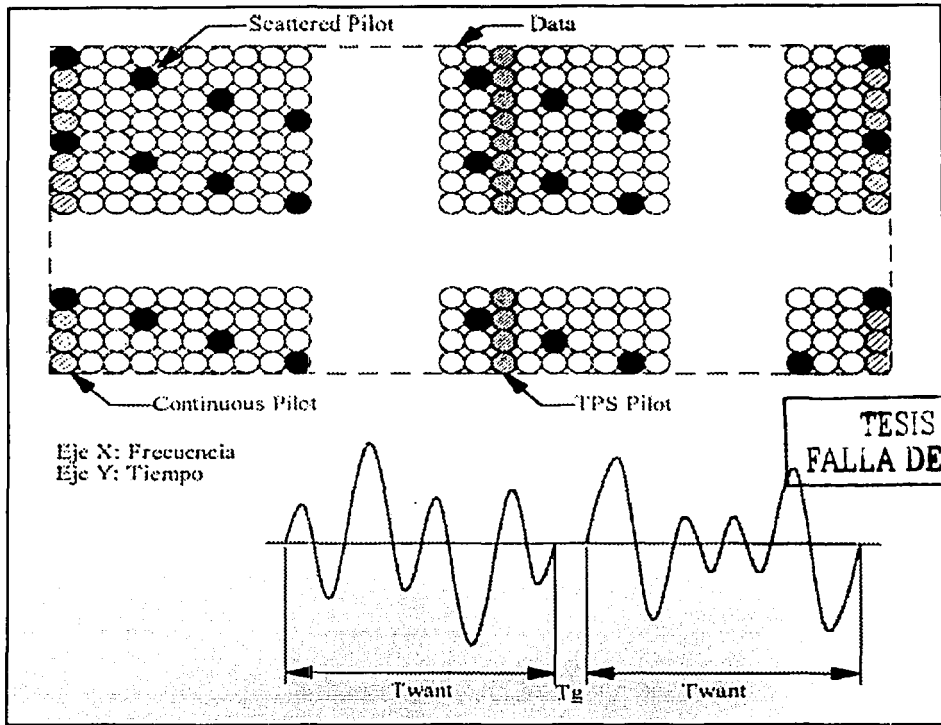


Figura 2.4 Método de modulación para el sistema DVB.

COFDM (Code Orthogonal Frequency Division Multiplex). En DVB se utiliza este método de codificación especial que es utilizado también en el sistema **DAB (Digital Audio Broadcasting)**.

Las principales características son:

- Cada grupo de N bits de datos de banda base se transmite sobre portadoras individuales separadas en frecuencia (FDM).
- En la Figura 2.4 el eje horizontal muestra N frecuencias en una banda de 6 u 8 MHz.
- El tiempo se divide en intervalos de multiplexación (T_{want}) que corresponde a N intervalos de bits (símbolo de N bits).
- Cada uno de los N bits codifica a una frecuencia distinta (S_0, S_1, \dots, S_N).
- La presencia o no de la frecuencia depende del valor del bit a ser codificado.
- El resultado de la sumatoria es una forma de onda particular para cada secuencia de N bits (ver la figura).

- Esta codificación es resistente al desvanecimiento (fading) selectivo por caminos múltiples y a interferencias de banda angosta.
- Para lograr este efecto el valor de N (T_{want}) debe ser de duración lo suficientemente alta (cerca de 1 mseg).
- Para reducir el efecto de las reflexiones (interferencia intersímbolo) se coloca entre 2 símbolos una banda de guarda (T_g).

La señal TPS (Transmission Parameter Signaling) transporta información sobre las selecciones que se han realizado. Entre ellas se encuentra el tipo de codificador convolucional FEC y la modulación adoptada, así como el tiempo de guarda. Para poder demodular correctamente esta señal TPS ocupa una frecuencia fija en forma permanente. También las señales piloto de temporización (Scattered Pilot) y frecuencia (Continuous Pilot) ocupan frecuencias fijas.

2.1.3 El estándar Japonés ISDB.

En Japón, el grupo DiBEG (Digital Broadcasting Experts Group), desarrolló también un sistema para transmisión de TV Digital. "Integrated Services Digital Broadcasting – Terrestrial" (ISDB-T), el cual está destinado a proveer los servicios de voz, audio, datos y servicios interactivos en un mismo canal. Está basado en compresión de video MPEG-2, pero con multiplexación más compleja y necesaria para manejar la variedad de servicios propuestos a transmitir. Este sistema también usa un sistema de transmisión basado en modulación OFDM (Orthogonal Frequency Division Multiplexing). La transmisión de OFDM básicamente utiliza los esquemas de modulación digital siguientes: QPSK, 16QAM, 64QAM, DQPSK.

Este estándar es muy similar a DVB-T, algunas diferencias con el estándar DVB-T son las siguientes:

- ISDB-T tiene tres modos de operación en los cuales OFDM utiliza 3 diferentes frecuencias de subportadora para su transmisión: 2K, 4K y 8K. (DVB-T no incluye el modo a 4K).
- ISDB-T usa DQPSK además de otras constelaciones usadas en DVB-T para evitar los problemas ocasionados en el receptor por el defasamiento de la señal.
- Segmentación de banda ancha para proveer posibilidad de recepción parcial.
- Estructura de frame complicada con gran cantidad de ajuste de retardo.

El sistema ISDB-T puede construir transmisión de señales con distintos anchos de banda (de una señal de banda estrecha a una señal de banda ancha) para combinar señales de banda estrecha llamados "segmentos OFDM" como se muestra en la figura 2.5. El ancho de banda de un segmento de OFDM es alrededor de 430 kHz. Los datos del programa son divididos en segmentos de datos, los cuales son procesados individualmente y transformados en señales de transmisión como se muestra en la figura 2.5. Un programa de HDTV o muchos de SDTV pueden ser acomodados en los 13 segmentos con lo cual se obtendría una señal con ancho de banda de 5.6 MHz. Ver figura 2.6.

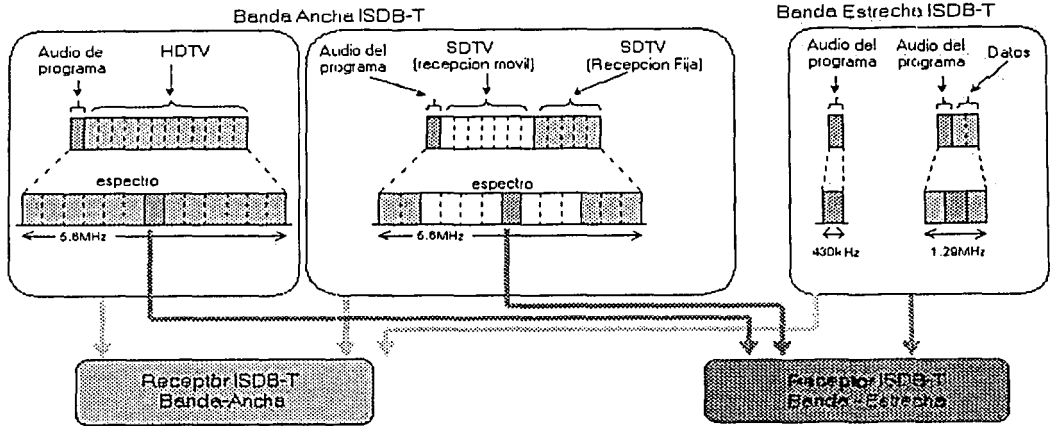


Figura 2.5. Ejemplo del esquema de transmisión y aplicación.

En el caso del estándar ISDB-T se le llama Banda-estrecha a la paquetización de audio digital sin video, en este tipo de banda solo se trasmite audio digital, por lo tanto cuando nos referimos tanto a transmisor como a receptor de banda estrecha, debemos interpretar que se trata de un dispositivo que solo es capaz de recibir o transmitir audio digital, de igual forma cuando se habla de banda-ancha sé esta considerando que existe una transmisión o recepción tanto de video digital como de audio en el canal.

Multiplexacion de Datos
(Ejemplo de la construcción de las capas jerarquicas, N=13)

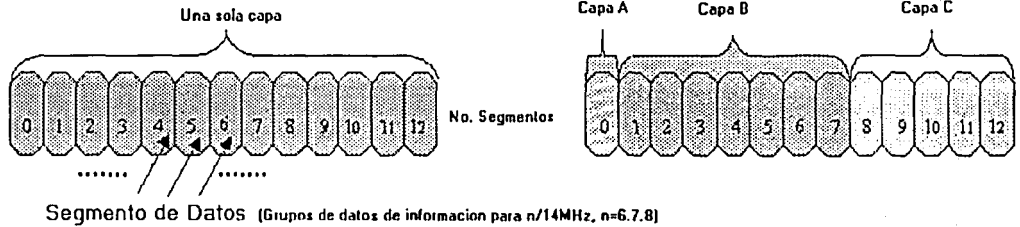


Figura 2.6. Jerarquías de transmisión y recepción parcial en ISDB-T

❖ **Concepto básico de OFDM**

OFDM^[6] (Orthogonal Frequency Division Multiplexing) es un método que divide la información para enviarla en múltiples portadoras moduladas digitalmente mediante QPSK o QAM. Figura 2.7. OFDM utiliza la transformada de Fourier rápida (FFT Fast Fourier Transform) para la demodulación. Cuenta con un intervalo extra llamado, intervalo de guardia, figura 2.8, el cual es sumado para proporcionar redundancia para el tiempo requerido por la FFT, con lo cual OFDM es menos susceptible a recepción múltiple o defasamiento, lo que da como resultado un alto desempeño en recepción móvil.

Espectro de transmisión

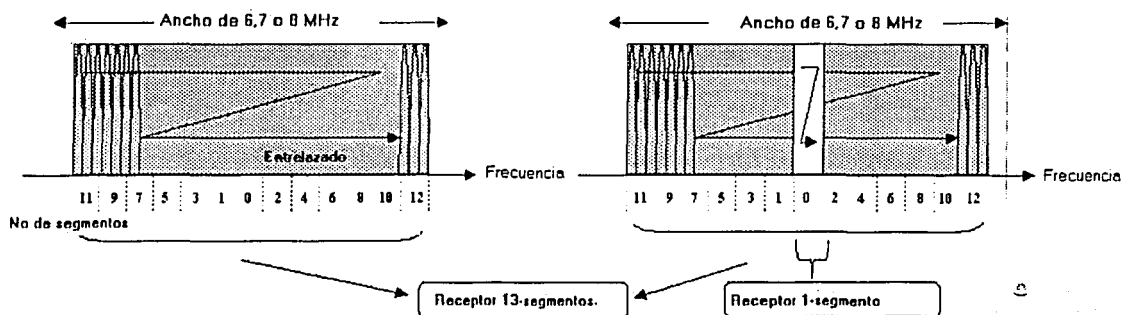


Figura 2.7 Portadora múltiples de OFDM

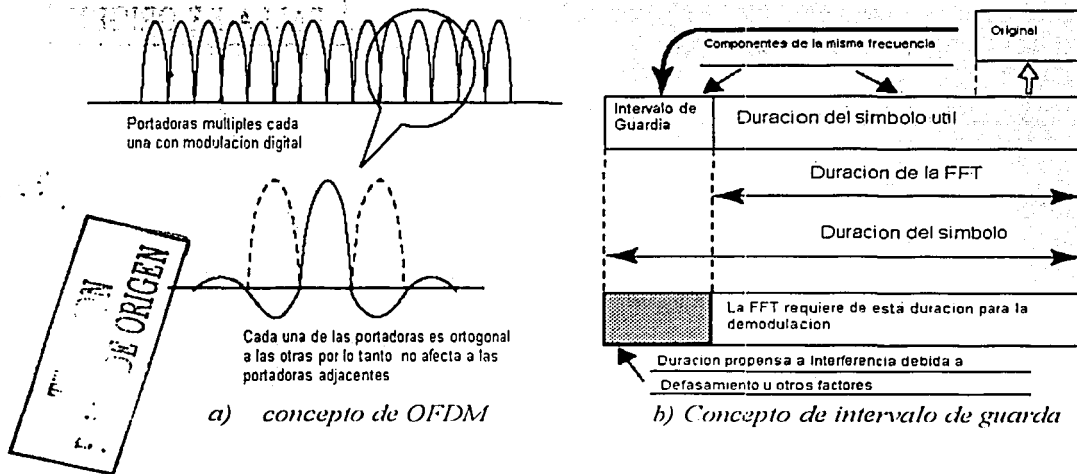


Figura 2.8 Características de OFDM

❖ Parámetros de ISDB-T

Modo ISDB-T		Modo 1	Modo 2	Modo 3
Numero de segmentos de OFDM	6,7,8 MHz	N (menor o igual a 13)		
Ancho de banda	6 MHz	5575 kHz	5573 kHz	5572 kHz
	7 MHz	6504 kHz	6502 kHz	6501 kHz
	8 MHz	7433 kHz	7431 kHz	7429 kHz
Espaciamiento de portadoras	6 MHz	3.968 kHz	1.984 kHz	0.992 kHz
	7 MHz	4.629 kHz	2.314 kHz	1.157 kHz
	8 MHz	5.291 kHz	2.645 kHz	1.322 kHz
Numero total de Portadoras	6,7,8 MHz	1405	2809	5617
Método de modulación de la portadora	6,7,8 MHz	QPSK, 16 QAM, 64 QAM, DQPSK		
Numero de símbolos por frame	6,7,8 MHz	204		
Duración efectiva del símbolo	6 MHz	256 μ s	504 μ s	1008 μ s
	7 MHz	216 μ s	432 μ s	864 μ s
	8 MHz	189 μ s	378 μ s	756 μ s

TESIS CON FALLA DE ORIGI

Duración del intervalo de guardia	6,7,8 MHz	1/4, 1/8, 1/16, 1/32 de la duración efectiva del símbolo
Código interno (Trellis)	6,7,8 MHz	Código convolucional (1/2, 2/3, ¾, 5/6, 7/8)
Código externo (Reed-Solomon)	6,7,8 MHz	RS (204, 188)
Tiempo de entrelazado	6 MHz	0, 0.096, 0.19, 0.38 seg
	7 MHz	0, 0.082, 0.16, 0.33 seg
	8 MHz	0, 0.072, 0.14, 0.29 seg
Velocidad de transmisión	6 MHz	3.651~23.234Mbit/s
	7 MHz	4.259~27.107Mbit/s
	8 MHz	4.868~30.979Mbit/s

Tabla 2.2. Parámetros de operación del estándar ISDB-T.

TESIS CON FALLA DE ORIGEN

2.2 Comparación entre estándares.

En este apartado del capítulo se realizará un análisis comparativo entre los tres estándares internacionales existentes, para la operación de la TV digital, la comparación se llevará a cabo por medio de tablas en las cuales se analizaron algunos parámetros comunes e importantes de estos estándares, la comparación se realizará al nivel de parámetros de operación y transmisión, ya que varían entre cada estándar. Este es un punto importante dado que es necesario realizar una comparación entre estándares, para poder elegir aquel que satisfaga nuestras necesidades y capacidades de la mejor manera posible, cuando se tienen múltiples opciones, con lo cual se puede realizar una elección justificada y acertada.

2.2.1 Sistemas de modulación utilizados.

El esquema de modulación utilizado por cada uno de los estándares es distinto en cada uno de ellos, con excepción de DVB-T e ISDB-T que utilizan básicamente el mismo, esta

es la más importante característica que da a cada estándar un modo de operación distinto, en el sistema de modulación utilizado radica la principal diferencia de operación entre cada estándar, ya que los tres utilizan como compresión al MPEG-2 en video, también existen diferencias en la compresión de audio, pero la importancia del esquema de modulación es que influye en la mayoría de los parámetros de operación, y debido que para el estándar japonés y el europeo se tiene el mismo esquema de modulación, la comparación se realizará solamente entre los estándares ATSC y DVB tomando en cuenta que al hacer esto implícitamente, por ser similares DVB e ISDB, también se estará comparando ATSC con ISDB. Se realizarán algunas anotaciones cuando sea requerido o se cuente con una diferencia notable en ISDB, es importante hacer notar que éste último es considerado como una mejora de DVB y que próximamente podrá sustituirlo en Europa, a continuación se presentan los sistemas de modulación utilizados por cada uno de los estándares:

- ATSC (Estándar Americano) Utiliza 8-VSB (Banda Lateral Residual de 8 niveles).
- DVB-T (Estándar Europeo) Utiliza COFDM(Codec-Orthogonal Frequency Division Multiplexing).
- ISDB-T (Estándar Japonés) Utiliza OFDM(Orthogonal Frequency Division Multiplexing).

❖ Sistema ATSC 8-VSB.

El sistema de modulación de codificación Trellis 8-VSB (8 niveles de banda lateral Residual) desarrollado en los Estados Unidos fue diseñado para transmitir video y audio digital sobre un canal de televisión de 6 MHz. Entrega una velocidad de transmisión de 19.39 Mbps en un canal terrestre. Cuenta con dos modos de operación; el modo "simulcast" 8-VSB terrestre, diseñado para tener mayor inmunidad a la interferencia NTSC y el modo de alta emisión 16-VSB desarrollado principalmente para canales de transmisión menos propensos al ruido como lo es un cable.

❖ Sistema DVB e ISDB, OFDM.

El principio del funcionamiento de la tecnología OFDM consiste en la transmisión simultanea a través de múltiples canales de una fuente de información digital. Cada uno de estos canales se modula en una frecuencia distinta, pero la velocidad de transmisión binaria es mucho menor que la velocidad de transmisión de la fuente, en principio cada canal podría ser N veces más lenta que la original, siendo N el número de canales (o frecuencias) que se transmiten simultáneamente. OFDM consigue una alta eficiencia espectral mediante técnicas de procesado de señal, que permite utilizar frecuencias entre portadoras menores de lo que tradicionalmente se venia haciendo.

En pocas palabras el principio de funcionamiento de OFDM consiste en dividir el flujo de bits de la fuente de información en un elevado numero de canales. Cada canal, de menor velocidad binaria, se modula con una frecuencia que es ortogonal al resto de las frecuencias de los otros canales. El conjunto de todas las frecuencias asignadas a cada canal cubre la banda del espectro asignada. En los sistemas OFDM los distintos usuarios se multiplexan en tiempo y/o frecuencias (TDMA y/o FDMA), de forma que la capacidad la determinan el

ancho de banda disponible, la eficiencia espectral de la modulación y la velocidad de transmisión requerida para cada usuario. La tecnología OFDM permite que cada portadora se ajuste independientemente a las condiciones de contorno (señal y ruido) en su zona de espectro, al emplear un elevado número de portadoras y poder ajustar estas independientemente, el sistema ofrece la posibilidad de adaptarse en forma casi-óptima a las condiciones de propagación de cada enlace, esta capacidad de adaptarse le confiere una capacidad superior en canales adversos.

2.2.2 Sistema de compresión utilizado.

Los tres estándares utilizan para compresión de video el estándar de MPEG-2 y en compresión de audio difieren, ATSC utiliza Dolby AC-3, DVB utiliza Musicam e ISDB utiliza MPEG-2.

- ATSC, MPEG-2 (video) y Dolby AC-3 (audio).
- DVB, MPEG-2 (video) y Musicam (audio).
- ISDB, MPEG-2 (video) y AAC (audio).

Nota: Musicam y AAC son compresiones comprendidas en el estándar MPEG

❖ MPEG-2

MPEG (Moving Picture Experts Group) es un estándar de compresión de audio, video y datos establecido por la Unión Internacional de Telecomunicaciones. El estándar MPEG 2 está normalizada por: ISO 13818-1¹⁷¹ (MPEG2 System), 13818-2 (MPEG2 Video), 13818-3 (MPEG2 Audio). Lo importante de la norma es que no exige como construir el codificador o decodificador. Esto queda a criterio del fabricante, exige sólo cumplir los parámetros del TS (Transport Stream) fijados por la norma:

El estándar prevé la transmisión de distintos niveles de calidad y bit rates dado por la elección de los distintos niveles y perfiles para el tipo de servicio que se desea brindar. Ver tabla 2.3.

Normas Internacionales para TV Digital

Arquitectura Mpeg		Perfiles				
Niveles	SIMPLE	MAIN	SNR	SPATIAL	TRGM	4:2:2
	I & P	I, P & B	I, P & B	I, P & B	I, P & B	I, P & B
Muestreo :	4:2:0	4:2:0	4:2:0	4:2:0	4:2:2 & 4:2:0	4:2:2 & 4:2:0
<i>HIGH LEVEL</i>		<i>HDTV</i>				
Samples/Line		1920			1920	1920
Lines/Frames		1152			1152	1080
Frames/Sec		60			60	60
Max Bit-Rate (Mbps)		80			100	300
<i>HIGH 1440</i>						
Samples/Line		1440		1440	1440	
Lines/Frames		1152		1152	1152	
Frames/Sec		60		60	60	
Max Bit-Rate (Mbps)		60		60	60	
<i>MAIN</i>		<i>SDTV</i>				
Samples/Line	720	720	720	720	720	720
Lines/Frames	576	576	576	576	576	608
Frames/Sec	30	30	30	30	30	30
Max Bit-Rate (Mbps)	15	15	15	15	20	50
<i>LOW</i>						
Samples/Line		352				
Lines/Frames		288				
Frames/Sec		30				
Max Bit-Rate (Mbps)		4				

TESIS CON FALLA DE ORIGEN

Tabla 2.3 Perfiles y niveles MPEG-2.

❖ **Dolby AC-3.**

Dolby Digital AC-3 es un sistema denominado de 5,1 canales debido a que proporciona 5 canales independientes (izquierdo, derecho, central, surround izquierdo y surround derecho); todos ellos reproducen una gama de 20 a 20,000 Hz. Además, el sistema puede proporcionar un canal subwoofer opcional independiente. A pesar de que los cinco canales proporcionan un ancho total de banda que abarca todo el espectro audible, se añade un canal para los efectos sonoros de Baja Frecuencia. Además, el sistema Dolby Digital AC-3 mejora la separación entre canales y la capacidad de que sonidos individualizados lleguen desde múltiples direcciones al oyente.

El sistema Digital Dolby AC-3 varía la velocidad de transmisión de forma instantánea en función de la frecuencia, para controlar la resolución general de la banda sonora. Con la utilización de algoritmos complejos, calcula la distribución óptima de los bits sin ninguna degradación audible del sonido. La compresión digital produce de forma inherente más ruido que la transmisión de datos sin compresión. Cuanto menos son los bits utilizados para describir una señal de audio, mayor es el ruido. De este modo, el sistema Dolby Digital AC-3 compensa esta reducción de información de dos modos: por filtrado digital y por enmascaramiento.

El filtrado digital tiene lugar en las primeras etapas del proceso, la información es fragmentada en 50 bandas de frecuencia. Pasando la información de audio a través de una serie de filtros paso banda para cada uno de los cinco canales, se suprime la mayor parte del ruido al limitar la respuesta de frecuencia. En la etapa de salida, el enmascaramiento se produce en la mayoría de los sistemas analógicos de reducción de ruido desarrollados por Dolby, y tienen una amplia aceptación en la industria electrónica. Los estándares Dolby Digital AC-3 tienen 50 bandas independientes con filtros paso banda (agudos y graves) de 12 dB por octava. Estos filtros no son lo suficientemente pronunciados para reducir totalmente el ruido residual que se crea con una velocidad de bits tan baja; por ello; Dolby utiliza, conjuntamente, técnicas de enmascaramiento para reducir el ruido audible.

❖ **Musicam.**

MUSICAM permite la reducción de la tasa de bits de una señal de audio digitalizada hasta 48 [kHz]. Es un proceso de codificación de audio altamente eficiente que puede tomar la señal de CD stereo a $2 \times 44.1 \times 16 = 1411.2 \text{ kbit/s}$ y reducirla a una tasa de 192kBit/s. Lo anterior es posible basándose en técnicas de codificación psico-acústicas especificadas por MPEG-2 layer II:

- La señal de audio (24 kHz en banda base) digitalizada es dividida en 32 sub-bandas. Lo anterior se realiza con DSP's que realizan FFT de 1024 bits, entregando 512 valores del espectro, es decir, muestras cada 46 Hz del espectro de audio original.
- Conocidas las componentes de frecuencia. Se divide el espectro en las 32 sub-bandas, cada una con 16 bits (512/32). Esta división del espectro permite la

distribución óptima de los bits de acuerdo a los requerimientos psicoacústicos. Para aquellas sub-bandas que resultan completamente enmascaradas por otras, no hay necesidad de enviarlas, ya que no se escucharán.

- Una vez que se cuenta con la información necesaria, a cada muestra se le asigna un factor de escala de 6 bits (asegurando un rango dinámico de 120 dB), junto con información para reconstruir la distribución óptima y un header para cierta información. De esta forma, se arma la trama.
- El sistema contempla técnicas de protección de la información, ya que la destrucción por pequeña que sea, resulta desastrosa. Además de la eliminación de redundancias para los factores de escala.

2.2.3 Comparación de parámetros de transmisión.

A continuación se presenta en la tabla 2.4 la comparación de algunas características más sobresalientes de los estándares y en la tabla 2.5 se muestran algunos parámetros de operación, como ya habíamos dicho anteriormente solo se realizara la comparación entre ATSC y DVB por la similitud ya comentada entre DVB e ISDB

CARACTERISTICAS	ATSC	ISDB	DVB
Ancho de banda por canal	6 MHz	6,7 y 8 MHz	8 MHz
Tipo de modulación	8-VSB o 16 VSB	OFDM (QPSK,16QAM,64QAM)	COFDM (con mapeo QPSK o QAM)
Estándar de compresión	MPEG-2(video) Dolby AC-3(audio)	MPEG-2 (video) MPEG-4 (audio)	MPEG-2 (video) Musicam (audio)
Relación de aspecto	4:3 (SDTV) 16:9 (HDTV)	4:3 (SDTV) 16:9 (HDTV)	4:3 (SDTV) 16:9 (HDTV)
Máxima resolución	Vertical 1080 líneas Horizontal 1920 pixeles	Vertical 1152 líneas Horizontal 1920 pixeles	Vertical 1152 líneas Horizontal 1920 pixeles

Tabla 2.4 Características de los estándares.

Capítulo # 2

Parámetro de comparación	Descripción	ATSC	DVB-T
Número de portadoras		una	Grande
Relación de Potencia Pico a Potencia Promedio (PAR)	Potencia pico nivel de potencia particular que se encuentra debajo o encima de la potencia promedio en cierto porcentaje de tiempo. PAR proporciona idea de cuanta potencia extra. Amplificador alta potencia. Costos.	✓	Más potencia 60%
Bit Error Rate (BER)	Medición del número de bits que se han recibido de manera errónea	3×10^{-6} ✓	2×10^{-11}
Umbral C/N	Relación mínima entre potencia portadora y potencia ruido para obtener recepción adecuada de la información. Si es chica, la potencia de Tx es menor.	✓ 15.2 dB	19.2 dB
Relación E_b/N_0	Desempeño del sistema según la tasa de transmisión empleada	✓	Más potencia
Distorsión por multitrayectoria	Sensibilidad del sistema para poder discriminar las distintas señales de la misma fuente de Tx, que pertenecen a reflexiones. En el caso de la Cd. de México no afecta de manera evidente (analógica)	Retrasos pequeños y de nivel grande	Retraso es grande ✓
Movilidad			✓
Interferencia por tono	Adición de una señal de frecuencia única. En la etapa de Rx, elementos transitorios debidos al televisor o señales de aparatos electrodomésticos	Pérdidas en señal, puede o no afectar inf. clave	Afecta una o algunas de las portadoras, afecta su inf. ✓
Ruido de Fase	Rx heterodinos emplean osciladores. Degradaciones del oscilador local son variaciones de fase conocidas como ruido de fase. Mejorar el oscilador.	Provoca rotación en constelación Corregido con PLL	Sintonizador con un comportamiento mejorado de ruido de fase
Interferencia de co-canal	Dos o más canales de TV usan la misma frecuencia con una cobertura diferente. TV UNAM es canal 60 de UHF.	✓	
Sistema de Corrección de errores	Más inmune a los impulsos de ruido y a distorsión por ruido de fase	✓	
Ruido de impulso	Hay ráfagas de ruido en el sistema, ambiente urbano, equipo industrial, aplicaciones caseras, líneas de alto voltaje. Ocurre en la banda de VHF y en la parte baja de UHF.	Sistema de reordenamiento de 52 segmentos. ✓	Sistema de reordenamiento de 12 seg.
Eficiencia en el uso de espectro	Porcentaje del ancho de banda utilizado efectivamente para la transmisión de información	✓	Intervalos de guardia
Zona de cobertura	Ambos sistemas pueden proporcionar casi la misma cobertura		Terreno irregular ✓
Potencia de Transmisión	ATSC requiere una menor potencia de transmisión para conseguir una misma área de cobertura que el DVB-T	✓	
Situación geográfica, edificios, antena receptora dentro de casa	Estos fenómenos delimitan la capacidad del sistema		✓
Migración de TV analógica a digital	Hay que ver inconvenientes técnicos y económicos. La norma oficial de televisión existente en México está basada en NTSC.	✓	
Expectativas de servicio	TESIS CON FALLA DE ORIGEN	HDTV	Integración de servicios de telecomunicaciones en un solo dispositivo

Tabla 2.5 Parámetros de operación.

2.3 ¿Qué estándar adoptara México?

En México han tenido lugar varias conferencias de empresas dedicadas a la producción de accesorios para la operación de canales digitales como son: transmisores, antenas, decodificadores, equipos de prueba y medición, etc.,. Estas conferencias apuntan a la utilización del estándar ATSC. A estas demostraciones y conferencias han asistido ingenieros de empresas mexicanas interesadas y obligadas a realizar la transición como son las televisoras privadas sobre todo TV Azteca y Televisa, además de ingenieros, investigadores y estudiantes involucrados en el proyecto de TV UNAM, así como gente del gobierno, principalmente de la secretaria de comunicaciones y transportes que es la encargada de la regulación y normatividad.

Nosotros asistimos a un par de estas conferencias: una ofrecida por ROHDE & SCHWARZ, empresa alemana dedicada a fabricar transmisores e instrumentos de medición; y la otra por Harris, empresa de Estados Unidos dedicada a la fabricación de transmisores y que ofrece apoyo técnico de instalación y mantenimiento. Al asistir a estas conferencias pudimos observar que la elección de México apunta hacia el estándar americano ATSC. Las televisoras ya han realizado pruebas de transmisión en ATSC, y TV UNAM ya esta en proceso de adquirir el transmisor. De acuerdo a lo que se ha visto en TV UNAM se considera que México adoptara dicho estándar. Ya que todo el desarrollo e investigación que se ha llevado a cabo es sobre este estándar.

2.3.1 ¿Por qué ATSC?

Son muchas las razones para creer que México adoptará este estándar. Entre las principales destacan: la gran influencia tanto política, económica y social que ejerce Estados Unidos sobre nuestro país, además de la tecnológica que en este caso es la más importante ya que se requiere el apoyo tecnológico y respaldo técnico para iniciar esta importante transición y esto se puede obtener con mayor facilidad de los Estados Unidos debido a la cercanía y el gran mercado de importación tecnológica que se tiene con ellos.

También hay que mencionar que este estándar tiene mas ventajas sobre DVB que desventajas comparado con éste. De las tablas anteriores podemos resumir que ATSC utiliza menor potencia de transmisión, lo cual es una ventaja económica de ahorro de energía; ATSC es una norma hecha para poder transmitir la verdadera HDTV y DVB no, con esto se quiere decir que ATSC permite la TV interactiva, en la cual el usuario tendrá muchos servicios y se podrán realizar varias aplicaciones como las que se tienen en internet (compras, votar, interacción en tiempo real con los programas, etc.) que es algo que en DVB no se contempla ya que fue diseñado para proporcionar una televisión con mejor imagen y audio, razón por la cual se piensa en ISDB como una mejora para DVB ya que tienen los parámetros de operación muy similares con la diferencia de que permite lo mismo que el estándar ATSC y además permite recepción móvil; la desventaja de ISDB es que se ha retardado mucho su lanzamiento y no ha sido muy promovido, en comparación con ATSC, que le lleva ventaja por que ya se encuentra en operación en Estados Unidos desde hace algunos años. Por esto en los capítulos siguientes nos enfocaremos al análisis mas profundo de la operación de este estándar, como son su modulación, métodos de

compresión, envío de la señal digital, aspectos técnicos, etc., para concluir con la propuesta de un diseño para receptor de TV digital en ATSC.

2.4 Descripción general de ATSC.

Ahora se enunciarán algunas de las características generales que conforman al estándar ATSC, empezando por la serie de normas que rigen su operación que son las siguientes:

- A/52 Audio digital
- A/53 Norma DTV
- A/54 Guía para el uso de la norma DTV
- A/63 Norma para la codificación de vídeo a 25/50Hz
- A/64 Transmisión y cumplimiento
- A/65 Protocolo de información de sistemas y programas (PSIP)
- A/70 Acceso condicional
- A/80 Satélite (contribución y distribución)
- A/90 Emisión de datos
- A/92 Multidifusión IP
- A/93 Activador síncrono/asíncrono

Las normas están disponibles en www.atsc.org.

La tesis no se enfocará al estudio de estas normas y su reglamentación, pero cabe mencionar que la mayoría de los parámetros requeridos en estas se mencionan durante el desarrollo de esta, en este trabajo realizaremos un análisis sobre su operación. Las capas de operación de ATSC se muestran en la siguiente figura 2.9.

En los siguientes capítulos se describirán los procesos que se realizan en cada una de las capas con las cuales transmite este estándar, con el propósito de conocer como se recibirá la información y los procesos que ha sufrido esta para poder realizar los procesos inversos y así decodificar la señal original. Como sabemos, para decodificar una señal básicamente hay que realizar el proceso inverso al que se realizó para su transmisión y agregar algunos sistemas para corrección y detección de errores.

Normas Internacionales para TV Digital

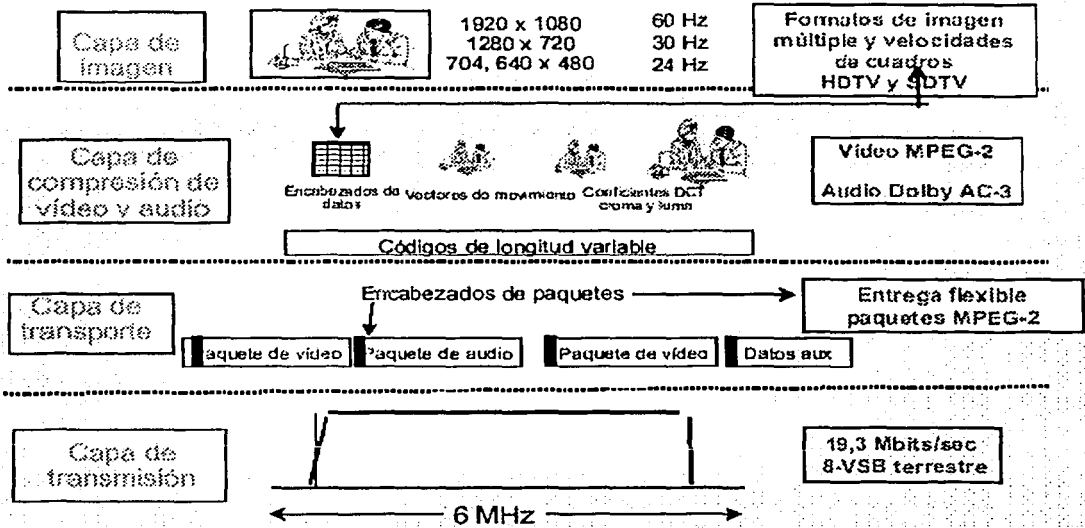


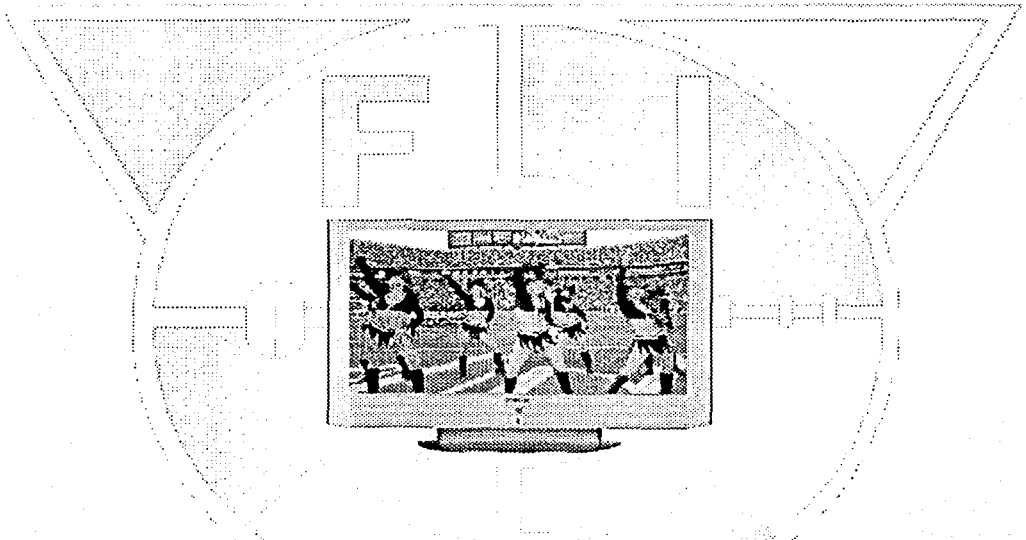
Figura 2.9. Capas de procesamiento de la señal para su transmisión.

TESIS CON FALLA DE ORIGEN

REFERENCIAS

- [1] Comité de Servicios de Televisión Avanzada (ATSC), organización internacional no lucrativa que desarrolla estándares voluntarios para la TV Digital. Sitio web www.atsc.org
- [2] Transmisión Digital de Video (DVB), consorcio de difusores, fabricantes, operadores, desarrolladores y organismos reguladores que diseñan los estándares globales para los servicios de televisión digital. Sitio web www.dvb.org
- [3] Grupo de Difusores Digitales Expertos (DiBEG), fundado para promover el estándar japonés de TV Digital, ISDB-T. Sitio web www.isdb.org
- [4] Índice de recursos MPEG en Internet, se enfoca en los estándares MPEG-1 y MPEG-2. Sitio web www.mpeg.org
- [5] Sociedad para la Industria y Matemáticas Aplicadas. Sitio web <http://epubs.siam.org>
- [6] Manual de Telecomunicaciones. Sitio web www.rares.com.ar
- [7] Plataforma para el desarrollo de software libre para la televisión digital. Sitio web www.linuxtv.org

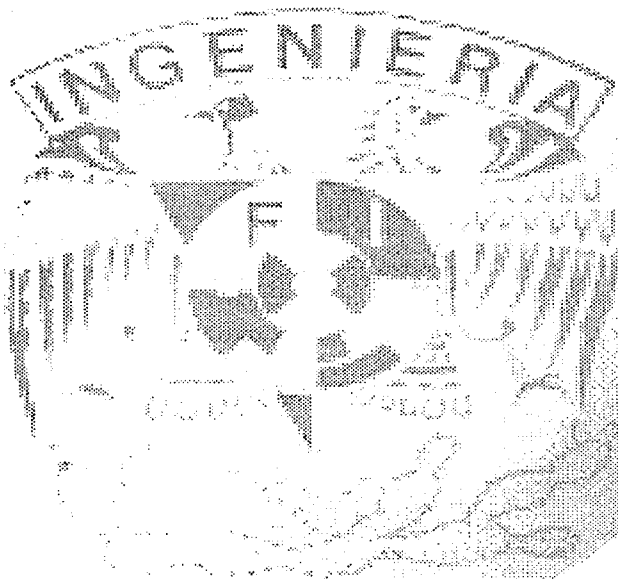




Capítulo # 3

*Compresión
de
Video*

TESIS CON
LA DE ORIGEN



TESIS CON
FALLA DE ORIGEN

44

3

Con base en el estándar ATSC se estudiara a continuación el formato MPEG aplicado a las imágenes de video, debido a que se trata de la parte medular de dicha norma.

3.1 Compresión de Datos

Las tecnologías digitales han tenido un rápido progreso en el audio y el video debido a muchas razones. La información digital es mucho más robusta y puede codificarse para eliminar los errores sustancialmente. Esto quiere decir que las pérdidas originadas durante las grabaciones y las transmisiones son depuradas. El Disco Compacto fue el primer producto de consumo que demostró este hecho.

Es cierto que el Disco Compacto ha mejorado la calidad del sonido con respecto a su predecesor vinílico, sin embargo, esta no es la mejora más importante. La ventaja más importante de las tecnologías de transmisión y grabación digitales es que permiten la manipulación de su contenido, a un grado que sería imposible con las técnicas analógicas. Una vez que el video o el audio son digitalizados se convierten en datos, con lo que llegan a entrar dentro del campo de la tecnología de las computadoras.

La convergencia de las computadoras y el audio/video es una consecuencia inevitable de las importantes invenciones del cómputo y la Modulación por Codificación de Pulso (PCM – Pulse Code Modulation). Una aplicación de esta tecnología de convergencia es multimedia, en donde se mezcla el almacenamiento de audio, video, gráficos, texto y datos en el mismo medio. Multimedia es imposible en el dominio analógico.

El éxito inicial del video digital ocurrió en las aplicaciones de post-producción, en donde el alto costo del video digital era compensado por su capacidad ilimitada. Sin embargo, producir una señal de video digital estándar genera unos 200 megabits por segundo de datos, y esta tasa de transmisión requiere una gran capacidad de almacenamiento y un gran ancho de banda para ser transmitida. El video digital solo podría ser utilizado en un mayor número de aplicaciones si los requerimientos de almacenamiento y el ancho de banda son disminuidos. Reducir estos requerimientos es el objetivo de la compresión.

La compresión es la manera de expresar el audio y el video digital usando menos datos. La compresión tiene las ventajas siguientes:

- Se necesita un espacio de almacenamiento más pequeño para una cantidad de material fuente dado.
- Cuando se trabaja en tiempo real, la compresión reduce el ancho de banda necesario. Además, la compresión permite una transferencia más rápida que la de tiempo real entre medios, por ejemplo, entre una cinta y un disco.

- Un formato de grabación comprimido puede proporcionar una densidad de grabado más baja y esto puede hacer que la grabadora sea menos sensible a los factores ambientales y el mantenimiento.

3.2 Introducción a la Compresión de Video

La compresión de video surge de la necesidad de transmitir imágenes a través de un canal que contenga un ancho de banda aceptable. Al digitalizar una señal analógica se tendrá un incremento en el ancho de banda de la señal resultante, ya que la frecuencia de muestreo debe ser por lo menos dos veces mayor a la frecuencia máxima de la señal analógica ($f_s \geq 2f_a$).^[1] Además si la señal transportara mas información, implicaría un aumento en el ancho de banda de la señal. El objetivo es aprovechar al máximo el espectro radioeléctrico otorgado para las transmisiones de televisión, con lo que resulta obvio que no es posible transmitir la señal digital de televisión en su formato original. Surge entonces, la necesidad de utilizar un proceso de compresión.

En todo material de un programa real, existen dos componentes de la señal: aquellos que son nuevos e impredecibles y aquellos que pueden ser previstos. Los componentes nuevos se llaman entropía y es la información verdadera de la señal. Los demás se llaman redundancia porque no son esenciales. Todas las imágenes de televisión contienen gran cantidad de información redundante, que implica una mayor cantidad de información.

Encontramos dos tipos básicos de redundancia: *espacial*, como en las superficies planas de una fotografía donde los pixeles adyacentes tienen casi el mismo valor; y *temporal* que son las similitudes que existen entre las imágenes sucesivas.

Redundancia Espacial: Una imagen puede tener grandes áreas de un mismo color, áreas de cielo o de hierba. En la siguiente imagen también se ve la hierba y el cielo. En ambas imágenes la información se repite, y no es necesario enviar toda la información, por lo tanto se puede reducir la cantidad de bits. Usando el código del primer píxel azul que describe el cielo, es posible hacer que se repita tantas veces como se necesite durante la línea de la imagen, y así solo se envía un píxel y un código con las veces que se repite. Para sacarle provecho se usa la correlación entre los elementos contiguos de una imagen. Se utiliza aquí la llamada compresión "intraframe" o "intra-coding".

"Intra-coding" es la técnica que explota la redundancia espacial. Esta técnica se basa en dos características que poseen las imágenes: no todas las frecuencias están presentes simultáneamente, y mientras más grande es la frecuencia espacial, menor es la amplitud. Para realizar esta técnica se requiere un análisis de las frecuencias espaciales de una imagen, el cual se puede llevar a cabo con las transformadas, como la DCT. Las transformadas producen coeficientes que describen la magnitud de cada frecuencia espacial. Típicamente muchos coeficientes tendrán un valor de cero, o cercano a cero, por lo que estos coeficientes pueden ser omitidos, dando como resultado una reducción en la tasa de transmisión.

Redundancia Temporal: No es necesario transmitir imágenes un total de 25 veces o 30 por segundo, solo las diferencias entre imágenes sucesivas son enviadas. Al hacer esto, se utiliza la compresión “interframe” o “inter-coding”.

“Inter-coding” es la técnica que explota la redundancia temporal y se refiere a encontrar similitudes entre imágenes sucesivas. Si se tiene una imagen disponible en el decodificador, la próxima imagen puede ser creada mandando solo las diferencias. Las diferencias pueden verse incrementadas cuando se mueven objetos, pero este aumento puede ser contrarrestado usando la compensación de movimiento, debido a que generalmente un objeto en movimiento no cambia mucho su apariencia de una imagen a la otra. Si el movimiento puede ser cuantificado, se puede crear una aproximación más cercana de la imagen mediante desplazar parte de la imagen previa a una localización nueva. El proceso de desplazamiento es controlado por un vector que se transmite hacia el decodificador. La transmisión de este vector requiere menos datos que enviar los datos de diferencia de imágenes. La compensación de movimiento minimiza pero no elimina las diferencias entre imágenes sucesivas.

La compresión de video puede tomar ventaja de la redundancia espacial y la temporal. En MPEG, la redundancia temporal se reduce primero, usando las similitudes entre imágenes sucesivas. Mientras sea posible la imagen actual es creada o “predecida” usando la información de las señales que fueron enviadas antes. Cuando se usa esta técnica, solo es necesario enviar la imagen diferencia, que elimina las divergencias entre la imagen real y la predicha. La imagen diferencia se sujeta a una compresión espacial. La compresión espacial aprovecha las similitudes entre píxeles adyacentes en áreas planas de una imagen y de las frecuencias espaciales dominantes en las áreas de patrones.

La codificación MPEG se divide en varios perfiles que tienen diferentes complejidades, y cada perfil puede ser implementado en niveles distintos dependiendo de la resolución de la imagen de entrada.

3.3 El sistema visual humano

En algunas aplicaciones del procesamiento digital de imágenes, como en la compresión digital, es necesario entender las características y las limitaciones que tiene el sistema ojo-cerebro, para ayudar a maximizar la efectividad de las operaciones de la compresión digital de imágenes.^[2]

Respuesta a la iluminación y discriminación. Debido a que las imágenes digitales se presentan como un conjunto de puntos brillantes, la capacidad del ojo de discriminar entre diferentes niveles de iluminación es una consideración importante para presentar los resultados del procesamiento de la imagen. La relación entre la intensidad de la luz que entra al ojo y su brillo percibido no es una función lineal. Esto significa que a medida que la intensidad de una fuente luminosa cambia, el observador no percibirá un cambio igual en el brillo. La respuesta de la intensidad real del ojo es más logarítmica, similar a la curva de la figura 3.1. De hecho, se ha mostrado experimentalmente que la intensidad de una fuente luminosa debe ser cercana al doble antes de que el ojo pueda detectar que esta ha cambiado.

Por lo tanto, los cambios ligeros en la intensidad en regiones oscuras de una imagen tienden a ser más perceptibles que los cambios iguales en regiones brillantes. Esta relación que hay entre la intensidad de la iluminación y el brillo percibido, es conocida como Ley de Weber.

El punto importante es que la respuesta logarítmica del ojo a la intensidad hace a este más sensitivo a los cambios de intensidad en las regiones oscuras que en las regiones brillantes de la imagen. En el procesamiento digital de una imagen, un simple oscurecimiento de las regiones brillantes pueden hacer indetectables los cambios minúsculos de intensidad. La capacidad del ojo para discriminar entre cambios de iluminación para cada nivel específico de adaptación también es de considerable interés.

TESIS CON
FALLA DE ORIGEN

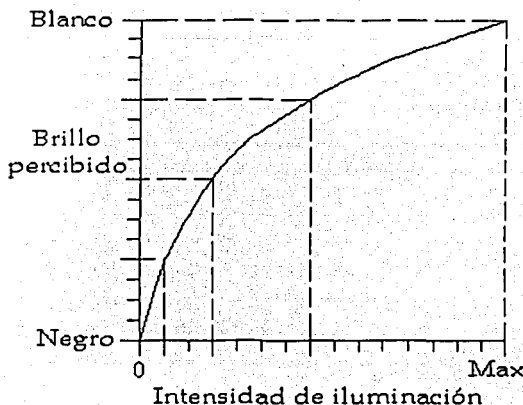
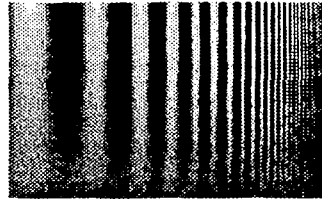


Figura 3.1 Respuesta logarítmica del ojo, Ley de Weber.

El sistema visual tiene limitaciones fundamentales en la respuesta en frecuencia. Como en cualquier sistema óptico, el ojo tiene límites de cómo puede resolver detalles finos, o transiciones de intensidad. Los factores limitantes son el número y organización de los fotorreceptores en la retina, la calidad de la parte óptica del ojo (córnea, humor acuoso, cristalino, y humor vítreo), y la transmisión y procesamiento de la información visual al cerebro. Generalmente, la respuesta en frecuencia del ojo disminuye a medida que se ven transiciones de intensidad, que se vuelven cada vez más finas, como se ve en la figura 3.2. Es también un factor el contraste, o diferencia entre niveles de gris, de la transición de intensidad. Entre más alto el contraste, más fino es el detalle que el ojo puede resolver. Finalmente, cuando las transiciones están demasiado finas o el contraste es demasiado bajo el ojo ya no puede resolverlos. En este punto, el ojo puede percibir sólo un promedio del nivel de gris del área detallada.



TESIS CON
FALLA DE ORIGEN

Figura 3.2 Patrón que incrementa la frecuencia de izquierda a derecha y decrementa el contraste de arriba a abajo.

El fenómeno discutido ilustra el complejo proceso que ocurre en el sistema visual humano. Combinando los conceptos de respuesta de intensidad no lineal, interacción del fotorreceptor y respuesta en frecuencia del ojo, se pueden hacer las observaciones siguientes:

- La intensidad del objeto visto es relacionada a la intensidad promedio alrededor del objeto. El objeto aparece más oscuro si el área circundante es brillante, o más brillante si el área circundante es oscura.
- Los cambios sutiles de intensidad son mas aparentes en las regiones oscuras que en las regiones brillantes de la imagen.

Las transiciones marcadas de intensidad se acentúan en una imagen. La respuesta a los detalles de la imagen decae cuando los detalles a resolver son demasiado finos. Los detalles con contraste alto se pueden resolver mas fácilmente que aquellos con contraste bajo.

Mezcla aditiva de colores. Los colores se pueden obtener haciendo una mezcla de los tres colores primarios, rojo, verde y azul, esta mezcla se denomina aditiva. Un ejemplo de la mezcla aditiva de estos tres colores se muestra en la tabla 3.1.

Rojo	+	Verde	=	Amarillo
Rojo	+	Azul	=	Magenta
Azul	+	Verde	=	Cián
Rojo + Azul	+	Verde	=	Blanco

Tabla 3.1 Mezcla aditiva de colores.

3.4 Fundamentos del vídeo digital

3.4.1 Vídeo analógico

En un sistema analógico, la información se transmite mediante alguna variación infinita de un parámetro continuo como puede ser el voltaje o la corriente (véase la figura 3.3).

El número de líneas de exploración horizontal de un campo es la mitad del total de las 525 líneas de un cuadro completo en el sistema NTSC o National Television System Committee [3], ya que un campo contiene la mitad de las líneas. Esto da por resultado 262.5 líneas horizontales para cada campo.

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

$$262.5 \times 60 = 15,750 \text{ líneas/s}$$

La frecuencia de 15,750 Hz es la velocidad con que el haz electrónico completa su ciclo de movimiento horizontal de izquierda a derecha y regresa nuevamente a la izquierda.

TESIS CON
FALLA DE ORIGEN

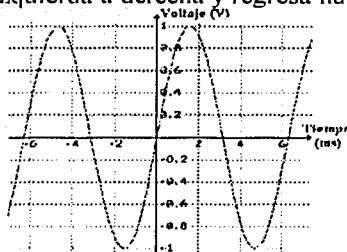


Figura 3.3 Variación infinita de un parámetro continuo en función del tiempo.

Las señales de color. [1.2.3] El sistema para la televisión en color es el mismo que para la televisión monocromática (conocida comúnmente como TV Blanco y Negro) excepto que en este también se utiliza la información de color. Esto se realiza considerando la información de imágenes en términos de rojo, verde y azul. Cuando es explorada la imagen en la cámara, se producen señales de vídeo separadas para la información de rojo, verde y azul de la imagen. Filtros de color separan los colores para la cámara. Sin embargo, para el canal estándar de 6 MHz de televisión, las señales de vídeo de rojo, verde y azul son combinadas de modo que se forman dos señales equivalentes, una correspondiente al brillo y otra para el color. Específicamente las dos señales transmitidas son las siguientes:

Señal de luminancia: Contiene solo variaciones de brillo de la información de la imagen, incluyendo los detalles finos, lo mismo que en una señal monocromática. La señal de luminancia se utiliza para reproducir la imagen en blanco y negro, o monocroma. La señal de luminancia o Y se forma combinando 30% de la señal de vídeo roja (R), 59% de la señal de vídeo verde (G) y 11% de la señal de vídeo azul (B). Estos porcentajes corresponden a la brillantez relativa de los tres colores primarios.

Señal de crominancia: La señal de crominancia o C es una combinación de las señales de color I y Q. La señal I o señal de color en fase se genera combinando el 60% de la señal de vídeo en rojo (R), 28% de la señal de vídeo en verde (G) invertida y 32% de la señal de vídeo en azul (B) invertida. La señal Q o señal de color en cuadratura se genera combinando el 21% de la señal de vídeo en rojo (R), 52% de la señal de vídeo en verde (G) invertido y 31% de la señal de vídeo en azul (B). Las amplitudes de las señales I y Q son

proporcionales a las señales de vídeo R, G y B. La señal de crominancia se define como $C = (I^2 + Q^2)^{1/2}$.

Las señales R-Y (señal diferencia de color rojo llamada Cr) y B-Y (señal diferencia de color azul conocida como Cb) se utilizan en la mayor parte de los receptores de televisión a color para demodular las señales de vídeo R, G y B. [11]

En el receptor NTSC, la señal C reproduce colores en proporción a las amplitudes de las señales I y Q. Con estas dos señales se consigue que los sistemas de color y monocromáticos sean completamente compatibles

3.4.2 Vídeo digital

La digitalización de una señal de vídeo tiene lugar en tres pasos: muestreo, cuantización y codificación. [1,2,4,5]

❖ Muestreo

A partir de una señal analógica, se toman muestras breves de esta cada determinado periodo de tiempo. El propósito del muestreo es obtener valores de la amplitud de la señal de entrada analógica de forma periódica, continuamente cambiante, y convertir las muestras en una serie de niveles PAM [11], es decir que la amplitud de un pulso varía de acuerdo a la amplitud de la señal analógica.

Ahora bien, una señal de vídeo está compuesta por un gran número de frecuencias formando un espectro continuo que va desde 0 a unos 6 MHz. Al muestrear esta señal, si vemos el resultado en el dominio de la frecuencia, cada frecuencia de vídeo aparecerá en las bandas laterales superiores e inferiores de cada armónico de la frecuencia de muestreo, incluyendo naturalmente la banda base (intervalo de frecuencia en el cual se encuentra contenida la información que deseamos transmitir). Esto se debe a que esencialmente, un circuito de muestreo y retención es un modulador de AM.

❖ Cuantización

Así se denomina al proceso mediante el cual se atribuye a cada muestra un valor de amplitud dentro de un margen de niveles previamente fijado. Este valor se representa por un número que será convertido a un código de ceros y unos en el proceso de codificación. Por razones de facilidad en los cálculos, el número de niveles se hace coincidir con una potencia de dos y los impulsos de la señal PAM se redondean al valor superior o inferior según sobrepasen o no la mitad del ancho del nivel en que se encuentran.

El error que se produjo con estas aproximaciones se llama error de cuantización [5], y equivale a sumar una señal errónea a los valores exactos de las muestras, como se ve en la figura 3.4.

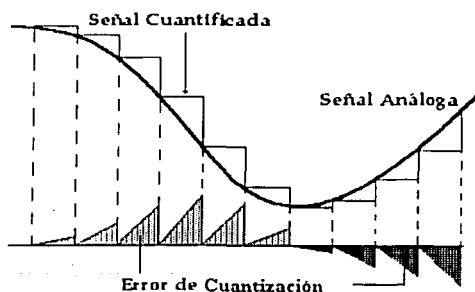


Figura 3.4 Error de cuantización.

Esta señal errónea aparecerá en el proceso de recuperación después de la decodificación digital-análoga, en forma de ruido visible. Cuantos más niveles existan menor será el ruido generado.

❖ Codificación

La codificación se define como el uso de un sistema de símbolos para representar información. El proceso de codificación representa las muestras obtenidas en la cuantización con algún código. Normalmente se usa un código binario (sus símbolos son el 1 y 0) ^[5].

Formatos de codificación: Dos planteamientos aparentemente contradictorios se mantienen aún hoy día acerca de la digitalización de la señal de televisión en color:

Codificación de las señales compuestas. Esta propuesta consiste en digitalizar directamente las señales compuestas existentes (NTSC, PAL, SECAM) ^[2,3]. Con ello persiste el problema de la incompatibilidad de las distintas normas internacionales, aun manteniendo la misma frecuencia de muestreo y codificación. La decodificación devolvería las señales NTSC, PAL o SECAM, respectivamente.

La ventaja fundamental de digitalizar la señal compuesta radica en que el equipo puede incluirse como una unidad mas en los Estudios análogos actualmente en servicio, sin necesidad de codificar o decodificar el NTSC, PAL o SECAM.

Codificación en componentes. Por este método se digitalizan las tres señales Y, $K_1(R-Y)$, $K_2(B-Y)$ donde K_1 y K_2 son factores de ponderación que imponen el sistema digital. Estos factores no tienen los mismos valores que los coeficientes ponderados de NTSC, PAL o SECAM. La primera y gran ventaja que se deriva de esta codificación es que siendo estas tres señales comunes a todos los sistemas, la compatibilidad puede alcanzarse por regulación internacional de los parámetros de muestreo, cuantificación y codificación. En tal sentido el CCIR (Comité Consultatif International des Radiocommunications o Comité Consultivo Internacional de Radio Comunicaciones) emitió en 1982 la norma 4:2:2 CCIR 601 de televisión digital en componentes.

La segunda ventaja de esta codificación es que una vez alcanzada la digitalización plena de la producción, sólo se requiere un paso final de conversión D/A y una codificación NTSC, PAL o SECAM según el sistema adoptado de transmisión. Se añade a las ventajas ya señaladas que el tratamiento digital en componentes elimina los efectos perturbadores mutuos de luminancia y crominancia a la vez que en edición electrónica desaparecen los problemas derivados de la estructura de 4 y 8 campos NTSC y PAL respectivamente. Sólo habría de tenerse en cuenta la estructura de dos campos entrelazados como en televisión en blanco y negro.

❖ La norma de televisión digital 4:2:2

Como ya se dijo antes la señal de video tiene tres componentes: Y, Cb y Cr. Durante el proceso de digitalización, los tres parámetros de la señal de video son asignados con un valor de muestreo numérico. Se toman grupos de cuatro píxeles^[2] de video dentro de las tres componentes y se hace un muestreo. Con una señal de video 4:2:2, se toman los cuatro píxeles de luminancia, dos de Cb y dos de Cr. Esto nos da un muestreo 4:2:2.

La norma 4:2:2, también conocida como CCIR 601, define los parámetros básicos del sistema de televisión digital para asegurar la mayor compatibilidad mundial. Se basa en una señal Y, Cr, Cb en el formato llamado 4:2:2 (4 muestreos Y por 2 muestreos Cr y 2 muestreos Cb), con una digitalización sobre 8 bits, con posibilidad de ampliarla a 10 bits para aplicaciones más exigentes.

Cualquiera que sea el estándar de barrido, la frecuencia de muestreo es de 13.5 MHz para la luminancia Y. Para las señales de crominancia Cr y Cb, dado su ancho de banda más limitado se muestrean a la mitad de la frecuencia de la luminancia, es decir, 6.75 MHz. Lo que se corresponde con una definición de 720 muestreos por línea en luminancia y de 360 muestreos por línea de crominancia, cuya posición coincide con la de los muestreos impares de luminancia. Ver la figura 3.5.

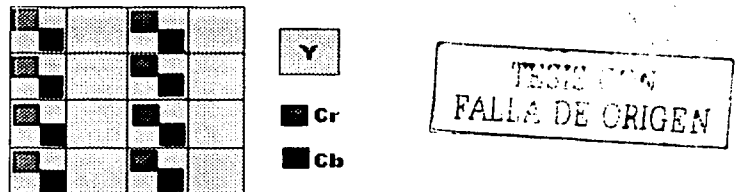


Figura 3.5 Posición de los muestreos en el formato 4:2:2.

El número de bits/muestra es de 8, tanto para la luminancia como para las señales diferencias de color, lo que corresponde a 2^8 niveles = 256 niveles de cuantización.^[2]

La luminancia utiliza 220 niveles a partir del 16 que corresponde al nivel de negro, hasta el 235 correspondiente al nivel de blanco. Se acepta una pequeña reserva del 10% para evitar que ocurran sobremodulaciones. Ver la figura 3.6.

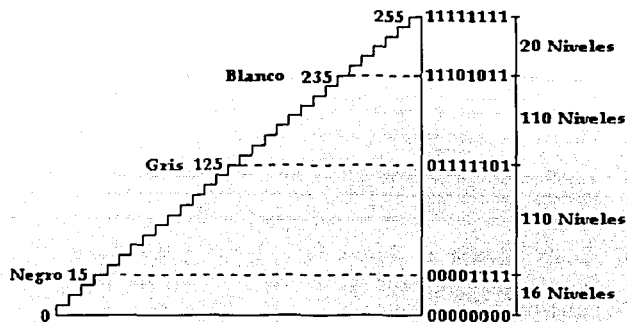


Figura 3.6 Cuantificación de la señal de luminancia.

Para las señales diferencias de color se utilizan 224 niveles, que se reparten a ambos lados del cero análogo, que se hace corresponder con el número digital 128. Así pues, la señal variará entre los valores extremos $128 + 112 = 240$ y $128 - 112 = 16$, con una reserva de 16 niveles a ambos lados. Ver la figura 3.7.

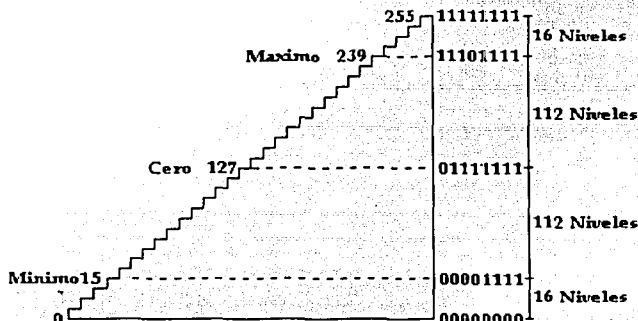


Figura 3.7 Cuantificación de la señal de crominancia.

Para digitalizar las señales analógicas de luminancia y crominancia, se requiere:

- Luminancia Y: Una frecuencia de muestreo de 13.5MHz.
- Para la señal diferencia de color B-Y (CB): Una frecuencia de muestreo de 6.75MHz.
- Para la señal diferencia de color R-Y (CR): Una frecuencia de muestreo de 6.75MHz.

En total se obtiene un total de 27,000,000 de muestras por segundo.

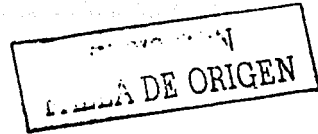
Si en la conversión A/D se utiliza para cada muestra un total de 8 bits, obtenemos un total de $27,000,000 \text{ muestras/s} \times 8 \text{ bits/muestra} = 216 \text{ Mbits/s}$ con un ancho de banda aproximado de transmisión de 108 MHz.

Para aplicaciones profesionales se utilizan 10 bits por muestra esto nos lleva al resultado de un total de $27,000,000 \text{ muestras/s} \times 10 \text{ bits/muestra} = 270 \text{ Mbits/s}$ ocupando un ancho de banda para transmisión de aproximadamente 135 MHz.

Una emisión de televisión digital tiene un ancho de banda superior al de una transmisión analógica.

Cabe señalar que la digitalización de la parte útil de la señal de vídeo solo requiere 166 Mbit/s, si se tiene en cuenta la inutilidad de digitalizar los intervalos de supresión del haz (también llamados "blanking") de línea y campo. Por tanto, estos tiempos libres pueden aprovecharse para transportar los canales de sonido digital, así como datos de servicio u otros.

La norma 4:2:2 CCIR 601 Se muestra en la tabla 3.2.



Parámetros	Sistema NTSC 525 líneas 60 campos
1. Señales codificadas	Y, Cr, Cb
2. Número de muestras por línea completa ➤ Luminancia ➤ Crominancia	858 429
3. Estructura de muestreo	Ortogonal, estructura idéntica de todos los campos y cuadros. Las señales Cr y Cb se muestran simultáneamente con las muestras impares de la luminancia (1, 3, 5, etc.)
4. Frecuencia de muestreo ➤ Luminancia ➤ Crominancia	13.5 MHz 6.75 MHz
5. Codificación	Cuantificación lineal. Codificación con 8 bits por muestra para la luminancia y cada señal de crominancia.
6. Número de muestras activas por líneas digitales: ➤ Luminancia ➤ Crominancia	720 360

<p>7. Correspondencia entre los niveles de video y de cuantificación:</p> <ul style="list-style-type: none"> ➤ Luminancia ➤ Crominancia 	<p>220 niveles de cuantificación. El nivel negro corresponde al número digital 16; el nivel nominal de blanco al número 235.</p> <p>224 niveles de cuantificación en la parte central del margen de cuantificación. El nivel de video cero corresponde al número 128.</p>
---	--

Tabla 3.2 Parámetros de la norma 4:2:2 para NTSC.

3.5. Compresión de vídeo bajo MPEG

Los métodos de compresión, recurren a los procedimientos generales de compresión de datos, aprovechando además la redundancia espacial de una imagen (áreas uniformes), la correlación entre puntos cercanos y la menor sensibilidad del ojo a los detalles finos de las imágenes fijas (JPEG^[2]) y, para imágenes animadas (MPEG^[2]), se saca provecho también de la redundancia temporal entre imágenes sucesivas.

En el año de 1990, la ISO^[6], preocupada por la necesidad de almacenar y reproducir imágenes de video digitales y su sonido estereofónico correspondiente, creó un grupo de expertos que llamó MPEG (Moving Pictures Expert Group) procedentes de aquellas áreas implicadas en el problema (telecomunicaciones, informática, electrónica, radio difusión, etc.).

El primer trabajo creado por este grupo es la norma ISO/IEC 11172, mejor conocida como MPEG-1, en el año 1992. La idea inicial era la de permitir el almacenamiento y reproducción en soporte CD-ROM con un flujo de transmisión de datos del orden de 1.5 Mbits/s, transportando tanto imagen como sonido.

El estándar MPEG además de aprovechar la redundancia espacial intrínseca de una imagen fija utilizada en la codificación JPEG, aprovecha la redundancia temporal que aparece en la codificación de imágenes animadas, permitiendo encontrar similitudes entre las imágenes sucesivas de video.

Debido a que la calidad en la compresión de video en el estándar MPEG-1 era de baja calidad y no servía para otras aplicaciones, se creó la norma ISO/IEC 13818, mejor conocida con el nombre de MPEG-2. Esta norma permite un flujo de transmisión hasta el orden de los 20 Mbits/s, transportando tanto imagen como sonido. Esta es la norma que se utilizará en la televisión digital de ATSC.

3.6 Compresión de Vídeo con MPEG-2

Para analizar el proceso de compresión de video de acuerdo con este estándar, se puede dividir en compresión Intraframe y compresión Interframe.

3.6.1 Compresión Intraframe

La imagen que se muestra a continuación (figura 3.8) contiene una gran cantidad de redundancia espacial o información innecesaria, la cual puede ser desechada antes de la transmisión.



TESIS CON
FALLA DE ORIGEN

Figura 3.8 Ejemplo de redundancia espacial.

Debido a la cantidad de datos involucrada y la redundancia psicovisual en las imágenes, MPEG emplea un esquema de compresión con pérdidas basado en la codificación por transformación.

❖ Procesos usados por la codificación intraframe

Para realizar la compresión intraframe de video, también llamada codificación espacial se utilizan los procesos que se describen a continuación.

Submuestreo de la entrada 4:2:2: La señal de entrada se asume como 4:2:2 SDI (Serial Digital Interface) que tiene una longitud de palabra de 8 o 10 bits. MPEG solo utiliza resolución de 8 bits, por lo que se necesita una etapa de redondeo cuando la señal contenga palabras de 10 bits. Al cambiar de 10 bits a 8, se reduce un 20% la cantidad de datos, debido a que al representar cada muestra con 10 bits se tiene una tasa de transmisión de 270 Mbits/s, mientras que al representarla con 8, la tasa de transmisión será de 216 Mbits/s.

Se tienen que eliminar también los intervalos de blanqueo horizontal y vertical. Al hacer esto se tiene una reducción de aproximadamente el 8% al eliminar el intervalo de blanqueo horizontal y de un 17% al eliminar el intervalo de blanqueo vertical, lo que se traduce en que de los 216 Mbits/s nos quedan 166 Mbits/s. Se tienen 720×480 pixeles de luminancia (Y) a la entrada, y se trabaja con 30 cuadros/s.

Además la mayoría de los perfiles operan con un muestreo de 4:2:0, por lo que se usa una etapa de filtro paso bajas vertical e interpolación, para convertir los datos de entrada

que están bajo un muestreo 4:2:2 a 4:2:0, logrando una disminución en los datos a transmitir.

Para convertir los datos de entrada de 4:2:2 a 4:2:0 los canales de crominancia Cr y Cb (ver figura 3.9), son submuestreados por 2 en ambas direcciones. Los canales de crominancia contienen la mitad de muchas líneas y píxeles por línea comparados al canal de luminancia.

TESTS CON FALLA DE ORIGEN

Y1	Y2	Y3	Y4	Cr1	Cr2	Cb1	Cb2
Y5	Y6	Y7	Y8	Cr3	Cr4	Cb3	Cb4
Y9	Y10	Y11	Y12				
Y13	Y14	Y15	Y16				

Figura 3.9 Submuestreo de los canales de crominancia.

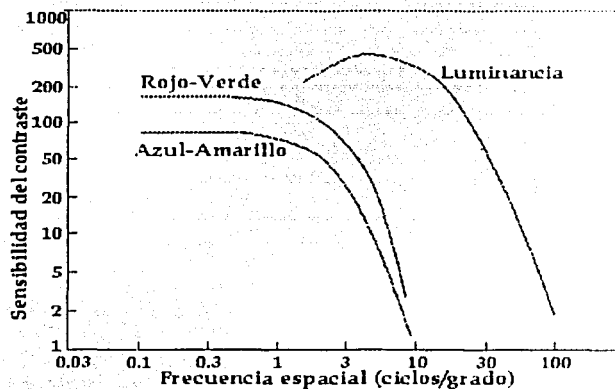


Figura 3.10 Respuesta visual a las variaciones de luminancia y crominancia.

El formato 4:2:0 indica que la señal de video es muestreada en sus componentes de luminancia y señales de diferencia de color a frecuencias distintas: la frecuencia de muestreo de la luminancia (Y) es dos veces mayor, lo que significa que por cada cuatro muestras de luminancia que se toman de una imagen se tendrán dos muestras de diferencia de color, una muestra de la señal de diferencia de color rojo (Cr) y una muestra de la señal de diferencia de color azul (Cb). Esto significa que este formato trabaja con una señal digital por componentes. En el formato 4:2:0 las muestras de crominancia que aparecen entre dos filas de luminancia significan que se toma el promedio de las muestras que corresponden a la misma columna de dos filas consecutivas, de tal forma que se presenta en la imagen el valor promedio obtenido.

Si cambiamos de 4:2:2 a 4:2:0 se reduce en un 25% la cantidad de datos que tenemos, quedándonos una tasa de transmisión de 124.5 Mbit/s de los 166 Mbit/s, lo que se traduciría en un ancho de banda de 62.25 MHz. La cantidad de datos todavía es muy grande si tomamos en cuenta que para transmitir video analógico se usa un ancho de banda de 5 MHz.

A fin de reducir la redundancia psicovisual en las imágenes, MPEG-2 incorpora las características del sistema visual humano en el proceso de compresión a través de la especificación de matrices de cuantificación. Se conoce que la respuesta en frecuencia del sistema visual humano decae con el incremento de la frecuencia espacial. Además, esta

caída es más rápida en los dos canales de crominancia. (véase la figura 3.10) Esto implica que una pequeña variación en la intensidad es más visible en regiones de variación lenta que en las regiones de variación rápida, y también más visible en la luminancia comparada con una variación similar en la crominancia.

Los elementos de estas matrices son basados en la visibilidad de funciones base individuales DCT de 8×8 con una distancia de observación igual a 6 veces el ancho de pantalla. Las funciones base fueron vistas con resolución de luminancia de 720 píxeles \times 576 líneas y una resolución de crominancia de 360×576 . Las matrices sugieren que estos coeficientes DCT que corresponden a imágenes base con baja visibilidad pueden ser mas toscamente cuantificados.

En el sistema básico, denominado a veces sistema básico secuencial, la precisión de los datos de entrada y de salida está limitada a 8 bits, mientras que los valores cuantificados de la DCT (Transformada Discreta del Coseno) están limitados a 11 bits.

La compresión de una imagen a color sigue los siguientes pasos:

Segmentación de la imagen en bloques o subimágenes. La imagen original en la forma Y, Cr, Cb se divide en bloques de 8×8 píxeles (véanse las figuras 3.11 y 3.12), que se procesan de izquierda a derecha y de arriba abajo. Siendo, para una imagen en formato CCIR 601 de 720×576 , un total de 6480 bloques de luminancia Y , y 1620 bloques para cada una de las componentes Cr y Cb . Con cada uno de estos bloques se forman tres matrices: una matriz de 64 números de 0 a 255 (para imágenes de 8 bits) para la luminancia, y dos que van de -128 a $+127$ para las componentes Cr y Cb .

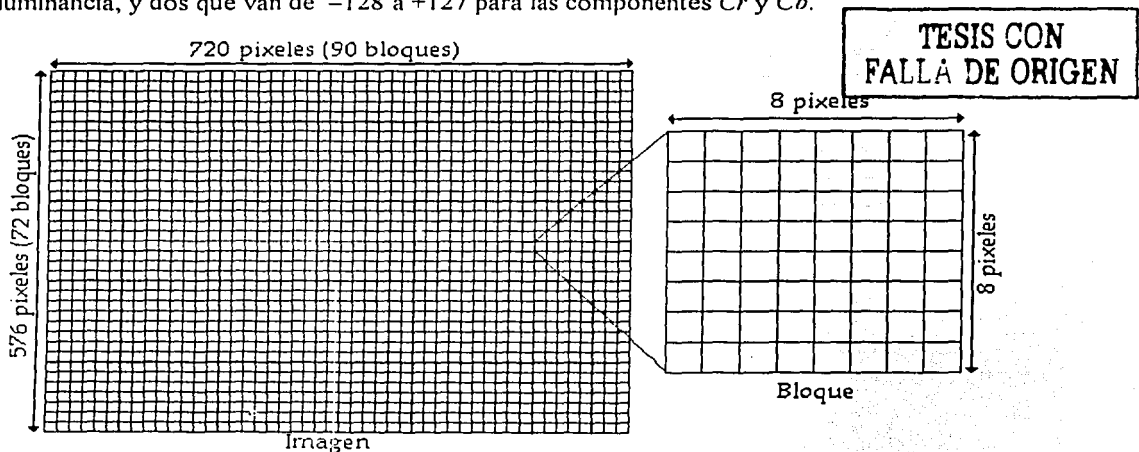


Figura 3.11 División en bloques o subimágenes de 8×8 píxeles

Cálculo de la DCT. El proceso matemático llamado Transformada Discreta del Coseno (DCT: Discrete Cosine Transform) convierte los valores de los píxeles del bloque de datos desde el dominio del tiempo al dominio de la frecuencia. ^[4.7.8]

La Transformada Discreta del Coseno (DCT) es una versión muestreada de la transformada del coseno y se usa extensamente en su forma bidimensional en MPEG. La DCT aplica el principio de la Transformada Rápida de Fourier, que indica que es posible obtener una representación muy aproximada de una señal física a través de un número finito de coeficientes de frecuencia, mediante la obtención de valores discretos de su espectro en el dominio de la frecuencia, los cuales representan una serie finita de cósenos. Los valores que contienen la información de la imagen están colocados en tres matrices, en donde cada una de ellas contiene uno de los componentes de la señal de video (Y, Cr o Cb).

Para realizar la DCT, es necesario segmentar cada matriz en bloques bien definidos. La segmentación de la imagen se realiza en bloques cuadrados cuyas dimensiones sean potencia de dos. Para este caso se utiliza el valor de 8, por lo tanto la imagen es segmentada en bloques de 8x8 píxeles. A cada uno de los bloques se le aplica el algoritmo de la DCT. Un bloque de 8x8 píxeles es transformado y llega a ser un bloque 8x8 coeficientes. Debido a que la transformada requiere multiplicación por fracciones, la extensión de la longitud de la palabra aumenta, pues los coeficientes tienen una longitud mayor que los valores de los píxeles. Típicamente un bloque de píxeles de 8 bits resulta en un bloque de coeficientes de 11 bits. Por lo tanto, la DCT no resulta en compresión alguna, sin embargo convierte los píxeles fuente a una forma en que la compresión es más sencilla.

El proceso se aprovecha de la correlación que existe entre muestras del mismo bloque. Se aplica la DCT a cada uno de los bloques Y, Cr, Cb, generando para cada uno de ellos una nueva matriz de 8x8 compuesta por los coeficientes de las componentes de frecuencias espaciales. En la figura 3.13 se muestra un ejemplo de los coeficientes de la DCT. El valor de estos coeficientes disminuye rápidamente cuando se van alejando del origen de la matriz, terminando generalmente en una serie de ceros. De esta forma si un bloque es de luminancia y color uniformes, únicamente el primer coeficiente (coeficiente DC) no será nulo, y así solo habrá que transmitir un único coeficiente en lugar de 64.

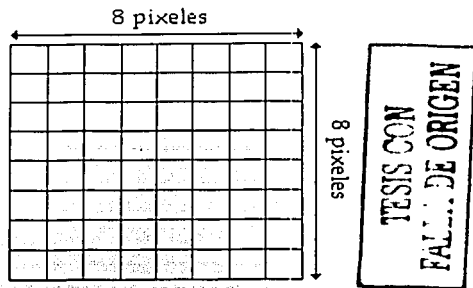


Figura 3.12 Bloque o subimagen de tamaño 8x8.

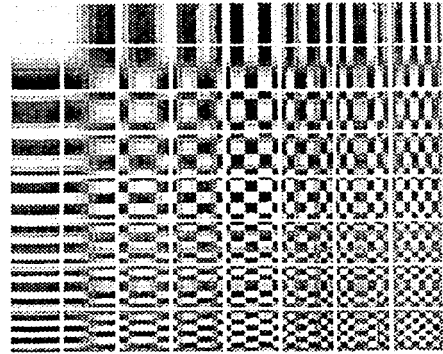


Figura 3.13 Coeficientes de la DCT.

Discriminación por cuantización. Se han realizado estudios que demuestran que el ojo humano es menos sensible a los cambios rápidos de contorno (frecuencia altas) que a los cambios de matiz (frecuencias bajas), por lo que al momento de eliminar la información menos importante de la imagen es preferible que sea la que pertenece a las frecuencias altas. De acuerdo con lo anterior, es necesario darle un nivel de importancia distinto a cada coeficiente obtenido, por lo cual se realiza una cuantización. Una DCT divide la señal en frecuencias diferentes, con lo que se hace posible el controlar el espectro de ruido: los coeficientes de baja frecuencia se entregan con mayor exactitud que los de alta frecuencia por medio de la cuantización. Esta se realiza dividiendo los coeficientes de la matriz DCT por una matriz de cuantización, de tal forma que los coeficientes de baja frecuencia serán divididos por números pequeños, mientras los coeficientes de frecuencias altas serán divididos por números grandes.

Después de la división los bits menos significativos son descartados o truncados. Este truncamiento es una forma de recuantización. Se toma el valor entero más cercano y se multiplica por unos valores de escala. Como resultado se tiene que los coeficientes que representan las frecuencias espaciales bajas son recuantizadas con escalones más pequeños y los afecta poco el ruido. Los coeficientes que representan las frecuencias espaciales altas se recuantizan con escalones más grandes y les afecta más el ruido. Sin embargo, menos escalones, significa que se necesitan menos bits para identificar cada escalón y se obtiene una compresión.

A la matriz de cuantificación se le pueden aplicar factores de escala para obtener diversos niveles de compresión. La información en las bajas frecuencias es codificada con mucho detalle y los valores en la parte izquierda de la parte más alta de la matriz son enviados con un alto nivel de resolución. Para darle la importancia que se merece a cada frecuencia, cada uno de los coeficientes son escalados por un factor, antes de la cuantificación, y así se da mayor peso a unos coeficientes que a otros.

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 se pueden enviar primero todos los coeficientes diferentes de cero, seguidos de un código que indique que los restantes son cero. La exploración es una técnica que aumenta la probabilidad de alcanzar este resultado, porque envía los coeficientes en orden descendente según su probabilidad.

La figura 3.14 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. El primer coeficiente del barrido en zig-zag es conocido como el coeficiente DC mientras que el resto son los coeficientes AC (ver figura 3.14). Con la excepción del coeficiente DC que se trata por separado, los 63 coeficientes AC se leen en zig-zag para transformar la matriz en una cadena de datos en serie, adaptada a la próxima etapa del proceso.

TESIS CON FALLA DE ORIGEN

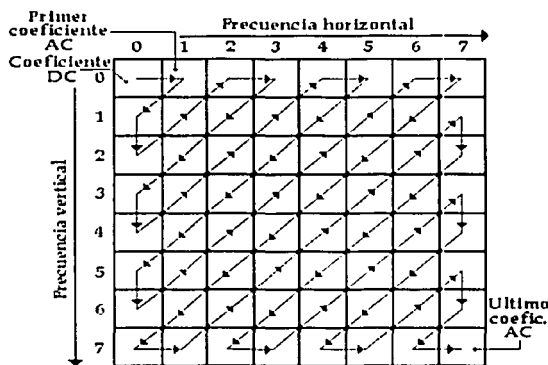
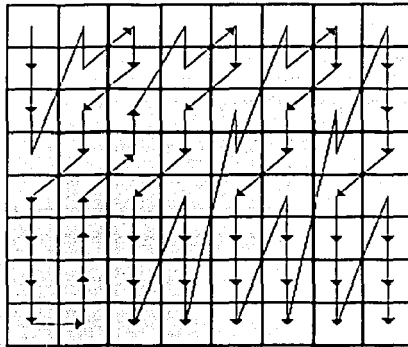


Figura 3.14 Barrido o exploración en zig-zag.

En la figura 3.15 se observa una exploración alternativa que presenta mejores resultados para las imágenes de video entrelazado, 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°. En la figura se muestra que esta forma de exploración, entrega primero las frecuencias espaciales verticales y luego las frecuencias espaciales horizontales.

Para convertir los coeficientes en una cadena unidimensional se reordenan usando la técnica de exploración adecuada, dependiendo del tipo de fuente. El coeficiente DC de cada bloque es codificado usando Differential Pulse Code Modulation (DPCM). Esta técnica

codifica la diferencia entre los coeficientes DC cuantizados del bloque actual y los coeficientes DC cuantizados del bloque anterior. Es decir, se codifica la diferencia entre coeficiente DC del presente bloque y el del bloque previamente codificado. Esto permite aumentar la precisión de cuantificación para un número de bits dado, de manera que se puede minimizar la visibilidad de los bloques sobre la imagen reconstruida, ya que el ojo, si bien es poco sensible a los detalles finos, es, por el contrario, muy sensible a pequeñas diferencias de luminancia sobre áreas uniformes.



TESIS CON
FALLA DE ORIGEN

Figura 3.15 Exploración alternada.

VLC y RLC. En este proceso se aplica la codificación entrópica de Huffman (Asignación del Código de Longitud Variable o VLC). Los datos resultantes de la DCT, cuantificación, escalado y la técnica de exploración pueden ser ampliamente reducidos aplicando la variable de longitud.

La probabilidad de la ocurrencia de valores de coeficientes particulares se puede estudiar. En la práctica, algunos valores ocurren muy seguido, otros ocurren con menor frecuencia. El nombre del bloque VLC o Codificación de Longitud Variable, se debe a la aplicación del Código de Huffman, el cual aplica conceptos de teoría de la información para reducir la cantidad de bits a transmitir. Este código consiste en asignar un nuevo código binario a los valores existentes, donde los valores con mayor probabilidad de aparecer tendrán un nuevo número binario con la menor cantidad de dígitos posibles y mientras la probabilidad de ocurrencia disminuya en los demás valores, el número binario tendrá más dígitos. Para prevenir confusiones, ninguna palabra codificada puede ser el prefijo de otra. Con esto la compresión de la señal de video es todavía mayor.

En el video real, no todas las frecuencias espaciales están presentes simultáneamente, por lo que la matriz de coeficientes DCT tendrá ceros dentro de ella. A pesar del uso de la técnica de exploración, los coeficientes que valen cero aparecerán todavía entre los valores significativos, además de que frecuentemente se encuentra una larga cadena de ceros al final de un bloque de datos. Para utilizar de una forma más eficiente estos coeficientes se les aplica una Codificación de Longitud Continua o RLC (Run Length Coding). Donde se

repitan valores, tales como cadenas de ceros, RLC transmite simplemente el número de ceros en lugar de los bits individuales. MPEG-2 ha sido diseñado de modo que se beneficia de la existencia de estas largas series de ceros que se producen normalmente en la reordenación por medio de la técnica de exploración.

❖ Codificador espacial

La señal de entrada se asume como 4:2:2 SDI, la cual tiene palabras con longitudes de 8 o 10 bits. MPEG utiliza solo la resolución de 8 bits, por lo que se pone la etapa de redondeo. Como la mayoría de los perfiles de MPEG operan con un muestreo 4:2:0, se coloca una etapa de conversión a este formato. Una vez que la entrada está bajo este formato necesita ser almacenada para que pueda ser convertida a bloques de 8×8 píxeles. La etapa DCT transforma la información de la imagen al dominio de la frecuencia. Después de la transformada, los coeficientes son cuantizados y truncados, con lo que se obtiene una compresión significativa. Después se exploran los coeficientes con la técnica del "zig-zag", con la cual se incrementa la probabilidad de que en la exploración se encuentren primero los coeficientes significativos. Después del último coeficiente distinto de cero, se genera un código de fin de bloque (end of block -EOB).

Los datos de los coeficientes se comprimen más gracias a las codificaciones de longitud variable y de longitud continua. En un sistema con tasa de transmisión fija se utiliza una memoria buffer para que absorba las variaciones en la tasa de transmisión debida a la dificultad de codificación de determinada señal. Al final del proceso de compresión se pueden obtener tasas de transmisión de 1.15 Mbps.

3.6.2 Compresión Interframe

La compresión interframe está basada en la redundancia temporal. En la figura 3.16 pueden verse dos imágenes consecutivas, las cuales utilizan una gran cantidad de elementos de redundancia en su transmisión. Si bien existen algunos cambios, estos son relativos. Se puede aprovechar el hecho de que hay información que no ha cambiado, enviando solo los cambios para disminuir la cantidad de datos transmitidos.

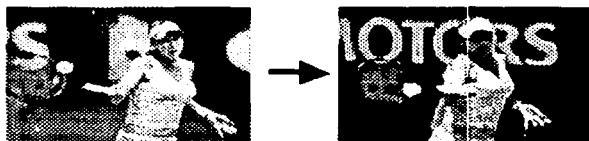


Figura 3.16 Ejemplo de la redundancia temporal.

Para obtener beneficios de la redundancia temporal se usa la "inter-codificación", es decir que se transmiten solamente las diferencias entre las imágenes. Un retraso de una imagen combinado con un restador puede calcular las diferencias entre imágenes. La imagen de diferencia es una imagen en toda la extensión de la palabra que puede ser comprimida más con la ayuda del codificador espacial. El decodificador revierte la

codificación espacial y añade la imagen diferencia a la imagen previa para obtener la imagen siguiente.

Existen algunas limitantes en este sistema de codificación, pues es imposible decodificar si la transmisión ya ha comenzado, además de que si hay datos incorrectos, estos se propagarán indefinidamente. La solución es utilizar un sistema que no sea completamente diferencial, es decir que se envíen imágenes completas periódicamente. Estas imágenes se llaman "Intra" (imágenes I) y se obtienen solamente mediante la compresión espacial. Si existe algún error o hay un cambio de canal, será posible hacer una decodificación correcta con la siguiente imagen I.

La esencia de este tipo de compresión, es mirar elemento por elemento entre escenas de forma individual. Para esto se agrupan en bloques los elementos de la imagen, típicamente en matrices de 8x8 píxeles. Un bloque está compuesto por un grupo de 8x8 con valores que representan información de la crominancia y luminancia de parte de la imagen. Los bloques de información son agrupados dentro de macrobloques consistentes en cuatro bloques de información de luminancia, por lo tanto cada macrobloque contiene 16 píxeles x16 líneas con valores de luminancia y un número de bloques de 8 píxeles x8 líneas de crominancia. Los macrobloques representan cada una de las diferencias de color (Cb y Cr). Para el formato de crominancia 4:2:0, son usados dos bloques de crominancia uno para Cb y otro para Cr (véase la figura 3.17). Varios de los macrobloques son agrupados en secuencias conocidas como Slices. Estas slices (rebanadas) combinados crean un fotograma. Las rebanadas (slices) son usadas para posibles detecciones de errores.

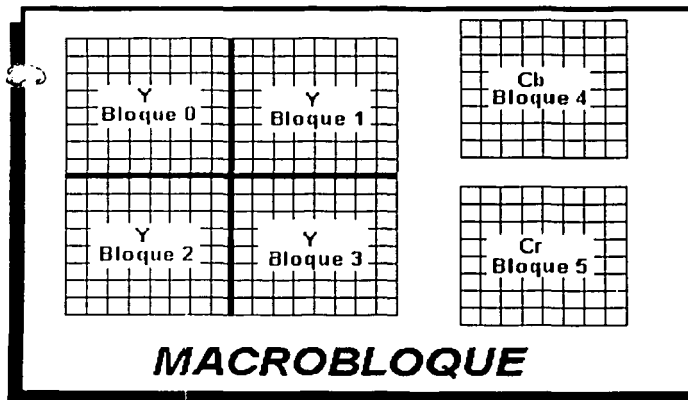


Figura 3.17 Composición de un macrobloque

TESIS CON FALLA DE ORIGEN

❖ Compensación de movimiento

Cuando las secuencias contienen movimiento, grandes cantidades de datos se pueden comprimir usando un proceso conocido como "Compensación de movimiento" (Motion Compensation).^[2] Por medio de este proceso se mide eficazmente el movimiento de los objetos de una imagen a otra (véase la figura 3.18). De este modo se consigue medir qué tipos de movimientos se repiten entre imágenes. En este se utiliza el hecho de que las 25 imágenes que aparecen durante 1 segundo pueden dar la posibilidad de identificar la velocidad y dirección en la cual una parte de la imagen se está moviendo en un fotograma además de donde acabará 25 fotogramas después un segundo más tarde.

El movimiento reduce las similitudes entre imágenes e incrementa la cantidad de datos necesarios para transmitir las imágenes. La compensación de movimiento se utiliza para incrementar las similitudes. Cuando un objeto se mueve a lo largo de la pantalla del televisor, aparece en diferentes lugares en cada imagen, pero su apariencia no cambia demasiado. Las diferencias pueden ser reducidas al medir el movimiento en el codificador. Este enviará al decodificador un vector. El decodificador utilizará el vector para desplazar parte de la imagen previa al lugar adecuado de la imagen nueva.

Un vector controla el desplazamiento de una superficie completa de la imagen que se conoce como macrobloque. El tamaño de un macrobloque está determinado por la codificación DCT y la estructura de submuestreo de color. En un sistema 4:2:0, el espaciado vertical y horizontal de las muestras de color es exactamente el doble del de la luminancia. Un solo bloque de DCT de 8x8 muestras de color se extiende sobre la misma área de cuatro bloques de luminancia de 8x8; por lo que es el área de menor tamaño que puede ser desplazada por un vector. Un macrobloque 4:2:0 contiene cuatro bloques de luminancia, un bloque Cr y un bloque Cb.

El estimador de movimiento trabaja comparando los datos de luminancia de dos imágenes sucesivas. Un macrobloque en la primera imagen es utilizado como referencia. La correlación entre la referencia y la imagen siguiente se mide en todos los desplazamientos posibles con una resolución de la mitad de un píxel (1/2 píxel) sobre el rango de búsqueda completo. Cuando se encuentra la correlación más grande, se asume que representa el movimiento correcto.

El vector de movimiento tiene componentes vertical y horizontal. En un programa de televisión típico, un objeto en movimiento puede extenderse sobre un número de macrobloques. Un factor de compresión mayor se obtiene si los vectores se transmiten diferencialmente. Cuando se mueven objetos grandes, los macrobloques adyacentes tienen los mismos vectores y la diferencia entre vectores es igual a cero. Los vectores de movimiento están asociados con macrobloques, no con objetos reales.

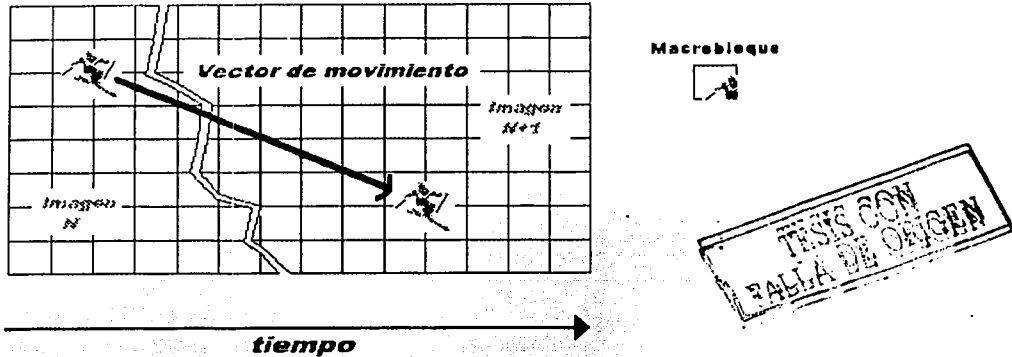


Figura 3.18 Medición del movimiento de una imagen a otra

Los Macrobloques se utilizan como las partes apropiadas de este tipo de codificación.

El flujo de datos MPEG consiste en una serie continua de imágenes codificadas, y consisten en varias escenas predichas y referencias de imágenes I. Algunas de las imágenes predichas, pueden ser codificadas más eficientemente que las imágenes I, por lo que sería más deseable transmitir imágenes de este tipo. Algunas de las imágenes transmitidas en el flujo de datos son críticamente dependientes de las predecesoras y sucesoras, por lo tanto no se pueden saltar aleatoriamente para comenzar la decodificación.

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

❖ Codificación bidireccional.

Cuando un objeto se mueve, oculta el fondo en su extremo delantero y lo deja visible en su extremo trasero. El que se descubre el fondo requiere que se transmitan nuevos datos, pues esta área se cubrió por el objeto y no se puede obtener información de la imagen previa, lo mismo sucede si la cámara hace un "alejamiento". MPEG ayuda a minimizar este problema al usar la codificación bidireccional, la cual permite que la información sea

tomada de imágenes anteriores y posteriores a la imagen actual. Si un fondo se está revelando, estará presente en una imagen posterior, y la información puede ser transportada hacia una imagen anterior para recrear dicha imagen.

En una base de macrobloques individuales, una imagen codificada bidireccionalmente puede obtener la información de la compensación de movimiento de una imagen anterior o posterior, o incluso usar un promedio de los datos posteriores y anteriores. La codificación bidireccional reduce significativamente la cantidad de datos diferencia al mejorar el grado de predicción posible.

Un compresor inteligente sería aquel que intente las tres estrategias de compresión y seleccione aquella que resulte en la menor cantidad de datos transmitidos.

❖ Tipos de imagen MPEG

En MPEG se necesitan tres tipos diferentes de imágenes (I, P y B) para realizar la codificación diferencial y bidireccional, con lo que se reduce la propagación de errores.^[9]

Los tres tipos de imágenes se encadenan según el esquema de la figura 3.19. Los cuales son el soporte de la codificación diferencial y bidireccional, posibilitando una predicción temporal de pérdidas o imágenes incompletas que recibe el receptor, minimizando la propagación de errores.

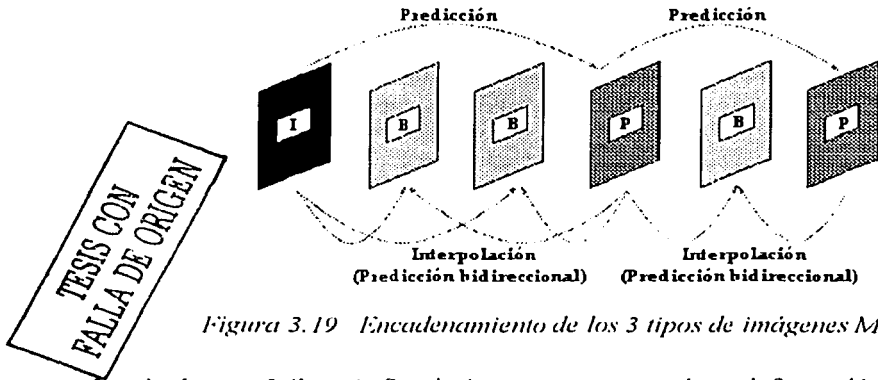


Figura 3.19 Encadenamiento de los 3 tipos de imágenes MPEG

Las imágenes I (intra). Son imágenes que no requieren información adicional para su decodificación. Son codificadas sin ninguna referencia a otras imágenes, es decir, que contiene todos los elementos necesarios para su reconstrucción por el decodificador y son, por ello, el punto de entrada obligatorio para el acceso a una secuencia. La tasa de compresión de imágenes I es relativamente pequeña. Las imágenes I están simplemente formadas por los coeficientes de la transformada y no contienen vectores de movimiento. Requieren una gran cantidad de datos, si se comparan con los otros tipos de imágenes.

Las imágenes P (previstas). Se codifican con respecto a las imágenes de tipo I o P anteriores, gracias a las técnicas de predicción con compensación de movimiento. Como la compensación de movimiento no es perfecta, no se podrá multiplicar indefinidamente el número de imágenes P, ya que, como se utilizan para decodificar otras imágenes P o B, se propagan los errores de codificación. Su tasa de compresión es mayor que la de las imágenes I. Las imágenes P requieren aproximadamente la mitad de los datos de las imágenes I. Una imagen P utiliza la imagen previa más próxima (I o P), la cual es base para sus predicciones y es llamada Forward Prediction (Predicción hacia adelante). La imagen P servirá como referencia para las futuras imágenes P o B, pero si existieran errores en una imagen P particular, buscaría imágenes futuras derivadas de esta. Las imágenes P contienen vectores de movimiento que describen de donde, en la imagen previa, debe ser tomado cada macrobloque. También contiene coeficientes de la transformada que describen la corrección o datos diferentes que deben ser añadidos al macrobloque.

Las imágenes B (Bidireccionales). Son obtenidas por interpolación bidireccional de predicción del movimiento para que el decodificador reconstruya una imagen que está alojada entre dos imágenes de tipo I o P. Incluyen vectores que describen de donde se deben tomar los datos de las imágenes posteriores o anteriores, y coeficientes de la transformada que proporcionan correcciones. Las imágenes B utilizan las imágenes pasadas y futuras (precedentes y siguientes) para hacer predicciones, para ello necesitan más de dos imágenes en memoria. Como no se utilizan para describir otras imágenes, las imágenes B no propagan los posibles errores de codificación. Este tipo de imágenes es el que ofrece el factor de compresión más alto, que generalmente es de una cuarta parte de los datos de las imágenes I.

Grupo de imágenes (GOP – Group of Pictures) Un GOP comienza con una imagen I y luego contiene imágenes P distribuidas a lo largo del grupo. Las imágenes restantes son imágenes B. El GOP termina con la última imagen antes de la siguiente imagen I. La longitud de un GOP es flexible, aunque tiene un valor común entre 12 y 15 imágenes.

Resulta lógico pensar que si las imágenes B van a ser tomadas de una imagen futura, estos datos deben estar disponibles en el decodificador. De modo que la codificación bidireccional requiere que los datos sean enviados fuera de secuencia y sean almacenados temporalmente.

El enviar datos fuera de secuencia requiere memoria adicional en el codificador y en el decodificador causando un retraso. El número de imágenes bidireccionales codificadas entre imágenes I o P debe estar restringido para reducir los costos y retardos.

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

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

Para alcanzar un flujo de video de 1.15 Mbits/s con una calidad satisfactoria, los parámetros comúnmente utilizados son $M=3$ y $N=12$ como se muestra en la figura 3.20.

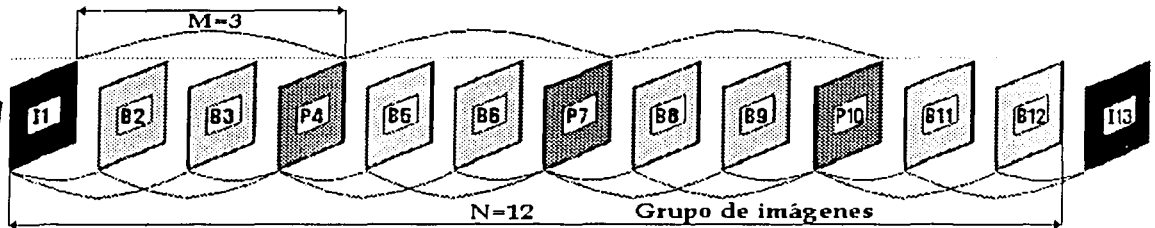


Figura 3.20 Ejemplo de grupo de imágenes, para $M=3$, $N=12$

En este caso, una secuencia de video se compone de $1/12$ (8.33%) de imágenes I, $1/4$ (25%) de imágenes P y de $2/3$ (66.66%) de imágenes B. El factor de compresión global se ve favorecido por el hecho de que las imágenes más frecuentes son las que tienen un factor de compresión mas alto. A la hora de recuperar las imágenes en nuestro receptor, 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 $M=3$ y $N=12$ el modo de codificación de imágenes sucesivas se traduce por la correspondencia que hay entre el número y el tipo de imagen de la forma siguiente:

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

Sin embargo, para codificar o decodificar una imagen B (Bidireccional), el codificador y el decodificador necesitarán la imagen I o P que la precede y la imagen P o I que la sigue. El orden de las imágenes será, por tanto, modificado antes de la codificación, de forma que el codificador y el decodificador dispongan, antes que las imágenes B, de las imágenes I y/o P necesarias para su tratamiento, o sea (Ver figura 3.21):

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

Los datos de las imágenes P se envían antes de los de las imágenes B. Las últimas imágenes B del GOP no pueden ser transmitidas hasta que se haya transmitido la imagen I del siguiente GOP, debido a que estos datos serán necesarios para efectuar la decodificación bidireccional. Para poner en orden a las imágenes de nuevo, se incluye una referencia temporal en cada una de ellas.

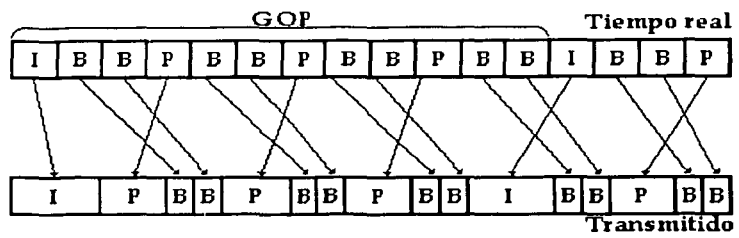


Figura 3.21 Comparación de las imágenes antes y después de la compresión, mostrando un cambio de secuencia.

TESIS CON
 FALLA DE ORIGEN

❖ **Construcción de la trama elemental de video MPEG-2**

Una secuencia de video o trama elemental (Elementary Stream, conocida como ES) de MPEG es básicamente la salida del material en bruto de un codificador y contiene lo necesario para que un decodificador restablezca la imagen original.^[9] La sintaxis de la señal comprimida es definida de manera rigurosa por MPEG, así se asegura que el decodificador cumpla con esta.

En la figura 3.22 se muestra la jerarquía de capas de la trama elemental de video. El proceso de codificación de las imágenes MPEG-2 está formado por capas bien definidas, las cuales se describen a continuación.

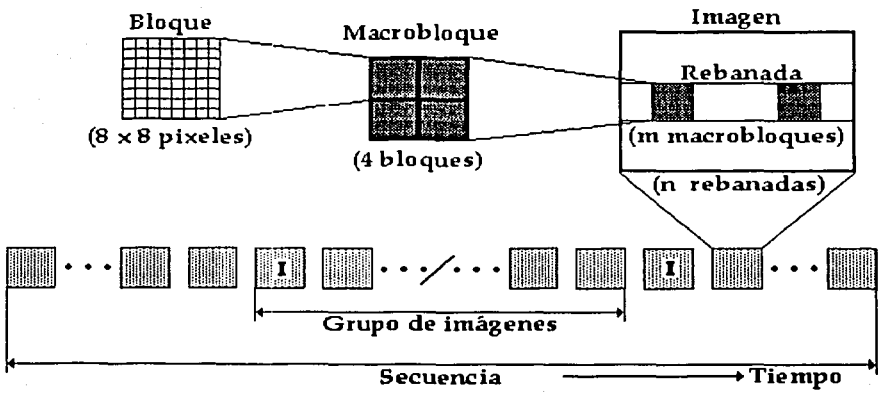


Figura 3.22 Jerarquía de capas de la secuencia

Bloque (Block). Es la unidad fundamental de la información de la imagen y esta representada por un bloque de coeficientes DCT, que tienen un tamaño de 8x8 píxeles, los cuales representan datos Y, Cr o Cb. Aquí el coeficiente DC es enviado primero ya que este

representa con mayor precisión la información de este bloque. Los demás coeficientes son enviados al final de este.

Macrobloque (Macroblock) Los bloques se agrupan en macrobloques, que son la unidad fundamental de la imagen y pueden estar compensadas en movimiento. Cada macrobloque tiene en su encabezado un vector de movimiento de dos dimensiones. En una imagen B, el vector puede ser hacia adelante o hacia atrás. La escala utilizada para la recuantificación de los coeficientes también es indicada. Usando los vectores, el decodificador obtiene información acerca de las imágenes anteriores y las posteriores, produciendo así una predicción de imágenes. Los bloques son transformados inversamente para producir una imagen de rectificación que es adicionada a la imagen prevista que ha sido producida a la salida del decodificador. En un formato de codificación 4:2:0, cada macrobloque tendrá 4 bloques Y, y dos bloques de diferencia de color. Para hacer posible la identificación de cada bloque y sus componentes, estos se envían en un orden específico. Cada macrobloque tiene un tamaño de 16×16 píxeles.

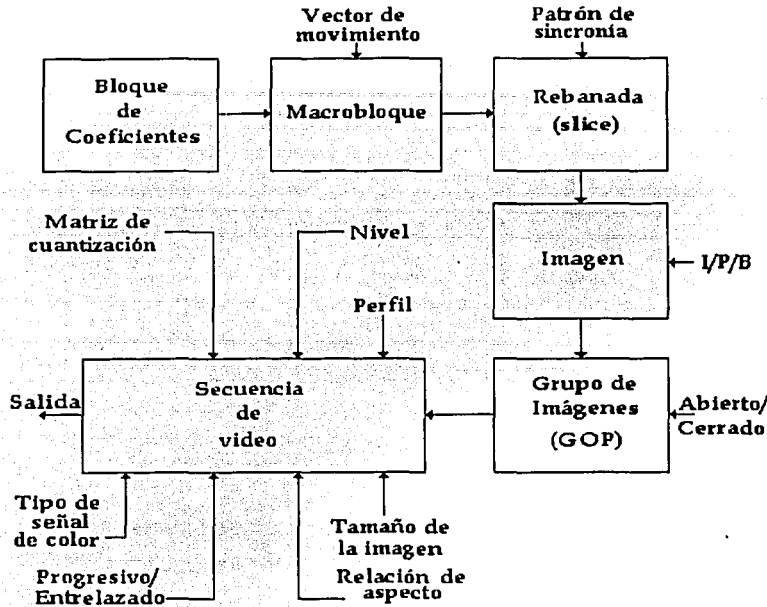
Rebanada (Slice). Los macrobloques son reunidos en rebanadas, y estas siempre deben representar una fila horizontal que está ordenada de izquierda a derecha. En MPEG-2 pueden ser de tamaño arbitrario, y además deben estar compuestas únicamente de macrobloques contiguos situados en la misma línea horizontal. 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. Para la televisión digital de la norma norteamericana, el ATSC establece que las rebanadas deben comenzar en el borde izquierdo de la imagen.

En caso de un bit erróneo en la trama elemental, puede ocurrir una falla en los símbolos del código de longitud variable, o que los coeficientes y vectores codificados diferencialmente sean incorrectos. La estructura de rebanadas permite una recuperación al proporcionar un punto de resincronización en la trama de bits.

Imagen (Picture) de tipo I, P o B. Cuando un número de rebanas se combinan, construyen una imagen, la cual es la parte activa de un campo o un cuadro. El encabezado de la imagen define qué imágenes I, P o B codifica e incluye una referencia temporal para que la imagen pueda ser representada en el momento adecuado. En el caso de tomas panorámicas e inclinaciones, los vectores en cada macrobloque serán los mismos. Un vector global puede ser enviado para toda la imagen, y luego se pueden enviar vectores individuales que lleguen a crear la diferencia en el vector global.

Grupo de imágenes (GOP o Group of Pictures). 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. Entre imágenes I, un número variable de imágenes P y/o B pueden ser colocadas como ya se ha descrito.

Secuencia. En la figura 3.23 se muestra la construcción de una secuencia de video MPEG.



TESIS CON FALLA DE ORIGEN

Figura 3.23 Construcción de una secuencia de video MPEG

Quando 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 de terminación. Códigos de soporte adicional pueden ser situados al inicio de la secuencia. El encabezado de la secuencia contiene parámetros como el tamaño horizontal y vertical de la imagen, norma de barrido, si se usa un barrido progresivo o entrelazado, el perfil, nivel, velocidad de transferencia de bits, y cuales matrices de cuantificación se usan para codificar imágenes espaciales y temporales. La secuencia de video es la trama elemental de video o ES.

Sin los datos del encabezado de la secuencia, un decodificador no puede entender la trama de bits, y por lo tanto estos llegan a ser los puntos de entrada en que los decodificadores pueden comenzar su correcta operación.

3.6.3 Diagrama de Bloques de un Codificador de Video MPEG-2

Un compresor de video MPEG-2, incluiría todos los procesos descritos anteriormente, combinando la codificación espacial (intraframe) y la temporal (interframe). En la figura 3.24 se presenta un diagrama de bloques del proceso de codificación de video.

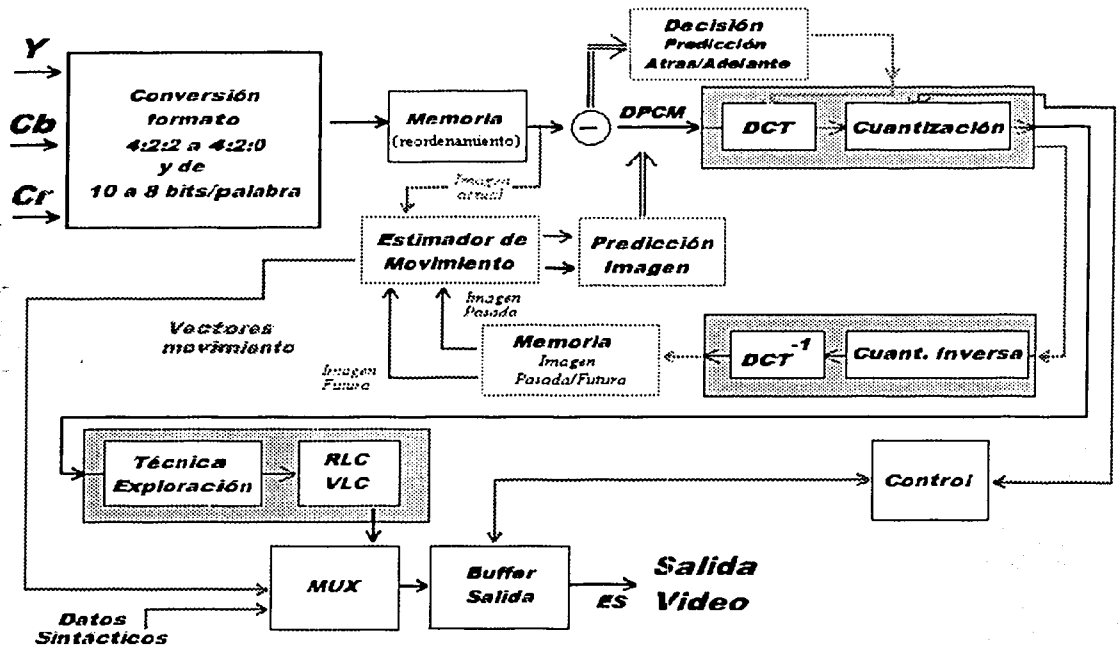


Figura 3.24 Diagrama de bloques de un codificador de video MPEG-2

La entrada pre-procesada (convertida al formato de la señal de entrada en palabras de 8 bits en lugar de 10, se suprimen los intervalos de blanqueo horizontal y vertical, y la conversión del submuestro de señales de video de 4:2:2 a 4:2:0) de video entra en una serie de almacenes de memoria que pueden ser utilizados para cambiar el orden de las imágenes. 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). 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.

TESIS CON FALLA DE ORIGEN

Los datos entran entonces a un restador y al estimador de movimiento, en donde se obtienen los vectores de movimiento para cada macrobloque de 16×16 píxeles. El número de vectores depende del tipo de imagen y del modo de codificación escogido para cada bloque.

Para crear una imagen I, se selecciona el final del retraso de entrada y se apaga el restador para que los datos puedan pasar y sean codificados espacialmente. Después se compara la imagen codificada con la original, con el restador y se envía una señal de error. La salida del restador pasa también a un almacén de frames que puede contener varias imágenes. La imagen I se queda guardada en la memoria.

Para codificar una imagen P, las imágenes B en el buffer de entrada son ignoradas, para que se seleccionen las imágenes futuras. El estimador de movimiento compara la imagen I que se tenía almacenada en memoria a la salida con la imagen P almacenada a la entrada para crear los vectores de movimiento. La imagen I es desplazada con estos vectores para hacer la imagen predicha P. La imagen P predicha es restada de la imagen P real para producir la predicción de errores, los cuales serán espacialmente codificados y serán transmitidos junto con los vectores. La predicción de errores también se añade a la imagen P predicha para crear una imagen P decodificada localmente que también entra a la memoria de salida.

La memoria a la salida contiene una imagen I y una imagen P. Se puede seleccionar una imagen B del buffer de entrada. El compensador de movimiento comparará la imagen B con la imagen I que le precede y con la imagen P que le sigue, para obtener los vectores bidireccionales. La compensación de movimiento hacia delante y hacia atrás se lleva a cabo para producir dos imágenes B predichas. Estas son restadas de la imagen B real. Con una base de macrobloque por macrobloque, los datos hacia delante y hacia atrás son seleccionados de acuerdo con los que presenten las menores diferencias. Las diferencias son entonces codificadas espacialmente y enviadas con los vectores. Cuando todas las imágenes B intermedias están codificadas, se recurre a la memoria de entrada para crear una nueva imagen P a partir de la imagen P anterior.

La salida del compensador de movimiento está codificada espacialmente y los vectores se añaden en un multiplexor. Para hacer la codificación espacial se aplica la DTC. Cada bloque de coeficientes resultante se cuantifica y barre en zig-zag para formar un vector unidimensional de coeficientes. Cada uno de estos es codificado usando el código RLC, para que posteriormente sean codificados por medio de una tabla VLC (codificación Huffman). Después se codifica la información auxiliar necesaria para que el decodificador pueda reconstruir el bloque (modo de codificación, vectores de movimiento, etc).

El coeficiente DC de cada bloque (con valores de coeficientes de la transformada) es codificado usando Differential Pulse Code Modulation (DPCM). Esta técnica codifica la diferencia entre los coeficientes DC cuantizados del bloque actual y los coeficientes DC cuantizados del bloque anterior. Es decir, se codifica la diferencia entre coeficiente DC del presente bloque y el del bloque previamente codificado.

En el bloque IDCT (Inverse Discrete Cosine Transform) se realiza la transformada discreta del coseno inversa, con el fin de poder utilizar la señal de video en el proceso de estimación y predicción de movimiento.

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.

Los datos sintácticos también se añaden para identificar el tipo de imagen (I, P o B) y proporcionar alguna otra información que ayude al decodificador.

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. Las variaciones de información a transmitir se deben a que en una secuencia de video, ciertas escenas tienen más detalles que otras, lo que implica más información que es necesario transmitir, pero sobre todo las variaciones más importantes se dan en la estimación de movimiento. La retroalimentación del buffer con el bloque de cuantización es necesaria para poder hacer un control de la tasa de bits de salida y evitar sobreflujo o subflujo de bits, ya que la velocidad de transmisión que debe entregar este proceso de compresión debe ser constante. Se obtiene entonces a la salida del codificador un tren binario completo, ya utilizable para un decodificador.

Los datos de salida son almacenados en el buffer para permitir variaciones temporales en la tasa de transmisión. Si la tasa de transmisión experimenta un incremento muy grande, el buffer tenderá a llenarse. Para prevenir un sobreflujo el proceso de cuantización será más exigente. Igualmente cuando en el buffer haya un subflujo o "underflow", la cuantización será más relajada para mantener una tasa de transmisión promedio. Esto significa que el almacén en el codificador contiene exactamente lo que contendrá la memoria del decodificador.

3.6.4 Decodificador de video MPEG-2

La decodificación es esencialmente el proceso inverso al de la codificación. Se llevan a cabo los mismos procedimientos, pero en orden inverso. Las tablas de especificación usadas en el proceso de codificación se llevan junto con el flujo de datos después de la compresión y se usan para la descompresión. El esquema donde se muestran los bloques del decodificador de video MPEG-2 se encuentra en la figura 3.25.

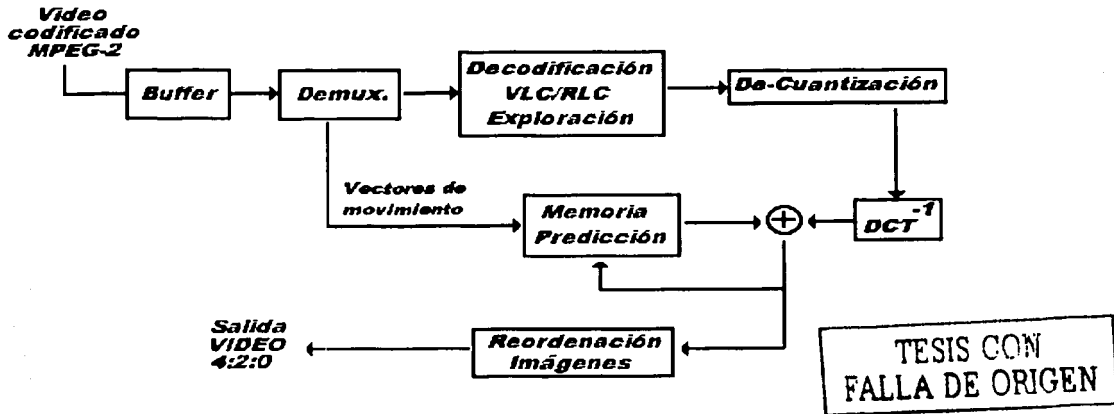


Figura 3.25 Esquema del decodificador de video MPEG-2

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.

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 (IDCT). 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. Por supuesto, este bloque de píxeles de 8x8 puede no reproducir exactamente el original ya que se perdió alguna información en el proceso de codificación.

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 (IDCT), 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.

En el decodificador, se extraen las matrices de la trama de bits en serie, se hace el proceso inverso en la codificación por entropía para obtener los coeficientes cuantizados. Se aplica la cuantización inversa y los coeficientes son colocados en la matriz de acuerdo con la exploración en "zig-zag" para recrear la matriz DCT. Después se efectúa la

transformada inversa, los bloques de 8x8 pixeles son recreados. Para obtener la salida, los bloques son almacenados en memoria RAM, en donde se lee una línea a la vez. Las muestras de color en 4:2:0 se posicionan entre las muestras de luminancia en el eje vertical para que se encuentren igualmente espaciadas cuando se utilice una fuente entrelazada. La necesidad de memoria para el decodificador es de unas 3 imágenes (dos imágenes de referencia más la imagen en vía de reconstrucción), siendo para una imagen 4:2:0, de aproximadamente 16 Mbits.

3.7 Perfiles y niveles MPEG-2

MPEG-2 se puede utilizar en un vasto rango de aplicaciones, requiriendo diferentes grados de complejidad y desempeño. El estándar MPEG-2 se encuentra dividido en perfiles y cada perfil es subdividido en niveles (véase la Tabla 3.3).

NIVELES	PERFILES					
	Simple	Principal	4:2:2	SNR	Espacial	Alto
Alto		4:2:0 1920x1152 80 Mb/2				4:2:0 ó 4:2:2 1920x1152 100 Mb/s
Alto 1440		4:2:0 1440x1152 60 Mb/s			4:2:0 1440x1152 60 Mb/s	4:2:0 ó 4:2:2 1440x1152 80 Mb/s
Principal	4:2:0 720x576 15 Mb/s Sin B	4:2:0 720x576 15 Mb/s	4:2:2 720x608 50 Mb/s	4:2:0 720x576 15 Mb/s		4:2:0 ó 4:2:2 720x576 20 Mb/s
Bajo		4:2:0 352x288 4 Mb/s		4:2:0 352x288 4 Mb/s		

Tabla 3.3 Niveles y perfiles de MPEG-2

Un perfil es básicamente el grado de complejidad esperada en la codificación, mientras que un nivel describe el tamaño de la imagen, la resolución de esta o la velocidad de transferencia de bits usada en ese perfil. 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 o Standard Profile at Main Level). 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. El perfil main es el que se utiliza en el sistema de televisión digital ATSC.

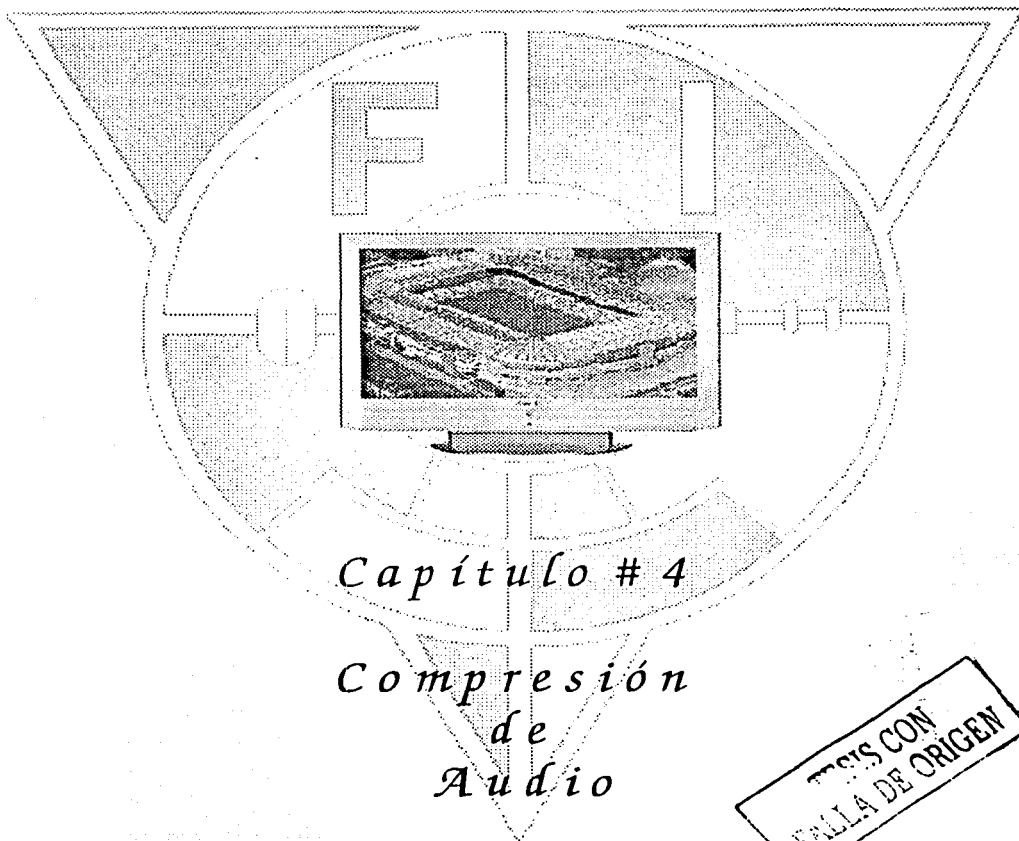
TESIS CON
 FALLA DE ORIGEN

El perfil principal se ha diseñado para una gran proporción de usos. La mayoría de las aplicaciones de broadcasting requieren el MP@ML (Main Profile at Main Level) o Perfil Principal con Nivel Principal, que es un subgrupo de MPEG que soporta la SDTV o Televisión de Definición Estándar (SDTV - Standard Definition TV). Tiene un submuestreo de las componentes de las señales de televisión de 4:2:0, una definición de 720x576 (720 píxeles x 576 líneas), una tasa de transmisión de 15 Mbps y maneja imágenes I, P y B.

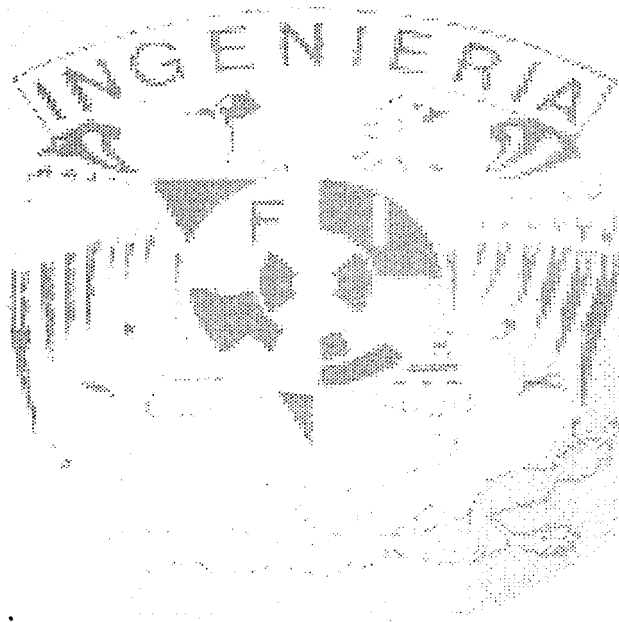
ESTA VERSIÓN NO SALE
DE LA COPIA ORIGINAL

REFERENCIAS

- [1] Libro de TOMASI, W.; "Sistemas de comunicaciones electrónicas"; Pearson Educación; Segunda Edición; México, 2001.
- [2] Libro de INGLIS, A.; "Video Engineering"; Mc Graw-Hill; E.E.U.U., 1996.
- [3] Libro de LIMANN, O.; "Fundamentos de televisión"; Marcombo; España, 1989.
- [4] Libro de PORAT, B.; "A course in Digital Signal Processing"; John Wiley & Sons; E.E.U.U., 1997.
- [5] Libro de CASTRO, A.; "Teleinformática para Ingenieros en Sistemas de Información"; Editorial Reverté; Segunda Edición; España, 1999.
- [6] Organización Internacional para la estandarización (ISO) Sitio web <http://www.iso.org/>
- [7] Libro de INGLE, V.; "Digital Signal Processing using Matlab V.4", International Thomson Publishing; E.E.U.U., 1997.
- [8] Enciclopedia gratuita en la red Wikipedia
http://www.wikipedia.org/wiki/Discrete_cosine_transform
- [9] Artículo de Tektronix llamado "A guide to MPEG fundamentals and Protocol Analysis";



MAIS CON
FALLA DE ORIGEN



TESIS CON
FALLA DE ORIGEN

4

Para una transmisión y almacenamiento más eficiente de las señales de audio, la cantidad de información requerida para representarlas debe ser reducida. En el caso de las señales de audio digital, la cantidad de información necesaria para reproducir con exactitud las muestras de la modulación por codificación de pulsos (PCM^[1]) originales pueden ser reducidas mediante la aplicación de un algoritmo de compresión digital, resultando en una representación digitalmente comprimida de la señal original. El propósito de un algoritmo de compresión digital es producir una representación digital de una señal de audio, la cual cuando sea decodificada y reproducida, sonará igual que la señal original, usando una cantidad mínima de información digital para la representación comprimida (o codificada).

Dolby AC-3 es un sistema de codificación perceptual de audio desarrollado en 1992 para que las impresiones de las películas cinematográficas de 35 mm pudieran contener audio digital multicanal además de la pista de sonido analógica óptica estándar. Desde entonces este sistema ha sido adoptado para su uso en el disco láser, la televisión de alta definición (High Definition Television – HDTV^[2]) de ATSC y la televisión digital de definición estándar (Standard Definition Television – SDTV^[3]) de ATSC, la televisión digital por cable, la transmisión digital por satélite, el Video DVD, el Audio DVD, el DVD-ROM, y la distribución de audio en Internet.

Dolby AC-3 es una tecnología flexible de compresión de datos de audio capaz de codificar varios formatos de canales de audio en una trama de bits con una velocidad baja. El algoritmo de compresión digital AC-3 puede codificar de 1 (audio monofónico) a 5.1 canales fuente de audio a partir de una representación PCM en una trama de bits serial con tasas de transmisión que van de los 32 kbps a los 640 kbps. El canal 0.1 se refiere a un canal con un ancho de banda fraccional destinado para transportar solo las señales de baja frecuencia (subwoofer). Se pueden incluir también servicios asociados de audio. Basado en una transformada por un banco de filtros y psicoacústica, AC-3 incluye las nuevas características de transmisión de una envolvente espectral con una resolución de frecuencia variable y una técnica llamada asignación de bits.

En este capítulo se explica el funcionamiento básico de los procesos de compresión y descompresión de audio utilizados por la televisión digital basada en el estándar ATSC. Se analizan los mecanismos de audición humana, así como los procesos que aprovecha la técnica de compresión de audio AC-3. También se mencionan las ventajas que tiene y los servicios de esta tecnología de audio digital. El documento de ATSC que regula todos los aspectos relacionados con la compresión de audio, es el “ATSC Standard A/52, Digital Audio Compression (AC-3)”.

4.1 Introducción a la Compresión de Audio

La compresión de audio saca ventaja de dos hechos. Primero, que en una señal de audio típica, no todas las frecuencias están simultáneamente presentes. Segundo, debido al fenómeno psicoacústico de enmascaramiento, por el cual ciertos sonidos no pueden ser

escuchados en la presencia de otros, el oído humano no puede percibir cada detalle de una señal de audio. En el proceso de compresión de audio se divide el espectro de audio en bandas mediante filtrado o transformadas, e incluye menos datos cuando se describen bandas en las que el nivel es bajo. Donde el enmascaramiento evita o reduce la audibilidad de una banda en particular, aún menos datos necesitan ser enviados.

La compresión de audio no es tan fácil de conseguir como la compresión de video debido a la agudeza del oído. El enmascaramiento solo trabaja apropiadamente cuando los sonidos que enmascaran y los enmascarados coinciden espacialmente. Otro factor que complica la compresión de audio es que las resonancias retrasadas en los altavoces de mala calidad enmascaran la compresión real. Si se prueba un compresor con unos altavoces de mala calidad da un resultado erróneo, y las señales que aparentemente son satisfactorias pueden decepcionarnos al escucharlas en un buen equipo.

La compresión de audio se basa enteramente en las características de la audición humana, lo cual debe ser considerado antes de cualquier descripción posible de la compresión. Sorprendentemente, la audición humana, particularmente en estéreo, en realidad discrimina más críticamente que la visión humana, y en consecuencia la compresión de audio debe hacerse con cuidado.

4.2 El Sistema Auditivo Humano

La audición es el resultado de una serie de procesos acústicos, mecánicos, nerviosos y mentales dentro de la combinación oído/cerebro que dan a una persona la impresión de sonido.^[4] La impresión que un humano recibe no es idéntica a la forma de onda acústica verdadera presente en el canal auditivo porque parte de la entropía de la onda se pierde. Los sistemas de compresión de audio que pierden solo esa parte de entropía que se perderá en el mecanismo de audición producirán buenos resultados.

La agudeza del oído humano es asombrosa, ya que puede detectar cantidades minúsculas de distorsión y aceptar un enorme rango dinámico. El único criterio de calidad de que se dispone consiste en el hecho de que si el oído es incapaz de detectar distorsión alguna, se dice que el sonido es perfecto. Por tanto, el criterio de calidad es completamente subjetivo y sólo se puede comprobar mediante pruebas de audición.

El mecanismo físico de la audición está formado por el oído externo, oído medio y oído interno. El oído externo consta del conducto auditivo y el tímpano. El tímpano convierte el sonido que entra en una vibración casi de la misma forma que lo hace un diafragma de un micrófono. El oído interno trabaja al detectar las vibraciones transmitidas a través de un fluido. La impedancia del fluido es mucho más alta que la del aire y el oído medio actúa como transformador para acoplar impedancias que mejora la transferencia de potencia.

Las vibraciones son transferidas hacia el oído interno por medio del estribo, el cual actúa en la ventana oval. Las vibraciones en el fluido que hay en el oído viajan hacia la cóclea (caracol), una cavidad espiral en el cráneo. La membrana basilar se extiende a través de la cóclea, la cual se encuentra representada de forma extendida en la figura 4.1.

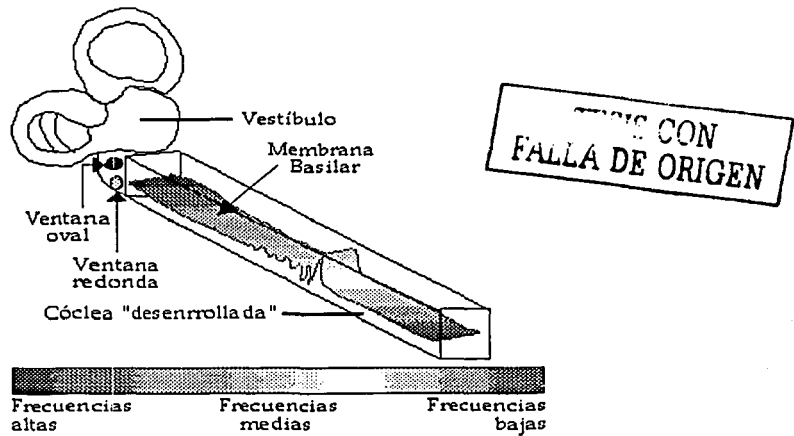


Figura 4.1 Membrana basilar "estirada".

Esta membrana varía en su masa y en su consistencia a lo largo de su extensión. En el extremo cercano a la ventana oval, la membrana es rígida y ligera, por lo que su frecuencia de resonancia es alta. En el extremo distante, la membrana es pesada y suave, y su resonancia ocurre en las frecuencias bajas. El rango de frecuencias de resonancia disponible determina el rango de frecuencias de la audición humana, que va desde los 20 Hz hasta los 20 KHz, sin embargo, en la práctica sólo llega hasta los 16 KHz aproximadamente.

Diferentes frecuencias en el sonido que entra, causa que áreas distintas de la membrana vibren. Cada área tiene terminales nerviosas diferentes para permitir que se ajuste la discriminación. La membrana basilar tiene también pequeños músculos controlados por los nervios que conjuntamente actúan como si fuera un sistema con realimentación positiva que mejora el factor de calidad de la resonancia.

El comportamiento resonante de la membrana basilar tiene su paralelo exacto con el comportamiento de un analizador de espectros. El oído analiza con bandas de frecuencia, conocidas como bandas críticas. Los anchos de bandas críticas dependen de la frecuencia, como se ilustra en la figura 4.2. Por debajo de los 500 Hz, el ancho de banda crítico es aproximadamente constante (alrededor de los 100 Hz), mientras que por encima de los 500 Hz crece en proporción a la frecuencia: el ancho de banda crítica centrada en una frecuencia superior a 500 Hz es de alrededor del 20% de la frecuencia central.

Basándose en los valores obtenidos mediante la figura 4.2, es posible subdividir el rango de frecuencias audibles en intervalos adyacentes de una banda crítica de ancho, que no se traslapan entre sí. Esta subdivisión se presenta en la figura 4.3, en el rango audible de 20 Hz a 20 KHz se encuentran 25 bandas críticas adyacentes.

TESIS CON
 FALLA DE ORIGEN

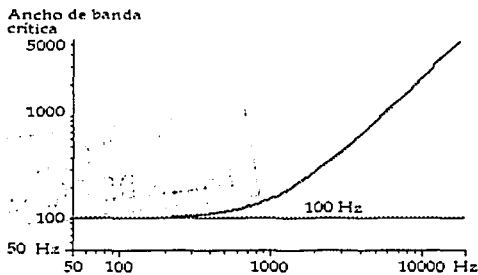


Figura 4.2 Ancho de las bandas críticas en función de la frecuencia.

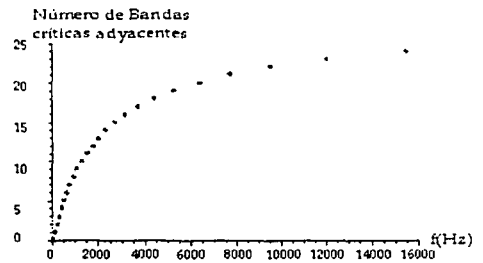


Figura 4.3 Bandas críticas adyacentes en el rango de frecuencias audibles.

De acuerdo con la incertidumbre de la teoría de las transformadas, si se conoce de forma más exacta una señal en el dominio de la frecuencia, se conoce de una forma menos exacta en el dominio del tiempo. Por consiguiente, mientras una transformada pueda discriminar mejor entre dos frecuencias, le será más difícil discriminar en el tiempo entre dos eventos. La audición humana se desarrolla con un cierto compromiso que equilibra la discriminación de incertidumbre en el tiempo y la discriminación de frecuencias; en equilibrio ninguna de las dos habilidades es perfecta.

La discriminación de frecuencias imperfecta resulta en que no se puedan separar las frecuencias muy cercanamente espaciadas. Esta incapacidad es conocida como enmascaramiento auditivo, que se define como una sensibilidad reducida hacia un sonido en la presencia de otro. El oído es incapaz de registrar energía en algunas bandas cuando existe más energía en otra banda cercana. La vibración de la membrana en sintonía con una sola frecuencia no puede ser localizada en una zona infinitamente pequeña, por lo que las zonas cercanas se ven obligadas a vibrar a la misma frecuencia con una amplitud que decrece con la distancia. Otras frecuencias son excluidas a menos que la amplitud sea lo bastante alta como para dominar la vibración local de la membrana.

Entonces, el enmascaramiento es la reducción total o parcial de la sensibilidad de un oyente para percibir un determinado sonido, provocado por la presencia simultánea de otro. Cuando un sonido hace que otro sea menos audible, porque ambos se producen al mismo tiempo, se dice que se produjo un fenómeno de enmascaramiento. El sonido cuyo umbral de audibilidad se ha modificado se denomina sonido enmascarado y al otro, sonido que enmascara.

Existen algunos lineamientos que rigen el enmascaramiento:

- Un sonido posee mayor poder para enmascarar, si se intenta enmascarar a otro que tenga una frecuencia parecida. En cambio, resulta muy difícil de enmascarar con otro de frecuencia diferente, alejada en el espectro. El enmascaramiento podrá realizarse, pero el nivel de presión sonora necesario tendrá que ser más importante que en el primer caso.

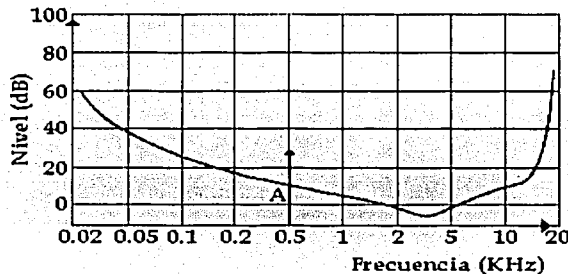
- Un sonido de determinada frecuencia tiene más poder para enmascarar otro de frecuencia más aguda (frecuencia mayor), que sobre otro de frecuencia más grave (frecuencia menor). Por lo tanto, si se está buscando el efecto de enmascaramiento, es más fácil lograrlo con un sonido que enmascara de frecuencia más grave que la del enmascarado.

El umbral de audición es una función de la frecuencia. No sorprende que la sensibilidad más grande esté en el rango del habla. En el campo auditivo el rango de frecuencia audible va de los 20 Hz a los 16 KHz (en la práctica), pero el oído no es igualmente sensible a todas estas frecuencias. Las más audibles son las ubicadas en el medio del espectro, aproximadamente entre 1 KHz y 5 KHz. El oído es menos sensible para frecuencias bajas y altas. Esta característica de menor agudeza para los tonos graves favorece el enmascaramiento de los sonidos que produce el cuerpo humano.

Se define como umbral de audición, para un tono puro de una frecuencia dada, a la mínima presión sonora eficaz que puede ser oída, en ausencia de todo ruido de fondo. El umbral auditivo representa la presión sonora mínima que produce la sensación de audición. La curva de sensibilidad que representa el umbral de audición de una señal en función de su frecuencia en ausencia de señal perturbadora se representa en la figura 4.4, donde la señal A es audible, puesto que sobrepasa el umbral de percepción.

En la presencia de un solo tono, el umbral se modifica. Hay que notar que el umbral es elevado para los tonos en las frecuencias más altas y en cierta porción de las frecuencias más bajas. En la presencia de un espectro de entrada complejo, tal como la música, el umbral es elevado en casi todas las frecuencias.

Una consecuencia de este comportamiento es que el silbido de un cassette analógico solo se escucha durante los pasajes de música silenciosa. La compansión (compresión/expansión) hace uso de este principio al amplificar las señales de audio con bajos niveles antes de grabarlas o transmitir las, y después regresarlas a su nivel correcto.



TESIS CON FALLA DE ORIGEN

Figura 4.4 Umbral de audición en función de la frecuencia.

Más recientemente, se ha observado que esta curva se ve afectada en presencia de múltiples señales: por ejemplo, en el caso de dos señales de frecuencias relativamente cercanas, la señal más fuerte hace subir el umbral de audición en sus proximidades, cuyo efecto es disminuir la sensibilidad del oído alrededor de estas frecuencias.

La figura 4.5 representa este caso, donde la señal A, antes audible, es ahora enmascarada por la cercana señal B, más potente que A. Este efecto recibe el nombre de enmascaramiento frecuencial.

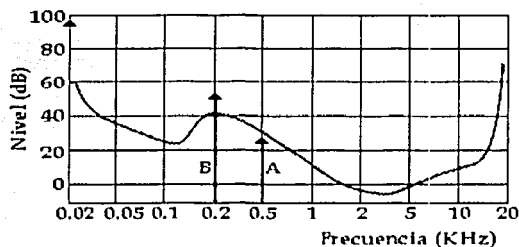


Figura 4.5 Enmascaramiento frecuencial (A enmascarado por B).

También existe un efecto de enmascaramiento temporal: un sonido de elevada amplitud enmascara igualmente los sonidos más débiles inmediatamente anteriores o posteriores. El pre-enmascaramiento y el post-enmascaramiento ocurre cuando el sonido que enmascara sigue enmascarando sonidos que están en niveles más bajos antes y después de la duración real del sonido que enmascara. Ver la figura 4.6.

La discriminación temporal imperfecta del oído se debe a su respuesta resonante. El factor Q es tal que un sonido tiene que estar presente por lo menos cerca de 1 milisegundo antes de que llegue a ser audible. Debido a su respuesta lenta, el enmascaramiento puede llevarse a cabo aún cuando las dos señales envueltas no sean simultáneas.



Figura 4.6 Enmascaramiento temporal.

El enmascaramiento aumenta el umbral de audición, y los compresores toman ventaja de este efecto aumentando el ruido de fondo, el cual permite expresar a la onda de audio con menos bits. El ruido del fondo solo puede ser aumentado en las frecuencias en las cuales hay un enmascaramiento efectivo. Para maximizar el efecto de enmascaramiento, es

TESIS COM
FALLA DE ORIGEN

necesario separar el espectro de audio en bandas de diferentes frecuencias que permitan la introducción de diferentes cantidades de compansión y ruido en cada banda.

4.3 Procesos de Compresión de Audio

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

Existen muchos tipos diferentes de compresión de audio y cada uno permite un factor de compresión diferente. En este apartado analizaremos la codificación de la sub-banda y la codificación por transformación.

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

En la codificación por transformación la forma de onda de audio en el dominio del tiempo es convertida a una representación en el dominio de la frecuencia, como la Transformada de Fourier^[5], la Transformada Discreta de Fourier^[6,7,8], o la Transformada Discreta del Coseno^[7,9]. La codificación por transformación toma ventaja del hecho de que la amplitud o envolvente de una señal de audio cambia relativamente despacio, con lo que los coeficientes de la transformada pueden transmitirse relativamente con poca frecuencia. Esta aproximación se estropea en presencia de transitorios y se requieren en la práctica sistemas adaptables. Los transitorios causan que los coeficientes se actualicen de forma mas frecuente mientras que en las partes estacionarias de la señal como las notas sostenidas la tasa de actualización puede reducirse.

4.3.1 Codificación de la sub-banda

La compresión de los datos de la sub-banda aprovecha el hecho de que los sonidos reales no tienen una energía espectral uniforme. La longitud de la palabra del audio PCM está basada en el rango dinámico requerido. Cuando una señal con un espectro no uniforme es transmitida por PCM, todo el rango dinámico es ocupado únicamente por la componente espectral más grande, y todas las demás componentes son codificadas con excesivo headroom (área entre el nivel normal de funcionamiento y el nivel de recorte). En su forma más simple, la codificación de la sub-banda funciona dividiendo la señal de audio en un número de bandas de frecuencia y comprimiendo y expandiendo cada banda de acuerdo con su propio nivel. Las bandas en las que hay poca energía dan como resultado amplitudes pequeñas que pueden transmitirse con una longitud de palabra corta. Por tanto, cada banda

se traduce en muestras de longitud variable, pero la suma de todas las longitudes de palabra de la muestra es inferior a las de la PCM, pudiendo obtenerse así una ganancia de codificación.

El número de sub-bandas que pueden utilizarse depende de qué otra técnica de compresión se vaya a combinar con la codificación de la sub-banda. Si se tiene la intención de utilizar la compresión basada en el enmascaramiento auditivo, es preferible que las sub-bandas sean más estrechas que las bandas críticas del oído y, por tanto, se requerirá un gran número. La figura 4.7 muestra la condición crítica en la que el tono del enmascaramiento se encuentra en el límite superior de la sub-banda. Se observará que cuanto más estrecha es la sub-banda, mayor es el ruido de recuantificación que puede enmascarse. No obstante, la utilización de un número excesivo de sub-bandas acentúa la complejidad y el retardo de codificación, y también produce pre-eco en los transitorios, que excede el enmascaramiento temporal.

Las figuras 4.8 y 4.9 muestran los diagramas de bloques de un codificador y un decodificador de sub-bandas, respectivamente. En la entrada, el rango de frecuencias es dividido en sub-bandas mediante un banco de filtros. Los datos descompuestos de la sub-banda se organizan en bloques de tamaño fijo, antes del proceso de reducción. Aunque todas las sub-bandas pueden utilizar bloques de la misma longitud, algunos codificadores pueden utilizar bloques que se hacen más largos a medida que disminuye la frecuencia de la sub-banda. Los bloques de las sub-bandas también se denominan bins de frecuencia.

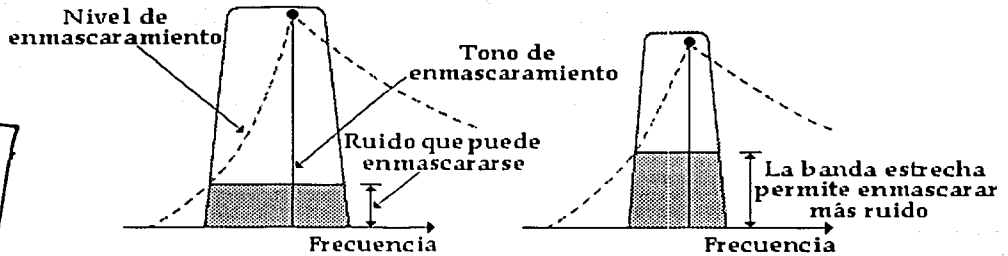


Figura 4.7 En la codificación de la sub-banda, el caso más desfavorable tiene lugar cuando el tono de enmascaramiento se encuentra en el límite superior de la sub-banda

La ganancia de codificación se obtiene cuando la forma de onda de cada banda pasa a través de un recuantificador. La recuantificación se consigue multiplicando los valores de las muestras por una constante y redondeando el resultado por exceso o por defecto de acuerdo con la longitud de palabra requerida. Cuanto menor es la longitud de palabra, mayor es la ganancia de codificación, pero más toscos resultan los escalones de cuantificación y, por tanto, el nivel de error de cuantificación

TESIS CON FALLA DE ORIGEN

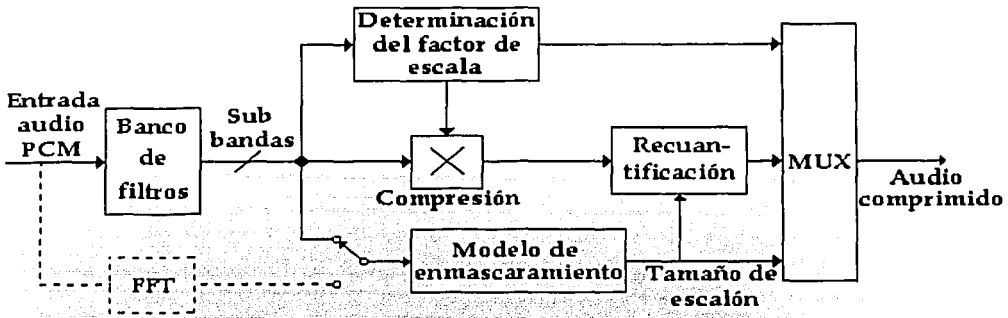


Figura 4.8 Diagrama de bloques de un codificador de sub-bandas

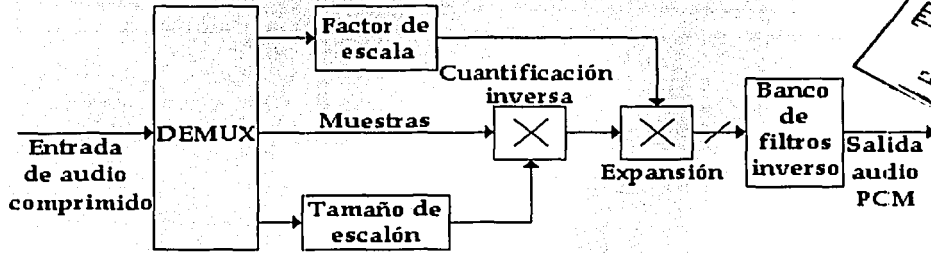


Figura 4.9 Diagrama de bloques de un decodificador de sub-bandas

4.3.2 Codificación por Transformación

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

Las transformadas prácticas requieren bloques de muestras en lugar de cadenas interminables. De ahí que se divida la forma de onda en cortos segmentos traslapados y, después se transforma cada uno de ellos individualmente tal como indica la figura 4.10. De este modo, cada muestra de entrada aparece en sólo dos transformadas, pero con una ponderación variable dependiendo de su posición en el eje temporal.

La DFT (Discrete Frequency Transform o Transformada de Frecuencia Discreta) requiere gran número de cálculos, debido al requisito de tener que utilizar una aritmética compleja para obtener la fase de las componentes, así como la amplitud. Una alternativa consiste en emplear la Transformada Discreta del Coseno (DCT). Esta presenta una ventaja cuando se utiliza con ventanas traslapadas.

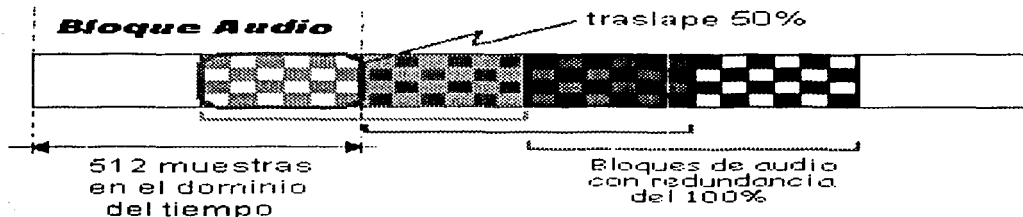


Figura 4.10 La entrada PCM es dividida en bloques traslapados en el codificador AC-3

En la Transformada Discreta del Coseno Modificada (MDCT)^[10], se usan ventanas con un traslape del 50%. De este modo, se obtiene el doble de coeficientes necesarios, que se submuestrean por un factor de dos para obtener una transformada muestreada críticamente, lo cual tiene como resultado un efecto de aliasing^[2,11] en el dominio de la frecuencia. Sin embargo, variando levemente la transformada, los productos de aliasing en la segunda mitad de una determinada ventana son iguales en tamaño, pero de polaridad opuesta a los productos de aliasing de la primera mitad de la siguiente ventana, por lo que se eliminarán en su reconstrucción. Éste es el principio de la eliminación del aliasing en el dominio temporal (TDAC, Time Domain Aliasing Cancellation)^[10].

La recuantificación realizada en el codificador eleva el ruido de cuantificación en el bin de la frecuencia durante todo el tiempo que dura el bloque. La figura 4.11 muestra que, si se produce un transitorio hacia el extremo final de un bloque, el decodificador reproduce la forma de onda correctamente, pero el ruido de cuantificación comenzará al principio del bloque y puede dar lugar a un pre-eco, en el que el ruido se oye antes que el transitorio.

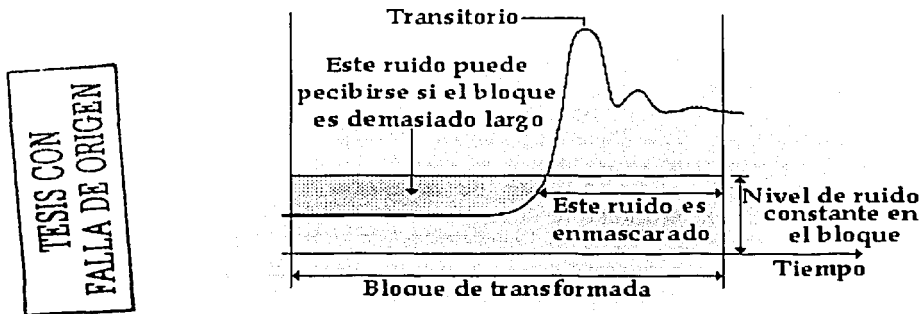


Figura 4.11 Transitorio en el final de un bloque de una transformada

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

4.4 Introducción a la técnica de compresión de audio AC-3

4.4.1 Desarrollo de AC-3

El origen de la tecnología AC-3^[12] vino del deseo de proporcionar una mejor calidad del sonido multicanal para la Televisión de Alta Definición. En los Estados Unidos, el proceso de estandarización de la Televisión de Alta Definición comenzó formalmente en 1987, con la creación, a través de la Comisión Federal de Comunicaciones (Federal Communications Commission – FCC), del Comité Consultivo para el Servicio de Televisión Avanzada (Advisory Committee on Advanced Television Service – ACATS). Las propuestas iniciales del sistema de HDTV se inclinaban hacia una transmisión de imagen analógica y una transmisión de sonido digital.

En 1990 nace el concepto de AC-3: un codificador de audio multicanal, operando en la misma tasa de transmisión de bits requerida por dos canales de audio codificados independientemente, ofreciendo al mismo tiempo desempeño multicanal. Si tenía éxito, el concepto de AC-3 permitiría que HDTV ganara en la transmisión discreta de audio, sin pagar el precio correspondiente de duplicar la tasa de transmisión de bits requerida.

Mientras que fue concebido como una solución para las necesidades de HDTV, el concepto de AC-3 se hizo realidad en respuesta a una necesidad similar del cine. En 1989 un subgrupo de SMPTE (Society of Motion Picture and Television Engineers – Sociedad de Ingenieros en Películas Cinematográficas y Televisión) estudió el asunto de cuantos canales de audio debería ofrecer el nuevo sistema de sonido digital. La conclusión fue que se proporcionarían 5.1 canales (izquierdo, central, derecho, surround izquierdo, surround derecho, subwoofer), el cual es idéntico a la organización del formato surround de 70 mm que se ha estado usando en el cine desde 1979. Para hacer que los datos de sonido digital en la cinta sea fiable, y que no interfiera ni con el área de la película ni con el área del sonido analógico, la tasa de transmisión disponible es limitada. Se determinó que 320 kbps de datos de audio con corrección de errores colocados y extraídos desde el área que hay entre las perforaciones en un lado de la película de 35 mm podrían ser fiables. Todo lo que se requería era llegar a entender el concepto de AC-3 como un codificador de 5.1 canales de audio operando a 320 kbps.

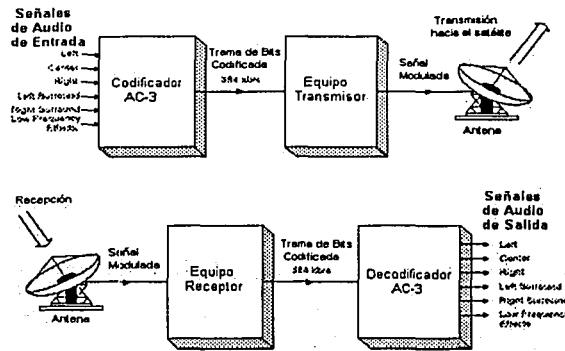
Debido a que el plan requerido para un producto de cine comercial debe tener un rápido desarrollo, AC-3 fue implementado primeramente como un sistema que trabajaba en tiempo real en múltiples tarjetas DSP. Las primeras demostraciones en la industria cinematográfica comenzaron en Mayo de 1991. Por Diciembre de 1991, la primera película digital

codificada por AC-3, *Star Trek VI*, se exhibió en tres cines. El lanzamiento formal del Sistema Dolby SRD fue en Junio de 1992, con la realización de la película *Batman Returns*.

A mediados de 1991 la existencia del sistema de codificación de audio se divulgó públicamente, y fue abrazado rápidamente por la comunidad de audio de HDTV en Estados Unidos. El grupo de trabajo 10-1 de la UIT (La UIT, Unión Internacional de Telecomunicaciones) BR se reunió en Junio de 1991 y aceptó el formato básico de 5 canales de audio, tomándolo como base para una recomendación. En Febrero de 1992 el Comité de Sistemas de Televisión Avanzada (Advanced Television Systems Committee) realizaron un documento que formalmente recomendaba la codificación de 5.1 canales de audio para el servicio de HDTV en los Estados Unidos. En Octubre de 1992, el GT 10-1 aceptó el canal 0.1 de frecuencia baja y modificaron su recomendación. En 1993 se le realizaron pruebas subjetivas al sistema AC-3 en los Estados Unidos para evaluar que tan conveniente era incluirlo en el sistema de HDTV que se estaba proponiendo por la Gran Alianza (un consorcio de los propulsores de HDTV en E.E.U.U. que han sido autorizados para colaborar en el sistema de transmisión de HDVT en E.E.U.U.). En Octubre de 1993 la Gran Alianza recomendó el uso de Dolby AC-3. En Noviembre del mismo año. Todo el comité A.CATS aprobó formalmente el uso de AC-3 para el sistema de HDTV de la Gran Alianza.

La intención de Dolby era hacer que AC-3 fuera aplicado universalmente para los codificadores de baja tasa de transmisión mediante satisfacer diversos requerimientos. Esto permite que AC-3 libere una señal de audio en una forma útil para toda la audiencia, aun cuando cada miembro de esta tenga distintas necesidades. Por ejemplo mientras los miembros más críticos de la audiencia pueden tener un sistema de reproducción multicanal, la mayor parte de la audiencia solo podrá escuchar en mono o en estéreo.

Una aplicación típica del algoritmo se muestra en la figura 4.12. En este ejemplo, un programa de audio de 5.1 canales se transforma de una representación PCM que requiere más de 5 Mbps (6 canales * 48 kHz * 18 bits = 5.184 Mbps) en una trama de bits serial de 384 kbps por medio de un codificador AC-3. El equipo de transmisión por satélite convierte esta trama de bits a una transmisión RF la cual es dirigida al transpondedor del satélite. La cantidad de ancho de banda y potencia requeridas para la transmisión han sido reducidas por un factor de más de 13 gracias a la compresión digital AC-3. La señal recibida desde el satélite es demodulada de nuevo en una trama de bits serial de 384 kbps, y decodificada por el decodificador AC-3. El resultado es el programa de audio original de 5.1 canales.



TESIS CON FALLA DE ORIGEN

Figura 4.12 Ejemplo de una aplicación de audio AC-3 transmitida por satélite

La compresión de audio digital es útil dondequiera que se obtenga beneficio económico, al reducir la cantidad de información digital requerida para representar el audio.

4.4.2 AC-3, sus 5.1 Canales de Audio y los Servicios de Audio

AC-3 divide el espectro de audio en bandas de frecuencia angostas usando modelos matemáticos derivados de las características del oído, y analiza cada banda para determinar la audibilidad de dichas señales. Para maximizar la eficiencia de datos, el número de bits más grande representa las señales más audibles y las señales con menos bits representan las señales menos audibles. Para determinar la audibilidad de las señales, el sistema realiza lo que se conoce como enmascaramiento. El enmascaramiento se refiere al fenómeno que ocurre cuando el oído no puede detectar sonidos de niveles más bajos si hay sonidos de niveles más altos en las frecuencias cercanas. Cuando esto ocurre, los sonidos de niveles altos enmascaran a los niveles bajos, lo cual se traduce en que se hagan inaudibles. Explotar este fenómeno permite que el audio sea codificado de forma más eficientemente que en otros sistemas de codificación. Esto hace que AC-3 sea una excelente opción para los sistemas donde se desea una alta calidad de audio, pero el ancho de banda o el espacio de almacenamiento se encuentra restringido. Esto es especialmente cierto para las bandas sonoras multicanal.

Los 5.1 canales de audio típicamente consisten en cinco canales principales discretos de rango completo (Izquierdo, Central, Derecho, Surround Izquierdo y Surround Derecho) más un canal opcional limitado en banda de Efectos de Baja Frecuencia (LFE - Low Frequency Effects) para las frecuencias bajas añadidas (el 0.1). Las tramas de Dolby Digital liberan canales principales con un ancho de banda completo de frecuencia, desde 3 Hz hasta 20 kHz, y un canal LFE limitado en frecuencia, que va de 3 Hz a 120 Hz. Todos los programas multicanal transportados dentro de la trama de bits de AC-3, pueden ser mezclados para lograr una compatibilidad con los sistemas estéreo, mono o Dolby Surround.

❖ **Distribución de los canales de audio**

Al tratarse de un sistema multicanal de audio, es necesario identificar y ordenar la codificación de los distintos canales de audio. En la figura 4.13 se muestra la distribución de un sistema de 5.1 canales de audio a la hora de la recepción. Los canales se muestran por sus siglas en inglés (L = left, C = center, R = right, S = surround, LFE = Low Frequency Effects).

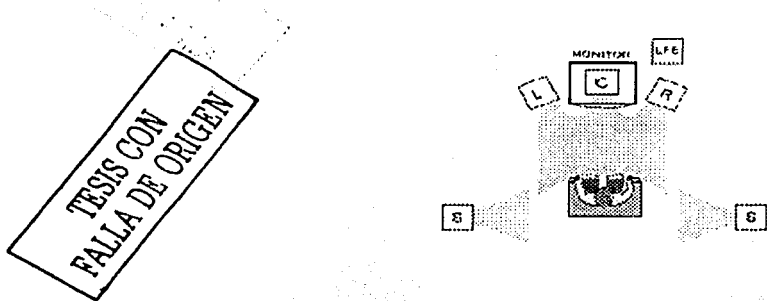


Figura 4.13 Distribución de los canales de audio en el estándar AC-3

Es posible recibir un número de canales de audio de acuerdo a las capacidades del equipo receptor. En la tabla 4.1 se muestran los posibles formatos de los canales de audio.

No. de canales	Orden del arreglo de canales
2	Dos canales cualesquiera de los 5 canales principales
1	Central
2	Izquierdo, Derecho
3	Izquierdo, Central, Derecho
3	Izquierdo, Derecho, Surround Derecho
4	Izquierdo, Central, Derecho, Surround Derecho
4	Izquierdo, Derecho, Surround Izquierdo y Surround Derecho
5	Los 5 canales principales

Tabla 4.1 Modos de codificación de los canales de audio

Existe una variedad de distribuciones de las salas de reproducción de 5.1 canales de audio. Todos requieren un espacio que pueda permitir la colocación de cinco altavoces de rango completo (20 Hz – 20 kHz) alrededor de la posición de los usuarios y la colocación de un altavoz adicional para la reproducción del canal LFE (Low-Frequency Effects), que va de los 3 Hz a los 120 Hz, en el espacio destinado. Una distribución de una sala típica para la producción del audio multicanal tendrá los altavoces Izquierdo, Central y Derecho

colocados al frente, mientras los dos o más altavoces surround se necesita que estén detrás del usuario.

Las especificaciones de la UIT-R para la distribución de una sala diseñada para escuchar programas multicanales se muestra en la figura 4.14. Aparte del alineamiento de las señales, se describe una geometría específica: el altavoz central está directamente enfrente, la posición de los altavoces L y R están a 30° del centro (formando un ángulo de 60° entre sí) y los altavoces surround están a 110° del centro.

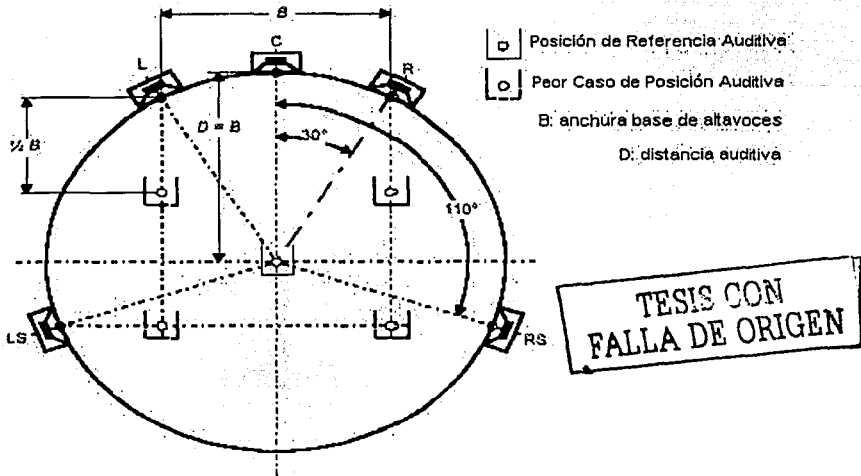


Figura 4.14 Distribución de una sala para escuchar programas multicanal recomendada por la UIT-R

❖ **Servicios de Audio disponibles en AC-3^[13]**

Por medio de esta técnica de codificación existe la posibilidad de tener servicios asociados de audio. La trama elemental de audio de AC-3 solo contiene los datos de un servicio único de audio. Es posible agregar otros servicios de audio mediante la multicanalización de varias tramas. Existen dos servicios principales de audio y seis servicios asociados de audio. En forma general, las tramas de audio de un programa de televisión estarán conformados por un servicio principal de audio y un servicio asociado, el cual complementa la mezcla de los canales de audio. Otra posibilidad es un servicio principal de audio combinado con servicios asociados.

Servicios principales de audio:

Servicio principal de audio completo (Complete Main Audio Service—CM). El servicio principal de audio completo tiene la información de un programa de audio completo. Es el tipo de servicio de audio que normalmente es transmitido. Puede contener desde 1 canal de audio hasta 5.1 canales.

Servicio principal de audio, música y efectos (Main Audio Service, Music and Effects—ME). Este servicio principal contiene el audio de música y efectos de un programa sin incluir el diálogo principal. Puede contener desde 1 canal hasta 5.1 canales de audio.

Servicios asociados de audio:

Vista disminuida (Visually Impaired—VI). Este servicio contiene una descripción narrativa del contenido del programa. Es posible tener la reproducción del servicio de vista disminuida y del servicio de audio principal. El control del servicio de audio principal lo tendrá el servicio de vista disminuida para asegurar que la narración sea inteligible.

Nivel auditivo disminuido (Hearing Impaired—HI). Contiene sólo un diálogo que puede ser reproducido simultáneamente con el servicio principal de audio completo. El diálogo que contiene el servicio de nivel auditivo disminuido está procesado de tal forma que sea inteligible para personas que tienen afectada su sensibilidad auditiva. Es posible que este servicio contenga una combinación de música y efectos, que también han sido procesados.

Diálogo (Dialogue—D). Este servicio está formado por un diálogo de programa, el cual puede ser reproducido simultáneamente con el servicio principal de audio, música y efectos.

Comentario (Commentary—C). Contiene un diálogo que puede ser reproducido con un servicio principal de audio, el carácter de este servicio es opcional. El control de nivel de audio se encuentra a cargo de este servicio, para que la información que transmite sea escuchada correctamente.

Emergencia (Emergency—E). Este servicio asociado transporta información sobre anuncios de emergencia o de alerta para la población.

Voz adicional (Voice-Over—VO). Es un servicio de audio asociado que transmite una señal de voz externa a la señal del programa, el cual puede ser codificado de manera independiente por el equipo transmisor, es decir, que es un servicio de audio que puede ser agregado a una trama de bits ya codificada por el proceso de compresión.

4.5 Compresión de Audio Digital mediante AC-3

AC-3 es una tecnología de compresión de datos de audio flexible capaz de codificar una variedad de formatos de canales de audio en una sola trama de bits de baja velocidad. Son

soportadas ocho configuraciones de canales las cuales van desde la convencional mono ó estéreo hasta el formato surround con seis canales discretos (Ver la tabla 4.1). La especificación de trama de bits de AC-3 permite frecuencias de muestreo de 48 kHz, 44.1 kHz o 32 kHz, y soporta tasas de transmisión desde 32 kbps hasta 640 kbps.

El diagrama de bloques de un codificador AC-3 se muestra en la figura 4.15. La codificación se efectúa en el dominio de la frecuencia, usando una transformada del coseno discreta modificada (modified discrete cosine transform – MDCT) de 512 puntos con un traslape del 50%. En el caso de señales transitorias, se obtiene un mejor desempeño al usar una técnica de switcheo de bloques, en la que se calculan dos transformadas de 256 puntos, en lugar de una de 512 puntos. Un proceso de conversión del punto flotante descompone los coeficientes de la transformada en pares de exponente y mantisa. Las mantisas son entonces cuantizadas con un número variable de bits, basándose en un modelo de asignación paramétrica de bits.

El modelo de asignación de bits de AC-3 utiliza los principios de enmascaramiento psicoacústico para decidir cuantos bits se le asignarán a cada mantisa en una banda de frecuencias dada. Dependiendo del grado de enmascaramiento, algunas mantisas recibirán muy pocos bits o incluso ninguno. Esto reduce el número de bits necesarios para representar a la fuente, a expensas de ruido añadido inaudible.

En AC-3 se toma una aproximación paramétrica, en la que el codificador construye su modelo de enmascaramiento basándose en los exponentes de los coeficientes de la transformada y unos pocos parámetros clave dependientes de la señal. Estos parámetros se envían hacia el decodificador en la trama de bits, usando mucho menos bits de los que serían necesarios para transmitir los primeros valores de bits de la asignación. En el decodificador, la asignación de bits se reconstruye basándose en los exponentes y los parámetros de asignación de bits.

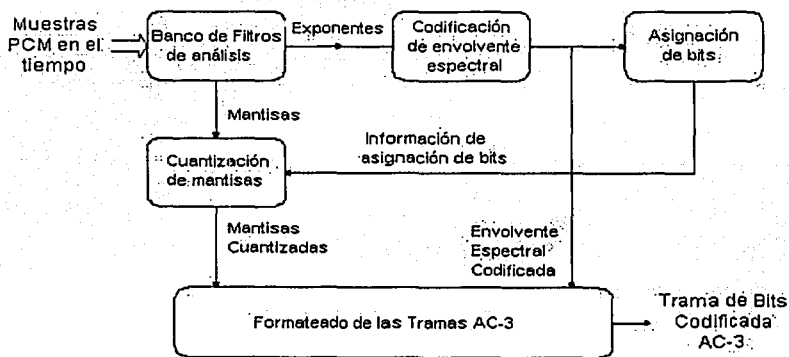


Figura 4.15 Diagrama de bloques de la compresión AC-3

La eficiencia de codificación de AC-3 mejora si el número de canales fuente aumenta. Esto se debe a dos rasgos principales: una unión global de bits y un acoplamiento de alta frecuencia. La técnica de unión global de bits permite al asignador de bits descomponer los bits disponibles entre los canales de audio según sea necesario. Si uno o más canales se encuentran inactivos en un instante de tiempo específico, los canales restantes recibirán más bits de los que recibirían si todos los canales estuvieran con una alta demanda de bits. El acoplamiento se utiliza para una mayor compresión de la tasa de transmisión de datos necesaria para el rango de las altas frecuencias del espectro de audio. En las frecuencias altas, el oído no puede detectar los ciclos individuales de la forma de la señal de audio, y en cambio responde solo a su envolvente. El acoplamiento reduce las componentes de alta frecuencia de los canales correlacionados a un solo canal acoplado, y entonces genera una cadena adicional de datos que describe la envolvente espectral por canal.

El codificador AC-3 acepta audio PCM y produce una trama de bits codificados de acuerdo con el estándar A/52 de ATSC. El proceso de codificación se describe brevemente a continuación: El primer paso en el proceso de codificación es el transformar la representación de audio de una secuencia de muestras PCM temporales a una secuencia de bloques de coeficientes de frecuencia. Esto se hace en el banco de filtros de análisis. Los bloques traslapados de 512 muestras temporales son multiplicados por una ventana temporal y transformados al dominio de la frecuencia. Debido a los bloques de traslape, cada muestra PCM de entrada es representada en dos bloques transformados secuencialmente. La representación en el dominio de la frecuencia debe entonces ser dividida por un factor de dos de tal forma que cada bloque contenga 256 coeficientes de frecuencia. Los coeficientes de frecuencia individuales son representados en una notación exponencial binaria como un exponente y una mantisa binaria. Los grupos de exponentes son codificados en una representación tosca del espectro de la señal, el cual es llamado envolvente espectral. Esta envolvente espectral es utilizada por la rutina de asignación de bits la cual determina cuántos bits usar para codificar cada mantisa individual. La envolvente espectral y las mantisas cuantizadas toscamente de 6 bloques de audio (1536 muestras de audio por canal) conforman un frame AC-3. La trama de bits de AC-3 es una secuencia de frames AC-3.

El codificador AC-3 real es más complicado que el indicado en la figura 4.15. Las funciones siguientes que no se mostraron anteriormente también están incluidas:

- Se asigna un encabezado de frame el cual contiene información (tasa de bits, frecuencia de muestreo, número de canales codificados, etc.) requerida para sincronizar y decodificar la trama de bits codificados.
- Códigos de detección de error que se insertan para permitir que el decodificador verifique que el frame de datos recibidos no tiene errores.
- La resolución espectral del banco de filtros de análisis puede ser dinámicamente alterada para adecuarlas a las características en tiempo y frecuencia de cada bloque de audio.
- La envolvente espectral puede ser codificada con una resolución en tiempo y frecuencia variables.

- Se puede ejecutar una asignación de bits más complicada, y modificar los parámetros de la rutina de asignación para producir una asignación de bits más óptima.
- Los canales pueden ser acoplados juntos en las altas frecuencias para alcanzar una ganancia de codificación mayor para el funcionamiento a tasas de transmisión más bajas.
- En el modo de dos canales, puede ser ejecutado selectivamente un proceso de "rematrizado" para proporcionar ganancia de codificación adicional, y permitir que mejoren los resultados que se obtengan si la señal de dos canales es decodificada con un decodificador surround matricial.

TESIS CON FALLA DE ORIGEN

4.6 Bloques principales del codificador AC-3

AC-3 es fundamentalmente un codificador adaptativo basado en la transformada y que utiliza un banco de filtros críticamente muestreado en una frecuencia lineal basado en la técnica de cancelación en el dominio del tiempo (TDAC – Time Domain Alias Cancellation) de Princen-Bradley. La ventaja primordial de la codificación basada en un banco de filtros es que los componentes de la señal y sus correspondientes componentes de ruido de cuantización (destinadas a ser enmascaradas) se mantienen dentro de un ancho de banda crítico de algún otro filtro, derivando el máximo beneficio de las características de enmascaramiento del oído humano, y minimizando la tasa de datos resultante necesaria para una codificación con una percepción libre de ruidos.

En la presencia de transitorios (de banda ancha), la longitud del bloque transformado por el banco de filtros puede ser reducido dinámicamente para contener el ruido de cuantización en una pequeña región temporal cerca del transitorio.

Una técnica de asignación de bits flexible aprovecha la multiplicidad de canales mediante asignar bits a través de las frecuencias y canales según sea necesario, desde un depósito de bits común, mientras toma en cuenta los efectos de enmascaramiento intra y entre canal. La ganancia de codificación es realizada según sea necesaria mediante separar y codificar independientemente la portadora de alta frecuencia y la envolvente de la información.

A continuación se describen los pasos principales del procesamiento de codificación con mayor detalle. Véase la figura 4.16, donde se muestra un diagrama de bloques de codificación AC-3 desglosado.

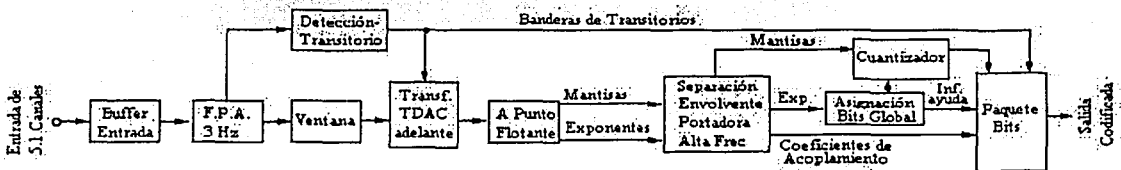


Figura 4.16 Diagrama desglosado de la técnica de codificación AC-3

4.6.1 Señal Digital de Entrada

El codificador AC-3 acepta audio PCM y produce una trama de bits codificados de acuerdo con el estándar A/52 de ATSC.^[14] Los codificadores de AC-3 actuales aceptan longitudes de palabras de 16, 18 o 20 bits a frecuencias de muestreo de 32, 44.1 o 48 kHz.

La frecuencia de muestreo de entrada debe estar ajustada a la frecuencia de muestreo de salida de tal forma que cada frame de sincronía de AC-3 contenga 1536 muestras por canal. Si la entrada de audio está disponible en un formato PCM con una frecuencia de muestreo distinto del requerido, se debe hacer una conversión de la frecuencia de muestreo para adecuarla a la frecuencia de muestreo.

4.6.2 Buffer de entrada

AC-3 es un codificador estructurado que trabaja por bloques, por lo tanto uno o más de los bloques de las señales en el dominio de tiempo, típicamente estos bloques constan de 512 muestras, son recogidos en un buffer de entrada para cada canal antes de proceder a realizar el procesamiento.

4.6.3 Filtrado en la entrada

Las señales de entrada del codificador son filtradas con un filtro paso altas^[1,15] individualmente con una frecuencia de corte de 3 Hz, para eliminar la componente de directa. La señal de Subwoofer es filtrada con un filtro paso bajas^[1,15] con una frecuencia de corte de 120 Hz.

4.6.4 Detección de transitorios

Antes de transformar la señal de audio del dominio del tiempo al dominio de la frecuencia, el codificador realiza un análisis de la naturaleza espectral y/o temporal de la señal de entrada y selecciona la longitud del bloque apropiada. Este análisis ocurre solo en el codificador. Las señales filtradas con un ancho de banda completo son analizadas con un filtro paso altas, para detectar la presencia de transitorios. Esta información es utilizada para ajustar el tamaño del bloque del banco de filtros TDAC, restringiendo el ruido de cuantización asociado con el transitorio dentro de una región temporal pequeña cercana al transitorio, evitando un desenmascaramiento temporal. El detector de transitorios es utilizado para determinar cuando cambiar de un bloque de transformada largo (con un tamaño de 512), a un bloque corto (con un tamaño de 256).

Una longitud de transformada grande es más adecuada para las señales de entrada cuyo espectro permanece estacionario, o varía lentamente en el tiempo. Una longitud de transformada grande proporciona una resolución frecuencial más grande mejorando el desempeño de codificación para dichas señales. Por otra parte, una transformada corta, posee una resolución temporal mayor, y se usa para señales que cambian rápidamente en el

tiempo. Por lo tanto, se debe hacer un compromiso entre la resolución en tiempo y frecuencia, al seleccionar la longitud de un bloque.

AC-3 adapta su resolución frecuencial/temporal de la transformada dependiendo de las características espectrales y temporales de la señal que está siendo procesada. Esta solución es muy parecida al comportamiento de la audición humana. En un código de transformada, la adaptación ocurre mediante switchear la longitud del bloque dependiendo de la señal.

4.6.5 Banco de Filtros TDAC y Transformada directa^[14]

La Transformada Discreta Modificada del Coseno (MDCT) es una transformada en frecuencia basada en la Transformada Discreta del Coseno de Tipo IV (DCT-IV), con la propiedad adicional de estar traslapada: ha sido diseñada para utilizarse en bloques consecutivos de grupos de datos más grandes, donde los bloques subsecuentes tienen un traslape de 50%. Este traslape, junto con las características de compresión de energía de la DCT, la hacen atractiva espacialmente para las aplicaciones de compresión de señales, debido a que ayuda a evitar errores en las fronteras de los bloques.

En AC-3, la MDCT no se aplica directamente a la señal de audio, sino a la salida de un banco de filtros de 50 bandas, con el objetivo de reducir el "aliasing". Una combinación de un banco de filtros con una MDCT se conoce como banco de filtros híbrido o MDCT por sub-bandas.

La elección de emplear un banco de filtros de análisis/síntesis en un codificador de audio es un trueque entre la resolución frecuencial, la resolución temporal, y el costo, donde el costo es medido en bits de memoria RAM y ciclos múltiples/acumulados. El beneficio de las señales de audio en estado estable se debe a una resolución frecuencial más fina, mientras que las señales transitorias requieren una resolución temporal más fina. No es posible alcanzar simultáneamente una resolución temporal y frecuencial grandes, por lo que debe hacerse un compromiso.

AC-3 utiliza un banco de filtros de cancelación de "aliasing". Cada señal de entrada filtrada en el dominio del tiempo para cada canal es individualmente ventaneada (el bloque de audio se multiplica por una función ventana para reducir los efectos de frontera en la transformada y para mejorar la selectividad del banco de filtros; una ventana simétrica de 512 puntos se forma con 256 coeficientes) y filtrada con un banco de filtros de análisis TDAC. AC-3 utiliza la ventana Kaiser-Bessel derivada (KBD: Kaiser-Bessel derived)^[16]. En el codificador y decodificador existe una tabla con los valores de la función ventana. Los bloques traslapados de 512 muestras son entonces transformados al dominio de la frecuencia. Cuando un bloque ventaneado normal es transformado, el resultado son 256 coeficientes de la transformada en el dominio de la frecuencia. Los bloques más chicos son construidos mediante tomar el segmento de audio ventaneado de 512 muestras y dividirlo en dos segmentos con 256 muestras cada uno. Cada mitad del bloque produce 128 coeficientes de la transformada para un total de 256. Este es idéntico al número de coeficientes producidos por solo bloque de 512 muestras, pero con una resolución temporal mejorada dos veces. Los coeficientes de la transformada de dos medios bloques son

colocados juntos (intercalados), coeficiente por coeficiente para formar un solo bloque de 256 muestras. Este bloque es cuantizado y transmitido de forma idéntica a un solo bloque largo.

Cada bloque de la transformada está formado de 10.66 ms representativos de audio (a una frecuencia de muestreo de 48 kHz), aunque las transformadas se llevan a cabo cada 5.33 ms. La tasa del bloque de audio es entonces 187.5 Hz. Durante las condiciones transitorias cuando es útil una resolución temporal más fina, el tamaño del bloque es reducido a la mitad de tal manera que las transformadas ocurran cada 2.67 ms. La resolución frecuencial del banco de filtros es 93.75 Hz. La mínima resolución temporal es 2.67 ms. La resolución completa del banco de filtros se utiliza; los filtros individuales no se combinan en bandas más anchas (o bandas críticas), excepto durante una porción de la rutina de asignación de bits. La asignación de bits puede ocurrir abajo del nivel de coeficientes de la transformada individuales, con los coeficientes vecinos recibiendo diferentes asignaciones.

De acuerdo con las banderas de encendido del bloque, cada bloque de audio es transformado al dominio de la frecuencia realizando una transformada MDCT, esta puede ser una transformada larga de $N=512$ puntos, o dos transformadas más cortas de $N=256$ puntos.

Si $x[n]$ representa la secuencia temporal de entrada ventaneada. La secuencia de salida en el dominio de la frecuencia está definida por

$$X_D[k] = \frac{-2}{N} \sum_{n=0}^{N-1} x[n] \cos\left(\frac{2\pi}{4N}(2n+1)(2k+1) + \frac{\pi}{4}(2k+1)(1+\alpha)\right)$$

para $0 \leq k < N/2$

en donde

$$\alpha = \begin{cases} -1 & \text{para la primera transformada corta} \\ 0 & \text{para la transformada larga} \\ +1 & \text{para la segunda transformada corta} \end{cases}$$

El procesamiento posterior se lleva a cabo en el conjunto de los coeficientes de la transformada MDCT de los 5.1 canales, tomada como una sola entidad. Al final, el decodificador recupera las señales en el dominio del tiempo al reconstruir los coeficientes de la transformada a través de un banco de filtros de síntesis basado en la transformada inversa MDCT. Tanto el banco de filtros de análisis como el de síntesis son muestreados críticamente y reconstruidos de manera exacta, por lo tanto la información no tiene pérdidas. Adicionalmente ambos bancos de filtros utilizan un traslape entre bloques del 50%, evitando discontinuidades y fronteras en los bloques. Mientras la codificación no elimine o añada información audible, la salida decodificada sonará siempre igual que la entrada del codificador.

TESIS CON
 FALLA DE ORIGEN

4.6.6 Conversión a punto flotante^[17,18]

Los coeficientes de la transformada MDCT de AC-3 se convierten a la representación de punto flotante para el procesamiento que sigue, con mantisas que tienen un rango de magnitud de 0.5 a 1.0, y sus correspondientes exponentes. Los exponentes son valores de 5 bits que indican el número de ceros que hay adelante en la representación binaria de un coeficiente de frecuencia. El exponente actúa como un factor de escala para cada mantisa, igual a $2^{-\text{exp}}$. Los valores de exponentes permitidos tienen un rango que va desde 0 (para el valor más grande de coeficientes) hasta 24. Esto asegura que el procesamiento intermedio no imponga limitaciones prácticas al rango dinámico. El rango dinámico amplio es mantenido por las transformadas de entrada y salida mediante el uso de una etapa intermedia de adaptación según se necesite.

La representación de los datos por punto flotante, particularmente la presencia de los exponentes, también sirve como una ayuda en los cálculos para los procesos como la asignación de bits.

4.6.7 Estrategia de acoplamiento^[12]

De forma general, el promedio de demanda de bits de canales múltiples es aproximadamente proporcional a la raíz cuadrada del número de canales. Si toma 128 kbps codificar un solo canal, un promedio de 5.1 canales requerirá $128 * (5.1)^{1/2} = 289$ kbps, que caben cómodamente dentro de la tasa de transmisión mínima utilizada por AC-3 que es de 320 kbps.

Esto implica que la mayoría de las señales multicanal pueden ser comprimidas apropiadamente a través de la estrategia de acoplamiento y de la asignación global de bits. Esta técnica elimina la información redundante localizada en las frecuencias altas, y se basa en el fenómeno psicoacústico de que en las frecuencias altas el sistema auditivo humano localiza el sonido a través de las envolventes de las versiones de las señales filtradas en bandas críticas que llegan hasta los oídos, más bien que las señales mismas. En las frecuencias de audio grandes (arriba de unos 2 kHz) el oído es físicamente incapaz de detectar los ciclos individuales de una forma de onda de audio, y en su lugar responde a la envolvente de la forma de onda. La direccionalidad es determinada por el tiempo de retraso inter-aural de la señal envolvente, y por la respuesta en frecuencia percibida, la cual es afectada por la sombra o interferencia causada por la cabeza y la forma del oído. El acoplamiento toma ventaja del hecho de que el oído no es capaz de detectar independientemente la dirección de dos señales de alta frecuencia que se encuentran muy cercanamente espaciadas en la frecuencia.

Este comportamiento es explotado en AC-3 mediante separar las señales en las sub-bandas de alta frecuencia en las componentes envolvente y portadora, codificando generalmente la información de la envolvente con mayor precisión que la información de la portadora, y si es necesario, combinando (acoplando) selectivamente las componentes portadoras a través de los canales. Esto tiene un mínimo impacto audible, debido a que la

localización de la información se mantiene en los datos de la envolvente, y las portadoras se combinarían acústicamente en cualquier caso en los oídos de los escuchas, produciendo un resultado equivalente.

La información de la portadora codificada aumenta los arreglos de exponentes y mantisas, mientras la información de la envolvente se transmite como un arreglo de coeficientes de acoplamiento.

Esta técnica preserva de manera exacta las características sónicas espaciales así como otras, no solo de los sonidos que vienen de fuentes discretas (altavoces), sino también de imágenes fantasma que aparecen entre los altavoces. También mantienen tanto la potencia acústica de cada componente de la señal fuente, y la potencia acústica de cada alimentador de los altavoces.

El codificador AC-3 es el responsable de determinar la estrategia de acoplamiento. El codificador controla cuales canales de audio serán incluidos en el acoplamiento, y cuales permanecerán completamente independientes. El codificador controla a que frecuencia comienza el acoplamiento, la estructura de la banda de acoplamiento (los anchos de banda de las bandas acopladas), y cuando son enviadas las nuevas coordenadas de acoplamiento. A medida que las técnicas de codificación AC-3 han mejorado, la frecuencia de acoplamiento típica se ha incrementado hasta 10 kHz.

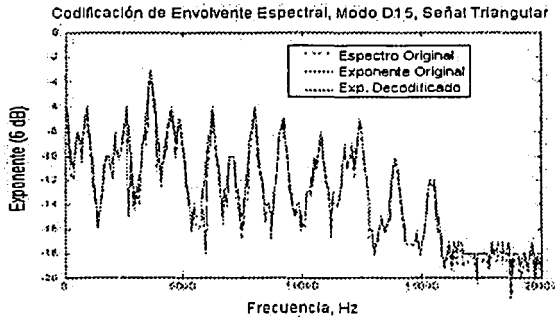
Los coeficientes de la transformada # 37 al #252 están agrupados en 18 sub-bandas de 12 coeficientes cada una. Debajo de la frecuencia de la primera sub-banda de acoplamiento, todos los canales se codifican de manera independiente. Arriba de la frecuencia indicada por la primera sub-banda de acoplamiento, los canales incluidos en el proceso de acoplamiento comparten un canal de acoplamiento común hasta la frecuencia indicada por la última sub-banda de acoplamiento. La última sub-banda codificada puede estar entre las sub-bandas 2 y 17. Las sub-bandas se combinan en bandas de acoplamiento, para las cuales las coordenadas de acoplamiento son generadas e incluidas en la trama de bits. La estructura de las bandas de acoplamiento se indica en la trama de AC-3.

4.6.8 Codificación exponencial y Asignación de bits^[12]

El asignador de bits analiza los coeficientes transformados en su representación de punto flotante con respecto a sus efectos de enmascaramiento y la relación con el umbral de audición absoluto, para calcular la codificación exponencial, así como la exactitud de la cuantización (número de bits) requeridos para codificar cada mantisa. El cálculo es realizado globalmente para tomar los canales como una entidad, usando un solo depósito de bits común, con algunos bits preasignados.

Tanto el codificador como el decodificador ejecutan la misma rutina central de asignación de bits, trabajando primeramente con la información de los exponentes de la MDCT. El codificador también tiene la opción de enviar la información no válida de asignación de bits, para permitir futuros refinamientos en los codificadores con los decodificadores existentes.

Para cada canal, la variación de los exponentes a través de la frecuencia y el tiempo es examinada, con el objeto de realizar la codificación exponencial más apropiada. Cuando se emplea la codificación D15 se tiene una envolvente espectral muy exacta, como se muestra en la figura 4.19. Si los exponentes indican un espectro relativamente plano, una estrategia de exponentes tal como la D25 o D45 puede ser usada.



TESIS
FALLA DE ORIGEN

Figura 4.19 Codificación Exponencial D15

En la figura 4.20 se muestra el resultado de aplicar el método D25 a la misma señal triangular mostrada en la figura 4.19. Si el espectro es muy tonal, entonces se usaría una estrategia exponencial de alta resolución espectral tal como la D15 y la D25. Si el espectro cambia poco a lo largo de los 6 bloques de un frame, los exponentes pueden ser enviados solo para el bloque 0, y reutilizados para los bloques 1 al 5. Si los exponentes cambian rápidamente durante el frame, los exponentes pueden ser enviados para el bloque 0 y para aquellos bloques en donde el grupo de exponentes difieran significativamente de los exponentes previamente enviados. Existe un intercambio entre la resolución frecuencial fina, la resolución temporal fina, y el número de bits requeridos para enviar exponentes. De manera general, cuando se opera a tasas de transmisión muy bajas, es necesario intercambiar la resolución temporal por la frecuencial.

TESIS CON
 FALLA DE ORIGEN

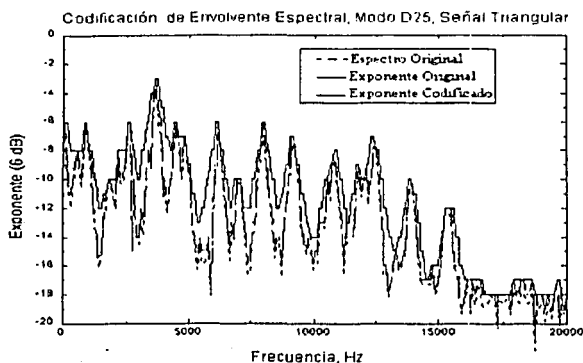


Figura 4.20 Codificación Exponencial D25

La transmisión de exponentes de AC-3 emplea codificación diferencial, en la cual los exponentes para un canal son codificados diferencialmente a través de la frecuencia. El primer exponente se envía como un absoluto, y el resto de exponentes se envía como diferenciales. Grupos de tres diferenciales se codifican en una palabra de 7 bits. Los exponentes absolutos encontrados al principio de la trama de bits en los grupos de exponentes codificados diferencialmente son enviados como valores de 4 bits, los cuales han sido limitados tanto en rango como en resolución para ahorrar un bit. Para los canales de ancho de banda completo y los LFE, el exponente absoluto inicial de 4 bits representa un valor de 0 a 15. Los exponentes más grandes que 15 son limitados a un valor de 15. Para el canal acoplado, los 5 bits del exponente absoluto están limitados a valores pares, y el bit menos significativo no se transmite. La resolución ha sido limitada a valores válidos de 0, 2, 4, ..., 24. Cada exponente diferencial puede tomar uno de cinco valores: -2, -1, 0, +1, +2. Esto permite deltas de hasta ± 2 (± 12 dB) entre exponentes. Estos cinco valores son mapeados en los valores 0, 1, 2, 3, 4 antes de ser agrupados.

En el modo D15, el mapeo se aplica a cada exponente diferencial individual para codificarlo dentro de la trama de bits. En el modo D25, cada par de exponentes diferenciales es representado por un solo valor mapeado en la trama de bits. En este modo el segundo exponente diferencial de cada par tiene un delta de 0 con respecto al primer elemento del par. El modo D45 es similar al D25 excepto por que cuatro exponentes diferenciales son representados por un valor mapeado. La figura 4.21 muestra el resultado de codificar un transitorio de castañas con el método D45.

Debido a que un solo exponente es efectivamente compartido por 2 ó 4 mantisas diferentes, los codificadores deben asegurar que el exponente elegido para el par o el cuarteto es el mínimo valor absoluto (que corresponde al exponente más grande) necesario para representar todas las mantisas.

Para todos los modos, grupos de tres valores mapeados (M1, M2 y M3) adyacentes (en frecuencia) se agrupan juntos y se codifican como valores de 7 bits de cuerdo con la fórmula siguiente

$$\text{valor agrupado codificado en 7 bits} = (25 * M1) + (5 * M2) + M3$$

El campo exponencial para un canal dado en un bloque de audio de AC-3 consiste de un solo exponente absoluto seguido por un número de estos valores agrupados.

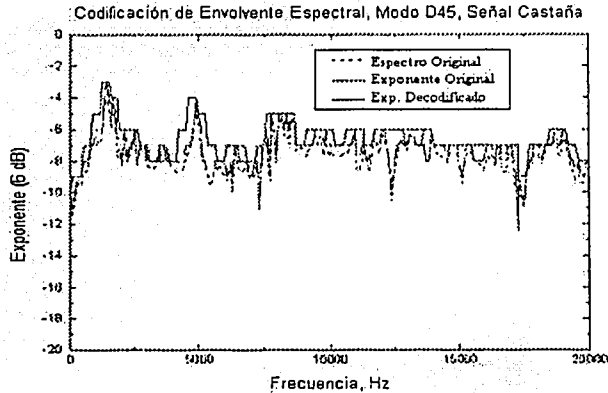


Figura 4.21 Codificación Exponencial D45

La rutina central de asignación de bits AC-3 comienza con la envolvente espectral decodificada, o exponentes, los cuales se considera que son la densidad espectral de potencia (power spectral density – psd) de la señal. La rutina de asignación de bits analiza la envolvente espectral de la señal de audio que está siendo codificada con respecto a los efectos de enmascaramiento para determinar el número de bits que asignará a cada mantisa del coeficiente de la transformada. La mayor parte de la rutina de asignación de bits es la convolución de la función de expansión que se acopla a la curva de enmascaramiento del oído (ver la figura 4.22), contra la densidad de potencia espectral. En el codificador, la asignación de bits se realiza de forma global en el conjunto de canales como una entidad, a partir de un contenedor de bits común. No hay bits preasignados en los exponentes y mantisas, permitiendo así que la rutina asigne los bits a través de los canales, frecuencias y bloques de audio, de manera flexible de acuerdo con la demanda de la señal.

TESIS CON FALLA DE ORIGEN

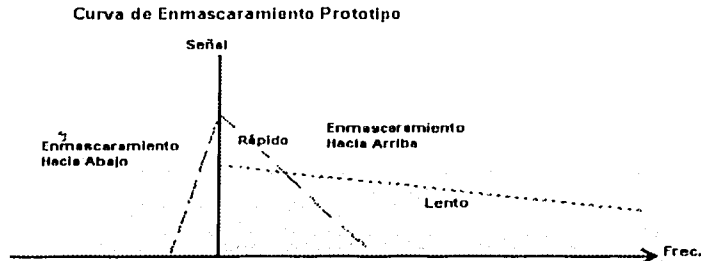


Figura 4.22 Curva de Enmascaramiento Prototipo

TESIS CON FALLA DE ORIGEN

La asignación de bits contiene un modelo paramétrico de la audición humana para estimar el nivel de umbral de ruido, expresado como función de la frecuencia, el cual separa las componentes espectrales audibles de las inaudibles. Varios parámetros del modelo de audición pueden ser ajustados por el codificador dependiendo de las características de las señales. La función expandida del oído está aproximada por tres curvas: una de forma muy empinada en la parte de las frecuencias bajas de la curva de enmascaramiento; una que decae muy rápido en la parte de las frecuencias altas; y una que decae lentamente en la parte de las frecuencias altas de la curva de enmascaramiento. La técnica ignora la curva de enmascaramiento de las frecuencias bajas por simplicidad a expensas de una sobre-asignación ocasional. De ahí que, una curva de enmascaramiento prototipo se define en términos de dos segmentos de línea que son los de las frecuencias altas (véase la figura 4.22), cada uno con su propia pendiente y su ordenada al origen. Una de muchas de estas pendientes y ordenadas al origen posibles son seleccionadas por el codificador para cada segmento de línea. El codificador puede hacer iteraciones en uno o en más parámetros como estos hasta que el resultado óptimo sea obtenido.

Cuando todos los parámetros utilizados para estimar el nivel de umbral de ruido hayan sido seleccionados por el codificador, la asignación de bits final se calcula. Los parámetros del modelo se transportan al decodificador con información adicional. El decodificador ejecuta la rutina en un solo paso.

Se ha estado desarrollando una técnica simplificada para ejecutar el paso de la convolución de la función expandida contra la densidad espectral de potencia. Se ejecuta una convolución simplificada de las dos diferentes curvas de enmascaramiento inclinadas hacia arriba contra la psd agrupada. El cálculo comienza en la frecuencia más baja del arreglo de bandas de psd y se mueve hacia las frecuencias más grandes. El umbral del nivel de ruido estimado se calcula en 50 bandas con un ancho de banda que no es uniforme. La estructura de las bandas, es independiente de la frecuencia de muestreo. Se realizan dos cálculos, uno para cada una de las curvas de enmascaramiento que decaen rápido y lento, a las cuales nos referiremos como escape rápido y escape lento. El cálculo se ejecuta en el dominio logarítmico, donde un decaimiento puede implementarse simplemente como un decremento. Mientras el cálculo se mueve hacia arriba banda por banda de frecuencia, cada psd agrupado nuevo es examinado. Si el psd en la nueva banda es significativo con respecto a los valores de escape corrientes, el nuevo psd es combinado en los términos del escape, y

es incrementado en valor. Si el psd nuevo es insignificante, los términos de escape anteriores son decrementados. El más grande de cada uno de los términos de escape en cada frecuencia es mantenido. El resultado de este cálculo es un arreglo que indica el enmascaramiento predicho, banda por banda.

Esta curva es comparada contra un umbral de audición, y el mayor de los dos valores se mantiene. El paso final es restar la curva de enmascaramiento predicha del arreglo psd original (sin agrupar) para determinar la relación señal a ruido (Signal to Noise Ratio - SNR) deseada para cada coeficiente de la transformada individual. Esto se muestra en la figura 4.23.

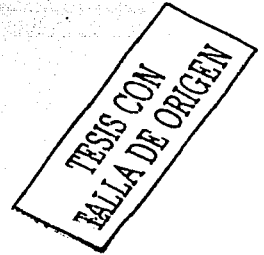
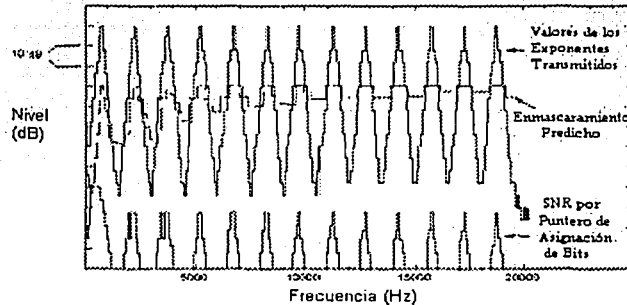


Figura 4.23 Cálculo de la asignación de bits

El arreglo de valores de SNR se convierte en un arreglo de apuntadores de asignación de bits (bit allocation pointers - *baps*). Los valores del arreglo *baps* apuntan hacia los cuantizadores que serán usados para cada mantisa de los coeficientes de la transformada. En este punto el codificador hace una cuenta de bits para determinar si la asignación de bits ha agotado el número de bits disponible. Todos los bits disponibles se toman de un banco común de bits, disponible para todos los canales. Si hay más bits disponibles, la mantisa individual de la SNR puede incrementarse, hasta que todos estos bits se agoten. Si se han asignado muchos bits, la mantisa individual de la SNR debe disminuirse, y/o debe llamarse al acoplamiento.

La asignación de bits requerida para cada mantisa se establece al hacer una búsqueda en una tabla que se basa en la diferencia entre la densidad de potencia espectral de la señal de entrada evaluada en una escala de frecuencia fina uniforme, y el umbral de nivel de ruido estimado evaluado en una escala de frecuencia tosca (por bandas). Por lo tanto, el resultado de la asignación de bits para un canal particular tiene la regularidad espectral correspondiente a la estrategia de exponentes empleada. Más específicamente, una asignación de bits separada será calculada para cada mantisa dentro del grupo de exponentes D15, cada par de mantisas dentro de un grupo de exponentes D25, y cada cuarteto de mantisas dentro de un grupo de exponentes D45.

4.6.9 Cuantización

Los resultados de los cálculos de la asignación de bits son utilizados para cuantizar los datos de la mantisa. Las mantisas son normalizadas, mediante un offset para proporcionar niveles de cuantización centrados en cero, con anchos iguales y simétricos (cuantización simétrica impar), que permitan minimizar la distorsión y facilitar el uso de un dither^[14] (ruido aleatorio que se le dan en lugar de valores cuantizados cuando el número de bits asignados a la mantisa es cero, con el objetivo de eliminar el ruido que se presentaría en una señal de muy poca amplitud).

Los *haps* son utilizados para la cuantización de la mantisa del bloque. Hay un *hap* por cada coeficiente de la transformada individual. Cada mantisa normalizada es cuantizada de acuerdo con el *hap* correspondiente. Todas las mantisas son cuantizadas en un nivel fijo de precisión. Las mantisas cuantizadas en 15 o menos niveles usan cuantización simétrica. Las mantisas cuantizadas en más de 15 niveles usa cuantización asimétrica que es una representación convencional en complemento a dos. Las mantisas cuantizadas asimétricamente son cuantizadas mediante el redondeo del número de bits indicado por el *hap* correspondiente. Las mantisas simétricamente cuantizadas son cuantizadas por medio del uso de una tabla de consulta. Algunos valores de mantisas cuantizadas son reunidas en grupos de dos o tres, y se codifican en una palabra código común, lo cual se hace para que exista una compresión mayor.

4.6.10 Empaquetado de datos

Los procesos antes descritos convierten cada bloque de 6 canales de señales temporales en series de arreglos derivados y valores escalares, incluyendo los exponentes transformados y las mantisas cuantizadas, información de la asignación de bits, coeficientes de acoplamiento y banderas de cuantización. Algunas de las mantisas cuantizadas son agrupadas juntas y codificadas por una sola palabra de código. En la última etapa del proceso de codificación, esta información es empaquetada en un solo bloque, junto con la información de sincronización, un encabezado e información diversa y de corrección de errores opcional. Se tiene cuidado de empaquetar los datos en una forma causal lógica, para que el decodificador pueda desempaquetarla. A continuación se encuentran las características de la trama de AC-3. La trama de bits se describe con todo detalle en el estándar de audio A/52^[14].

4.7 La trama de bits de AC-3

4.7.1 Características de la Trama de bits de AC-3.

La trama de bits de AC-3 está compuesta por paquetes (véase figura 4.24), que presentan un intervalo constante de tiempo de 1536 muestras PCM por todos los canales codificados. Cada paquete tiene un tamaño fijo, el cual depende solo de la frecuencia de muestreo y la tasa de datos codificada. Además, cada paquete es una entidad independiente, que no comparte ningún dato con los paquetes previos además del traslape que ocurre en la MDCT.

Al principio de cada paquete de AC-3 están los campos de Información de Sincronía (SI – Sync Information) y de Información de Trama de Bits (BSI – Bit Stream Information). Los campos SI y BSI describen la configuración de la trama de bits, incluyendo la frecuencia de muestreo, la tasa de transmisión, el número de canales codificados, y otros elementos a nivel del sistema. Hay también dos palabras CRC (Cyclic Redundance Check: Procedimiento de detección de errores por chequeo de redundancia cíclica)^[11] por paquete, uno al principio y otro al final, para la detección de errores.

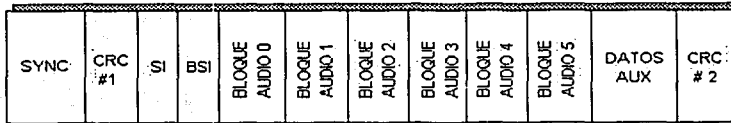


Figura 4.24 Estructura del Frame de AC-3

RECIBIDO CON FALLA DE ORIGEN

La unidad de tiempo fundamental para AC-3 está relacionada con el tamaño del bloque de la transformada. Mientras cada bloque transformado tiene un tamaño de 512 muestras, un traslape/suma de 100% de la transformada TDAC significa que los bloques son transformados cada 256 muestras, o cada 5.33 ms para una frecuencia de muestreo de 48 kHz. La tasa del bloque de la transformada es de 187.5 Hz. La sintaxis de AC-3 agrupa seis bloques de la transformada en una trama AC-3. La tasa de la trama es de 31.25 Hz, y cada trama dura 32 ms.

Cada trama AC-3 comienza con una palabra de sincronía de 16 bits. Después de la palabra de sincronía están 8 bits de información que indican la frecuencia de muestreo y el tamaño de la trama. A los primeros 3 bytes de la trama AC-3 se les conoce como *información de sincronía* (sync info) porque la información en estos bits es utilizada para obtener y mantener la sincronización con las tramas de AC-3.

Después de la información de sincronía se encuentra un grupo de datos a los que se les llama *información de la trama de bits* o BSI (bit stream info). Los datos BSI contienen información sobre el número de canales que están codificados, niveles de diálogo, código de lenguaje, información sobre los servicios asociados, etc. Todos los datos BSI son esencialmente estáticos, y son una descripción de los datos que se encuentran en los bloques de audio que le siguen.

A continuación de los BSI se encuentran los seis bloques de audio codificados. El primer bloque siempre contiene un recordatorio completo de los exponentes, las coordenadas de acoplamiento, y toda la demás información que se transmite condicionalmente. Los cinco bloques siguientes pueden o no contener información además de las mantisas cuantizadas. Los datos no acostumbrados al final de la trama AC-3 pueden ser considerados como *datos auxiliares* (aux data).

Con cada paquete hay seis bloques de audio, cada uno representando 256 muestras PCM por canal codificado (véase figura 4.25). El bloque de audio contiene las banderas de

encendido de los bloques, las coordenadas de acoplamiento, los exponentes, los parámetros de asignación de bits y las mantisas. Dentro de un mismo paquete está permitido compartir los datos, por lo que la información presente en el Bloque 0 puede ser reutilizada en los bloques siguientes.

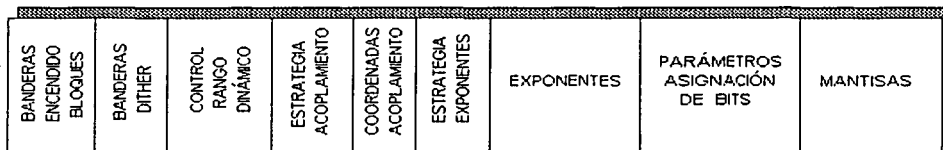


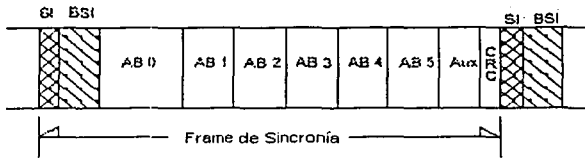
Figura 4.25 Estructura de un bloque de Audio AC-3

Un campo auxiliar de datos está localizado al final del paquete. Este campo permite a los diseñadores de sistemas implantar un control privado o estado de información en la trama de bits de AC-3 para un sistema de transmisión amplia.

El codificador AC-3 fue diseñado para que fuera una solución completa del subsistema de audio, incorporando muchas características que generalmente no están asociadas a la codificación de una baja tasa de transmisión. Estas incluyen un rango dinámico de compresión conveniente según el reproductor de audio del consumidor, la normalización de diálogos y el mezclado (downmixing) del audio multicanal en un número específico de canales de salida. Las palabras de control de rango dinámico están dentro de la trama de AC-3 y aplicadas en el decodificador, permiten diferentes modos de reproducción a partir de una trama de bits de una sola fuente.

4.7.2 Resumen de la Trama de Bits.

Una trama de bits de audio AC-3 serial codificado está hecha de una secuencia de frames de sincronía (ver la figura 4.26). Cada frame de sincronización contiene 6 bloques de audio codificado (AB – audio blocks), cada uno de los cuales representa 256 nuevas muestras de audio por canal. Un encabezado de información de sincronía (SI – synchronization information) al inicio de cada frame contiene la información necesaria para conseguir y mantener la sincronía. Un encabezado de información de la trama de bits (BSI – bit stream information) le sigue a SI, y contiene parámetros que describen al servicio de audio codificado. Los bloques de audio codificado pueden ser seguidos de un campo de datos auxiliares (Aux – auxiliary data). Al final de cada frame está un campo de verificación de errores que incluye una palabra CRC para detección de errores. Una palabra adicional CRC se localiza en el encabezado SI, cuyo uso en el decodificador, es opcional. La estructura de la trama de bits de AC-3 se resume en la Tabla 4.2.



TESIS CON FALLA DE ORIGEN

Figura 4.26 Frame de Sincronía de AC-3

Sintaxis	Información de sincronía { syncinfo }	Tamaño de palabra [bits]
Syncword:	palabra de sincronía	16
Cre1:	cyclic redundancy check 1	16
Fscod:	frecuencia de muestreo codificada	2
Frmsizecod:	tamaño de frame codificado	6
Información de trama de bits { bsi }		
Bsid:	identificación de trama de bits	5
Bsmod:	modo de trama de bits	3
Acmod:	modo de audio codificado	3
Cmixlev:	nivel de mezclado central	2
Surmixlev:	nivel de mezclado surround	2
Dsurmud:	modo Dolby surround	2
Lfeon:	canal LFE encendido	1
Dialnorm:	normalización diálogo	5
Compre:	existe palabra de ganancia de compresión	1
Compr:	palabra de ganancia de compresión	8
Langcode:	existe código idioma	1
Langcod:	código de idioma	8
Audprodi:	existe inf. de producción de audio	1
Mixlevel:	nivel de mezclado	5
Roomtyp:	tipo de sala	2
Dialnorm2:	normalización diálogo, canal 2	5
Compr2e:	existe palabra de ganancia de compresión, canal 2	1
Compr2:	palabra de ganancia de compresión, canal 2	8
Langcod2e:	existe código de idioma, canal 2	1
Langcod2:	código de idioma, canal 2	8
Audprodi2e:	existe inf. de producción de audio, canal 2	1
Mixlevel2:	nivel de mezclado, canal 2	5
Roomtyp2:	tipo de sala, canal 2	5
Copyrightb:	bit de marca registrada	1
Origbs:	trama de bits original	1
Timecod1e, timecod2e:	existe codificación de tiempo de primera y segunda mitad.	2
Timecod1:	codificación de tiempo, primera mitad	14

Sintaxis	Tamaño de palabra [bits]
Timecod2: codificación de tiempo, segunda mitad	14
Addbsie: existe información adicional de trama de bits	1
Addbsil: longitud de información adicional de trama de bits	6
Addbsi: información adicional de trama de bits	$(\text{addbsil}+1) \times 8$
Bloque de Audio {audblk}	
Blksw[ch]: bandera de encendido de bloque	1
Dithflag[ch]: bandera de dither	1
Dynrng: existe palabra de ganancia de rango dinámico	1
Dynrng: palabra de ganancia de rango dinámico	8
Dynrng2e: existe palabra de ganancia de rango dinámico, canal 2	1
Dynrng2: palabra de ganancia de rango dinámico, canal 2	8
Cplstre: existe estrategia de acoplamiento	1
Cplinu: acoplamiento en uso	1
Chincpl[ch]: canal en acoplamiento	1
Phsfginu: banderas de fase en uso	1
Cplbegf: código de frecuencia de inicio de acoplamiento	4
Cplendf: código de frecuencia de final de acoplamiento	4
Cplbndstrc[sbnd]: estructura de acoplamiento de banda	1
Cplcoe[ch]: existen coordenadas de acoplamiento	1
Mstrepco[ch]: coordenada de acoplamiento maestra	2
Cplcoexp[ch][bnd]: exponente de coordenada de acoplamiento	4
Cplcomant[ch][bnd]: mantisa de coordenada de acoplamiento	4
Phsflg[bnd]: bandera de fase	1
Rematstr: estrategia de "rematrixing"	1
Rematflg[rband]: bandera de "rematrix"	1
Cplexpstr: estrategia de acoplamiento de exponentes	2
Chexpstr[ch]: estrategia de exponentes de canal	2
Lfeexpstr: estrategia de exponentes de canal LFE	1
Chbwcod[ch]: código de ancho de banda de canal	6
Cplabexp: acoplamiento de exponentes absolutos	4
Cplexps[grp]: acoplamiento de exponentes	7
Exps[ch][grp]: exponentes de canal	4 ó 7
Gainrng[ch]: código de rango de ganancia de canal	2
Lfeexps[grp]: exponentes de canal LFE	4 ó 7
Baie: existe información de asignación de bits	1
Sdeycod: código de caída lenta	2
Fdeycod: código de caída rápida	2
Sgaincod: código de ganancia lenta	2
Dbpbcod: código de dB por bit	2
Floorcod: código de piso de enmascaramiento	3
Snroffste: existe offset de SNR	1
Csnroffst: offset SNR áspero	6
Cplfsnroffst: offset fino de SNR de acoplamiento	4

Sintaxis	Tamaño de palabra [bits]
Cplfgaincod: código de ganancia rápido de acoplamiento	3
Fsnroffst[ch]: offset fino de SNR de canal	4
Fgaincod[ch] : código de ganancia rápido de canal	3
Lfefsnoffst: offset fino de SNR de canal LFE	4
Lfelfgaincod: código de ganancia rápido de canal LFE	3
Cplleake: existe inicialización de acoplamiento de fuga	1
Cplfleak: inicialización de acoplamiento de fuga rápida	3
Cplslcak: inicialización de acoplamiento de fuga lenta	3
Deltbaie: existe información de asignación de bits delta	1
Cpldeltbae: existe asignación de bits delta de acoplamiento	2
Deltbae[ch]: existe asignación de bits delta	2
Cpldeltseg: número de segmentos de la asignación de bits delta de acoplamiento	3
Cpldeloffst[seg]: offset de asignación de bits de acoplamiento delta	5
Cpldeltlen[seg]: longitud de asignación de bits delta de acoplamiento	4
Cpldeltba[seg]: asignación de bits de acoplamiento delta	3
Deltseg[ch]: número de segmentos de asignación de bits delta de canal	3
Deltoffst[ch][seg]: offset de asignación de bits delta de canal	5
Deltlen[ch][seg]: longitud de asignación de bits delta de canal	4
Deltba[ch][seg]: asignación de bits delta de canal	3
Skiple: existe longitud de salto	1
Skipl: longitud de salto	9
Skipfld: campo de salto	skipl×8
Chmant[ch][bin]: mantisas de canal	0 a 16 bits
Cplmant[bin] : mantisas de acoplamiento	0 a 16 bits
Lfemant[bin] : mantisas de canal LFE	0 a 16 bits
Campo de datos auxiliares {auxdata}	
Auxbits : bits de datos auxiliares	n
Auxdatal : longitud de datos auxiliares	14
Sintaxis	Tamaño de palabra [bits]
Auxdatac: existen datos auxiliares	1
Campo de detección de errores en la trama {errorcheck}	
Crcsv: bit reservado CRC	1
Crc2: cyclic redundancy check 2	16

Tabla 4.2 Estructura del Frame de Sincronía de AC-3

4.8 El proceso de decodificación de AC-3

ERROR CON FALLA DE ORIGEN

El proceso de decodificación^[19] es básicamente el inverso del proceso de codificación. El decodificador, mostrado en la figura 4.27, debe sincronizar la trama de bits codificada,

verificar si hay errores, y cambiar el formato de los diferentes tipos de datos tales como la envolvente espectral codificada y las mantisas cuantizadas.

En el decodificador se ejecuta la rutina de asignación de bits y los resultados se utilizan para desempacar y decuantizar las mantisas. La envolvente espectral es decodificada para producir los exponentes. Los exponentes y mantisas son transformadas de nuevo al dominio del tiempo para producir las muestras PCM temporales decodificadas.

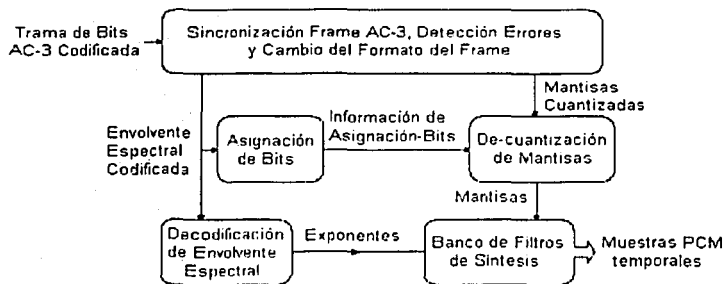


Figura 4.27 Diagrama de bloques de la decodificación AC-3

Un decodificador AC-3 real es más complicado que el mostrado en el diagrama anterior. Las siguientes funciones también deben incluirse:

- Debe aplicarse encubrimiento o silenciamiento de errores en caso de que se detecten errores en los datos.
- Los canales que tienen su contenido de altas frecuencias acoplado junto deben ser desacoplados.
- La resolución del banco de filtros de síntesis debe ser alterado dinámicamente de la misma manera que el banco de filtros de análisis ha sido alterado durante el proceso de codificación.

4.9 Bloques principales del decodificador AC-3

En los párrafos siguientes se describen los principales procesos que tienen lugar en la decodificación con mayor detalle. Véase la figura 4.28, donde se muestra un diagrama de bloques del proceso desglosado.

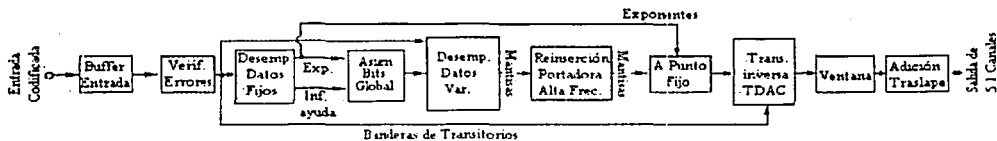


Figura 4.28 Diagrama desglosado de la técnica de decodificación AC-3

4.9.1. Buffer de entrada

El decodificador, tal como el codificador, se encuentra estructurado por bloques, y por lo tanto establece y mantiene la sincronización con el tren de bits entrante, reuniendo un bloque codificado entero en un buffer de entrada antes de proceder a decodificarlo.

4.9.2 Verificación de errores

Cada bloque de entrada del decodificador es verificado para que haya una consistencia interna, así como la presencia de la información de estado para un procesador de corrección de errores externo opcional. Si se indica una condición de error incorregible, el decodificador debe reutilizar el último bloque de entrada bueno en lugar del actual para ocultar el error. La naturaleza del proceso de reconstrucción de la señal de traslape/adición hace que esta forma de encubrimiento de errores relativamente benigno, y un bloque bueno dado puede ser generalmente repetido muchas veces, si es necesario, antes de que las condiciones de error extendido del decodificador requieran que ambos se queden en silencio o, en aplicaciones como las bandas sonoras de las películas, volver al uso de la banda sonora analógica coexistente.

4.9.3 Desempacado de los datos de formato fijo

El desempacado de datos se lleva a cabo en dos etapas. Primero se desempacan los datos de formato fijo, incluyendo los exponentes, los coeficientes de acoplamiento, y las banderas de modo. Las porciones relevantes de estos datos se utilizan entonces por el decodificador para recuperar la asignación de bits, la cual es utilizada para desempacar los datos de formatos variable, principalmente los arreglos de la mantisa de la MDCT.

4.9.4 Decodificación de la asignación de bits

La rutina de la decodificación de la asignación de bits es casi idéntica a la correspondiente del asignador del codificador, excepto que esta utiliza los resultados intermedios transmitidos para ahorrar tiempo, y opcionalmente para modificar la asignación de bits derivada como si fuera dirigida por el codificador. Este arreglo también permite al decodificador calcular la asignación de bits de un canal a la vez, reduciendo los requerimientos de memoria del decodificador.

La asignación de bits del decodificador debe coincidir exactamente con la del codificador para que los datos con formato variable sean desempaquetados apropiadamente, o se pueden introducir artefactos a la salida.

4.9.5 Desempacado de datos de formato variable

La asignación de bits decodificada recuperada, que especifica el tamaño cuantizado de cada mantisa, es utilizado para desempacar los datos de formato variable de la trama de bits codificada.

4.9.6 Conversión a punto fijo

En preparación para la transformada MDCT inversa, los datos de mantisas y exponentes son combinados para reconstruir los coeficientes de la MDCT de punto fijo. Si se está usando un ruido aleatorio (dither), se sustrae en este punto del proceso.

4.9.7 Reinserción de la portadora de alta frecuencia (Desacoplamiento)

Los coeficientes de alta frecuencia que han sido codificados como portadora y la información de la envolvente son reconstruidos mediante combinar las portadoras con sus correspondientes coeficientes de acoplamiento.

4.9.8 Transformada Inversa^[14]

Los coeficientes de la transformada MDCT recuperados de cada canal son transformados inversamente de nuevo hacia el dominio del tiempo, ventaneados, y traslapados/sumados para producir las señales digitales decodificadas en el dominio del tiempo. Los coeficientes de Subwoofer son llenados con ceros en las frecuencias medias y altas antes de la transformación, para que la salida Subwoofer en el dominio del tiempo tenga una frecuencia de muestreo completa.

4.9.9 Buffer de salida PCM

Los decodificadores típicos proporcionarán muestras PCM a la salida con la frecuencia de muestreo PCM. Debido a que los bloques de muestras resultan del proceso de decodificación, se requiere un buffer de salida.

4.9.10 Salida PCM

Las muestras PCM de salida pueden ser liberadas en una forma adecuada para su interconexión con un convertidor digital a analógico (DAC)^[1,2,11].

4.10 Algunas características adicionales de AC-3^[12,14,20]

AC-3 libera todos los canales codificados hacia el decodificador, el decodificador tiene completa flexibilidad para mezclar los canales como sea apropiado para determinada situación. Situaciones de audición diferentes requieren diferentes coeficientes de mezclado.

Para las situaciones que estén disponibles menos de 5.1 canales de presentación (altavoces), el decodificador puede mezclar los 5.1 canales fuente para obtener el número requerido de canales a la salida, por ejemplo un decodificador de solo dos canales de salida.

La uniformidad en el volumen es alcanzada mediante determinar el nivel subjetivo de un diálogo hablado normal, y codificar explícitamente este nivel en la trama de datos como una palabra de control llamada *nivel de diálogo* (dialogue level). El decodificador AC-3 puede entonces interactuar con el nivel de reproducción del sistema. Los programas

diferentes pueden tener distintos niveles de diálogo lo cual simplemente significa que tienen diferentes cantidades de altura disponible para el efecto dramático. Cuando el escucha ajusta el control de nivel, el nivel del diálogo normal reproducido (por ejemplo sin gritos o susurros) se fijará en nivel de presión subjetivo de sonido deseado. Cuando un nuevo segmento de programa comienza, con el diálogo codificado en un nivel diferente el sistema de reproducción puede usar la palabra de control del nivel de diálogo codificado para hacer el ajuste correspondiente en el volumen del sistema de reproducción. El resultado es que para todos los programas y todos los canales, el nivel del diálogo reproducido será uniforme.

El codificador AC-3 contiene un sistema de control de rango dinámico integral. Durante la codificación, o en cualquier punto posterior, las palabras de rango dinámico pueden ser colocadas dentro de la trama de bits de AC-3. Estas palabras de control son utilizadas por el decodificador para alterar el nivel del audio decodificado de acuerdo con los bloques (cada 5.3 ms)

REFERENCIAS

- [1] Libro de CASTRO, A.; "Teleinformática para Ingenieros en Sistemas de Información"; Editorial Reverté; Segunda Edición; España, 1999.
- [2] Libro de INGLIS, A.; "Video Engineering"; Mc Graw-Hill; E.E.U.U., 1996.
- [3] Compañía distribuidora de equipo de RF para estaciones televisoras. Sitio web http://broadcastengineering.com/ar/broadcasting_sdtv_format/
- [4] Libro de HIGASHIDA, B.; "Ciencias de la Salud"; Mc Graw-Hill; Segunda Edición; México, 1991.
- [5] Libro de OPPENHEIM, A.; "Señales y Sistemas"; Prentice-Hall; Segunda Edición; México, 1998.
- [6] Libro de PŠENIČKA, B.; "Procesamiento Digital de Señales –La transformada de Fourier y Filtros Digitales"; Facultad de Ingeniería, UNAM; México, 1994.
- [7] Libro de PORAT, B.; "A course in Digital Signal Processing"; John Wiley & Sons; E.E.U.U., 1997.
- [8] Libro de INGLE, V.; "Digital Signal Processing using Matlab V.4"; International Thomson Publishing; E.E.U.U., 1997.
- [9] Enciclopedia gratuita en la red Wikipedia Sitio web http://www.wikipedia.org/wiki/Discrete_cosine_transform
- [10] Enciclopedia gratuita en la red Wikipedia Sitio web <http://www.wikipedia.org/wiki/MDCT>
- [11] Libro de TOMASI, W.; "Sistemas de comunicaciones electrónicas"; Pearson Educación; Segunda Edición; México, 2001.
- [12] Publicaciones del laboratorio de investigaciones de Audio Dolby Todd, C; et. al.; "AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"; Audio Engineering Society, 96th Convention, 1994. Sitio Web <http://www.dolby.com/tech/ac3flex.html>
- [13] Documento oficial de la norma A/53 de ATSC "Digital Television Standard, Revision B with Amendments 1 and 2"; Estándar ATSC A/53B; ATSC; E.E.U.U. 2003.
- [14] Documento oficial de la norma A52 de ATSC "Digital Audio Compression (AC-3). Revision A"; Estándar ATSC A/52A; ATSC; E.E.U.U. 2001.

[15] Libro de PŠENIČKA, B.; "Filtrado de Señales Analógicas"; Facultad de Ingeniería, UNAM; México, 1994.

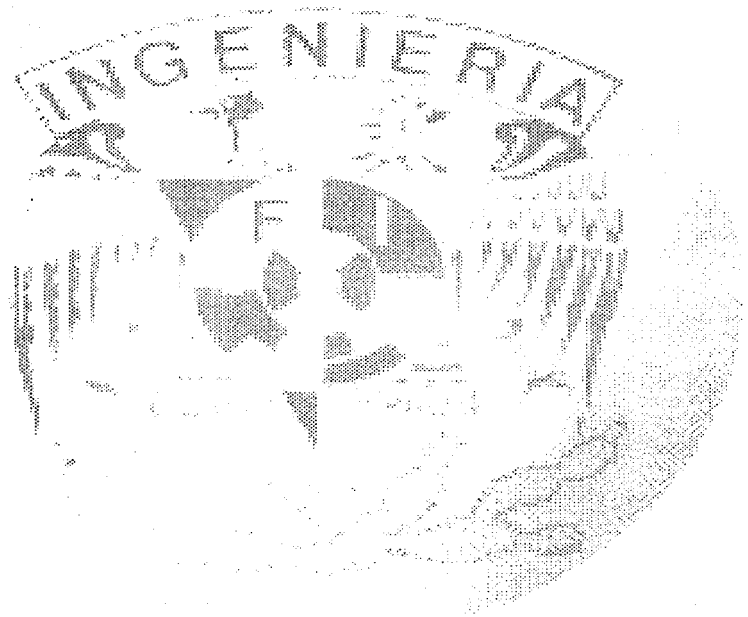
[16] Enciclopedia gratuita en la red Wikipedia Sitio web
http://www.wikipedia.org/wiki/Kaiser_window

[17] Libro de SPASOV, P.; "Microcontroller Technology -The 68HC11"; Prentice-Hall; Segunda Edición; E.E.U.U., 1996.

[18] Libro de Métodos Numéricos. Burden, R. ; "Análisis Numérico"; Grupo editorial iberoamérica; Segunda edición; México 1996.

[19] Libro de Vernon, Steve; "Design and implementation of AC-3 coders"; IEEE Tr. Consumer Electronics, Vol. 41, No. 3, August 1995.

[20] Plublicaciones del laboratorio de investigaciones de Audio Dolby sitio web
<http://www.dolby.com/tech/L.mn.0002.5.1Guide.s.pdf>



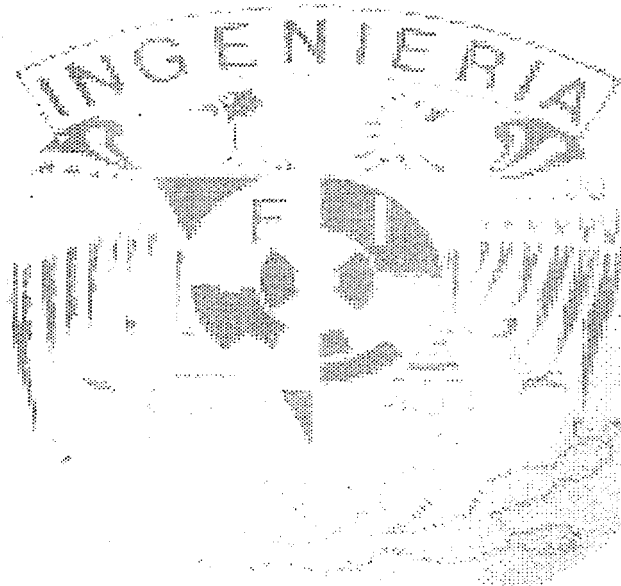
TESS CON
FALLA DE ORIGEN

124



Capítulo # 5
Transporte
y
Multiplexado

TESIS CON
FALLA DE ORIGEN



TESIS CON
TÍTULO DE INGENIERO

5

En este capítulo se tocará lo concerniente al multiplexado de la información procedente de los codificadores de audio y video en forma de trenes elementales de datos (Elementary Streams, ES) para que posteriormente por medio del sistema de transporte basado en las especificaciones del sistema MPEG-2, se genere el flujo de transporte (Transport streams TS), que permite realizar técnicas de protección de errores y proporciona gran flexibilidad para el servicio de multiplexado de audio, video y datos, entre otras ventajas.

Como se mencionó anteriormente, el transporte y el multiplexado están basados en MPEG-2, con lo cual se necesita saber el estándar antecesor al mismo, MPEG-1, del cual recupera gran parte de su estructura. A continuación se analizarán los puntos más importantes de MPEG-1, para así poder entender de una mejor manera los conceptos que se utilizan y llevar a cabo de forma exitosa el multiplexado y transporte de la señal digital, ya que son de vital importancia en los procesos que conforman la televisión digital.

5.1 Multiplexado de las señales en MPEG-1

La parte 1 (sistema) de la norma ISO/IEC 11172 (MPEG-1) define así las reglas para la constitución de una "capa de sistema" 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 "empaqueta" la capa de compresión son las siguientes:

- ✓ Sincronización de los trenes elementales en la reproducción.
- ✓ Combinación de estos trenes múltiples en un solo tren de datos.
- ✓ Inicialización de las memorias intermedias al comienzo de la reproducción.
- ✓ Gestión continua de estas memorias intermedias.
- ✓ Por ultimo, identificación del tiempo por marcadores.

Un conjunto de codificación MPEG-1 de "sistema" debe, por tanto realizar la codificación de los datos de video y audio, después el multiplexado de esta información y de los datos privados, añadiendo la información necesaria para su sincronización y otra información relativa 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 en ingles System Target Decoder, STD).

5.2 Generación de los flujos de la señal MPEG-2.

En la figura 5.1 se muestra de manera esquemática la secuencia de operaciones que da lugar a los flujos de programa y de transporte MPEG-2.

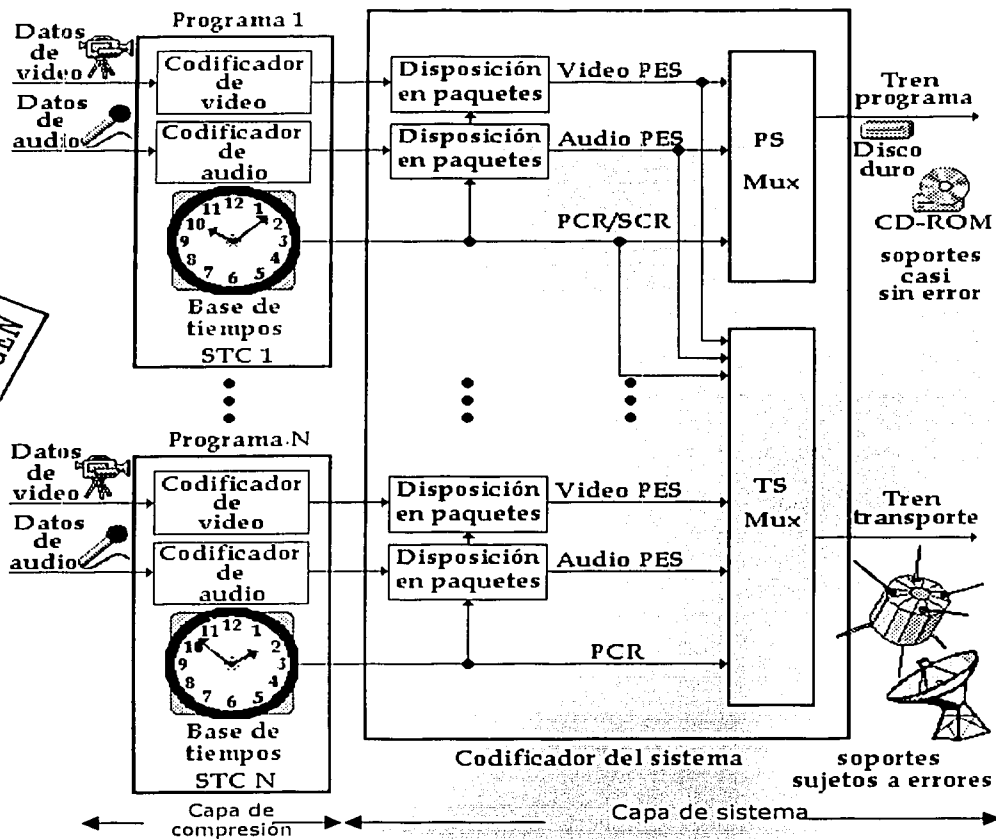


Figura 5.1.- Flujos de Programa y de Transporte MPEG-2

Dicha secuencia de operaciones puede dividirse en dos grandes bloques, denominados "Capa de Compresión" y "Capa de Sistema":

En la "Capa de Compresión" se realizan las operaciones propiamente dichas de codificación MPEG, recurriendo a los procedimientos generales de compresión de datos, y aprovechando además, para las imágenes, su redundancia espacial (áreas uniformes) y temporal (imágenes sucesivas), la correlación entre puntos cercanos y la menor sensibilidad del ojo a los detalles finos de las imágenes fijas.

En cuanto al audio, se utilizan modelos psicoacústicos del oído humano, que tienen en cuenta la curva de sensibilidad en frecuencia (máxima entre 1 y 5 KHz), los efectos de

enmascaramiento frecuencial (señales simultáneas a diferentes frecuencias) y enmascaramiento temporal (un sonido de elevada amplitud enmascara sonidos más débiles anteriores o posteriores), para reducir la cantidad de datos que hay que transmitir, sin deteriorar de forma perceptible la calidad de la señal de audio.

En la "*Capa de Sistema*" se realizan las operaciones que conducen a la obtención de los flujos de señal MPEG-2, consistentes en la organización en "paquetes", de los datos comprimidos y el posterior multiplexado de todas las señales asociadas al programa (vídeo, audio, datos).

En el caso de Flujos de Programa MPEG-2, se multiplexan todos los componentes del programa y se incorpora el reloj del sistema, pero solamente se transmite la información correspondiente a un único programa audiovisual. En cambio, en el caso de Flujos de Transporte MPEG-2, además de la posibilidad de multiplexaje de varios programas audiovisuales, se añaden diversas informaciones relativas al servicio como son: la tabla de asociación de programas (PAT), información para acceso condicional (CAT), el mapa de cada programa (PMT), la tabla de datos de la red (NIT), etc.

Características básicas de los flujos MPEG-2:

Conviene resaltar dos características notables de los flujos MPEG-2:

- No existen protecciones contra errores dentro del multiplexado. Las citadas protecciones y la subsiguiente modulación de los flujos MPEG son objeto de bloques de procesamiento posteriores, están en función del medio de transmisión elegido.
- No hay especificación física o eléctrica para los multiplexajes MPEG. El diseñador puede elegir los niveles de señal y tipo de conector que mejor se adapte a su aplicación.

Terminología empleada:

- "*Program*" o "*Programa*" Según la terminología MPEG, "*Program*" significa: Un servicio o canal simple de radiodifusión.
- "*Elementary Stream*" (E.S.) o "*Flujo Elemental*". Un "*Elementary Stream*" es el nombre dado a cada componente simple de un "*Programa*", después de que se ha codificado digitalmente y comprimido según MPEG. Así, un programa ya comprimido de TV se compone de varios "*Elementary Streams*": Uno para el vídeo, varios para sonido estéreo en diferentes idiomas, otro para el teletexto, etc.
- "*Packetised Elementary Stream*" (P.E.S.). Cada "*Elementary Stream*" se estructura en paquetes, dando lugar a un flujo que se denomina "*Packetised Elementary Stream*" (P.E.S.), y que está compuesto por "*PES-packets*". Como cada P.E.S. se obtiene directamente de cada E.S., y éstos son independientes, existirá por tanto un P.E.S. por cada E.S. original.

5.3 Procesamiento de la señal.

Como se ve en la figura 5.1, la parte de un programa digitalizado sin comprimir, se compone de varias informaciones simultáneas: vídeo, varias señales de audio estéreo en diferentes idiomas, teletexto, datos, etc.. Cada una de estas informaciones digitales, todavía sin comprimir, está agrupada en bloques elementales denominados "Unidades de Presentación".

❖ Unidades de Presentación de Vídeo:

En cuanto a la señal de vídeo digitalizada sin comprimir, se emplea el formato 4:2:2 con cuantificación de 8 bits recogido en CCIR Rec. 601-1, cuya "Unidad de Presentación" es una "Imagen o Cuadro" (830 kbytes para sistemas de 625 líneas).

❖ Unidades de Presentación de Audio:

En cuanto al audio, la "Unidad de Presentación" es una "Trama de Audio AES/EBU". En este caso, se requieren de 16 a 24 bits por muestra para proveer el rango dinámico y la relación señal/ruido deseados.

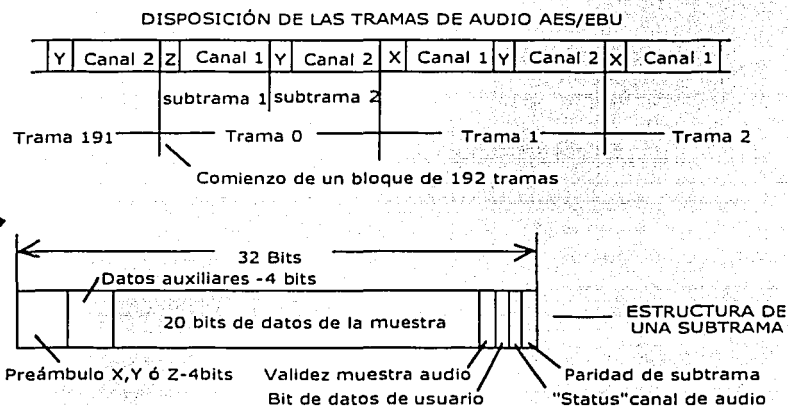


Figura 5.2.- Estructura de la "Trama de Audio AES/EBU"

El formato de los datos de audio sin comprimir puede verse en la figura 5.2 para una cuantificación de 20 bits. Consiste en secuencias de bloques de 192 tramas, cada una de las cuales está compuesta por dos subtramas correspondientes a los dos canales de audio estéreo.

A su vez, cada subtrama contiene los 20 bits de datos correspondientes a las muestras, precedidos por otros 4 bits de datos auxiliares, que pueden emplearse para extender la

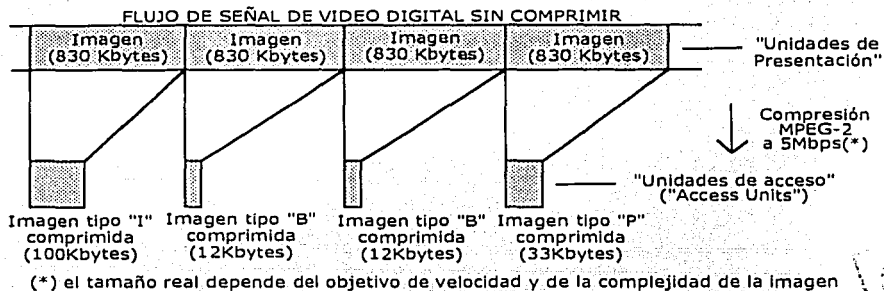
cuantificación a 24 bits. La subtrama comienza con un preámbulo de 4 bits que indica a qué canal pertenece la muestra y termina con 4 bits que aportan información sobre el canal incluyendo un bit de paridad. En total, cada subtrama tiene 32 bits.

El preámbulo de cada subtrama puede ser de tipo X, Y ó Z. Se emplea el tipo Z para indicar el comienzo de cada bloque de 192 tramas y después alternativamente los tipos X e Y para identificar las subtramas de los dos canales de audio: X para el canal 1 e Y para el canal 2.

5.4 Codificación MPEG.

El proceso continúa en la denominada "Capa de Compresión". Es aquí donde se reduce la velocidad binaria del flujo de señal, mediante la sustitución de las "Unidades de Presentación" por "Unidades de Acceso".

La figura 5.3 ilustra la operación asociada a la "Capa de Compresión".



TESIS CON FALLA DE ORIGEN

Figura 5.3.- Compresión MPEG-2 de la señal de video digital

En el caso de la señal de video, las "Unidades de Acceso" son de 3 tipos, correspondiendo a otros tantos tipos de imágenes MPEG:

- Imágenes tipo I (Intra): Se codifican sin ninguna referencia a otras imágenes, es decir: contienen todos los elementos necesarios para su reconstrucción.
- Imágenes tipo P (Previstas): Se codifican con respecto a la imagen de tipo I o de otra P anterior, gracias a las técnicas de predicción con compensación de movimiento. Su tasa de compresión es claramente mayor que la de las imágenes I.
- Imágenes tipo B (Bidireccionales): Se codifican por interpolación entre las dos imágenes de tipo I o P precedente y siguiente que las enmarcan. Ofrecen la tasa de compresión más alta.

En el caso de la señal de audio, las "Unidades de Acceso" típicamente contienen unas pocas decenas de milisegundos de audio comprimido.

El resultado de la codificación MPEG de una secuencia de vídeo, es una sucesión de "Unidades de Acceso de Vídeo". Dicha sucesión de unidades de acceso constituye el denominado "Flujo elemental de vídeo (Video E.S.)".

Paralelamente, el resultado de la codificación MPEG para el audio es una sucesión de "Unidades de Acceso de Audio" que componen el denominado "Flujo elemental de audio (Audio E.S.)".

5.5 Empaquetado P.E.S.

Ya en la "Capa de Sistema" nos encontramos con los "Empaquetadores P.E.S.", que constituyen el siguiente paso en la generación tanto del flujo de programa como del flujo de transporte MPEG-2.

Se trata de convertir cada "Flujo elemental (E.S.)" compuesto exclusivamente por "Unidades de acceso", en un "Paquetes de flujo elemental (P.E.S.)". Un P.E.S. está compuesto íntegramente por "Paquetes PES", como se muestra en la figura 5.4:

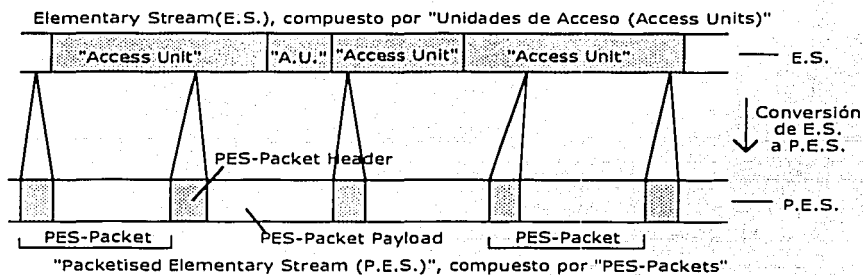


Figura 5.4.- Conversión de un E.S. en un P.E.S.

Un "Paquete PES" se compone de una "Cabecera" y de una "Carga Útil". Consiste simplemente en bytes de datos tomados secuencialmente desde el "Flujo elemental (E.S.)" original. No hay ningún requerimiento para alinear el comienzo de una "Unidad de acceso" y el comienzo de la "carga útil del paquete PES". Así, una nueva "unidad de acceso" puede comenzar en cualquier punto de la "carga útil" de un "Paquete PES", y también es posible que varias "unidades de acceso" pequeñas estén contenidas en un simple "Paquete PES".

TESIS CON FALLA DE ORIGEN

Los "Paquetes PES" pueden ser de longitud variable, limitados normalmente a 64kBytes, aunque la excepción a esta norma se da precisamente en los "Video PES", cuando se emplean en un "flujo de transporte", donde pueden tener longitud indefinida. Esta flexibilidad en la longitud de los paquetes PES puede ser explotada por los diseñadores de diferentes maneras: pueden usarse paquetes de longitud fija si interesa, o por ejemplo pueden ser de longitud variable de forma que siempre coincida el inicio de una unidad de acceso con el comienzo de la carga útil de cada paquete PES.

5.5.1 Cabecera del paquete PES

La figura 5.5 muestra los campos que comprende la "Cabecera" o "Header" de un "Paquete PES". Dicha cabecera tiene longitud variable, e incorpora a su vez datos para conocer su longitud.

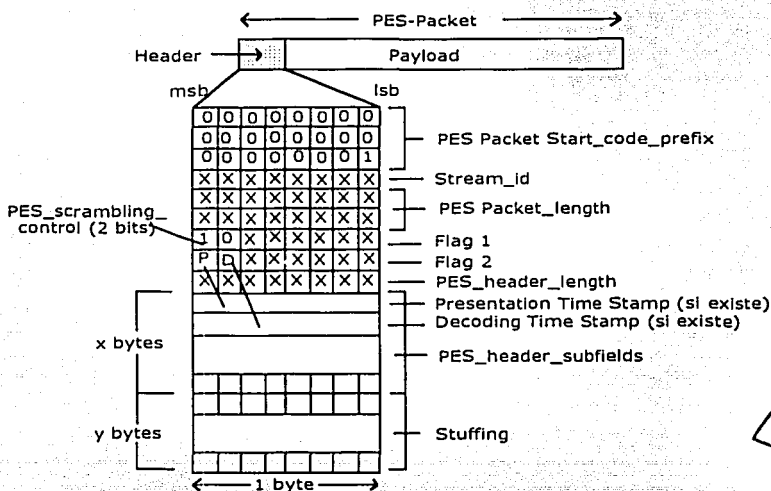


Figura 5.5.- Cabecera de un PES-Packet

En la tabla 5.1 pueden verse algunos detalles complementarios relativos a los campos de la cabecera de un paquete PES.

Campo	Definición (comentario)	Nº de bits
Start_code_prefix	prefijo del código de inicio (00 00 01 hex)	24
stream_id	Identificación del PES tipo (4 bits de mayor peso) y núm. (4 bits de menor peso) del PES	8
packet_length	longitud del PES-Packet (número de bytes restantes que siguen)	16
PES_scrambling_control	Indica si hay cifrado, y su código	2

TESIS CON FALLA DE ORIGEN

Flags	Marcadores	14
PES header length	Longitud restante de la cabecera (x+y)	8
PES header subfields	Campo variable, función de los flags	X bytes
stuffing bytes	bytes de relleno opcionales (valor FF hex)	Y bytes

Tabla 5.1 Estructura de la cabecera de un PES-Packet

5.5.2 PES-Packet start code

Los primeros 4 bytes, es decir: el “*PES-packet start code prefix*” más el “*Stream_id*” constituyen el “*PES-Packet start code*”. Debe garantizarse que esta combinación de 32 bits únicamente aparece en el comienzo de la cabecera de un PES-packet, excepción hecha de los datos de tipo “privado”.

El byte “*Stream_id*” permite distinguir los paquetes PES pertenecientes a los diferentes flujos elementales de un mismo Programa. MPEG especifica los valores permitidos para este campo, disponiéndose de 32 valores para E.S. de audio y de otros 16 valores para E.S. de vídeo.

5.5.3 Flags 1 y 2

Los bytes “*Flag 1*” y “*Flag 2*” son indicadores que muestran la presencia o ausencia de varios campos opcionales que pueden estar incluidos en la cabecera de un PES-packet. Estos campos opcionales llevan información complementaria relativa al PES, tales como: si está cifrado o no, prioridad relativa, datos de “copyright”, un campo para identificación de errores en el paquete, etc. Particularmente importantes son los 2 bits más significativos del “*Flag 2*”, marcados como *P* y *D* en la figura 5.5. Estos bits indican respectivamente la presencia de un campo denominado “*Marca de tiempo de presentación (PTS)*” y de otro campo denominado “*Marca de tiempo de decodificación (DTS)*” dentro de la cabecera del paquete PES.

5.5.4 Time Stamps

Los “Time Stamps” o “marcas de tiempo” son el mecanismo proporcionado por la “Capa de sistema” de MPEG-2 para asegurar una correcta sincronización en el decodificador entre “Flujos elementales” que están relacionados. Dichas “*marcas de tiempo*” son valores binarios de 33 bits expresados en unidades de 90kHz. El más importante de dichas marcas es el PTS. Especifica el momento en que una “Unidad de Acceso” debe retirarse del buffer del decodificador, ser decodificada y seguidamente presentada al espectador. MPEG supone que todo esto puede realizarse instantáneamente, aunque como en la práctica lleva algún tiempo, es responsabilidad del diseñador del decodificador resolver este problema.

En muchos tipos de flujos elementales es suficiente con los PTS, pero en el caso de flujos elementales compuestos por vídeo codificado MPEG, puede necesitarse un segundo tipo de marca de tiempo, que es la denominada “*Marca de tiempo de decodificación (DTS)*”.

Un “DTS” especifica el momento en que una unidad de acceso debe retirarse del buffer del decodificador y ser decodificada, pero no presentada al espectador. En lugar de esto último, la imagen decodificada se almacena temporalmente para ser presentada más tarde. Tal tratamiento es necesario únicamente para imágenes de tipo I y P de una secuencia de vídeo codificada MPEG; estas imágenes I y P son las que por interpolación permiten generar las imágenes de tipo B. Un DTS nunca aparece aislado, sino que debe ir acompañado de un PTS, que es el que informa del momento en que la unidad de acceso almacenada temporalmente debe presentarse al espectador. De esta forma, el PTS siempre será mayor que su DTS asociado (es decir, posterior en el tiempo), puesto que la presentación de la imagen al espectador ocurrirá más tarde.

No es necesario ubicar marcas de tiempo para cada unidad de acceso. El decodificador normalmente conoce de antemano el ritmo al que las unidades de acceso tienen que ser decodificadas, siendo suficiente proporcionarlas ocasionalmente, simplemente para asegurar el sincronismo del proceso de decodificación a largo plazo.

El requerimiento especificado en MPEG es que una marca de tiempo debe aparecer al menos cada 0.7 segundos en los PES de vídeo o audio. Por otra parte, si una unidad de acceso tiene una marca de tiempo asociada con ella, entonces esta última tiene que estar codificado en la cabecera del paquete PES en que dicha unidad de acceso comienza.

En los paquetes de audio existen marcas de tiempo del tipo PTS, los paquetes de audio nunca son transmitidos fuera de secuencia, entonces aquí no hay marcas de tiempo del tipo DTS. Estos marcadores sirven también para sincronizar las funciones de decodificación de audio y vídeo, 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.

5.5.5 PES_header_length

El “PES header data length field” es el último de los bytes obligatorios de una cabecera del paquete PES. Su valor indica el número de bytes opcionales presentes en la cabecera, hasta que se alcanza el primer byte de la carga útil del paquete PES o también llamado “payload”.

Hay 25 campos opcionales en una cabecera, que pueden contener por tanto un total de 200 bits de datos adicionales.

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 para obtener un flujo

constante o alinearse sobre los sectores físicos de un dispositivo de almacenamiento, así como los bytes de relleno en el interior de un paquete (máximo 16 por paquete).

Los paquetes se agrupan en flujos, cuya cabecera proporciona la información de sincronización y de flujo por medio de los campos de referencia de reloj de sistema (SCR) sobre 33 bits y del flujo del multiplexado. Los campos SCR se utilizan para sincronizar en el decodificador un reloj de sistema común a 90 KHz (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 5.6.

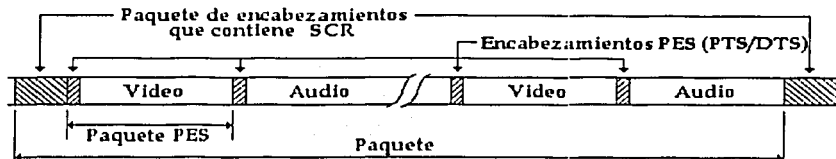


Figura 5.6 Un paquete es un conjunto de paquetes PES

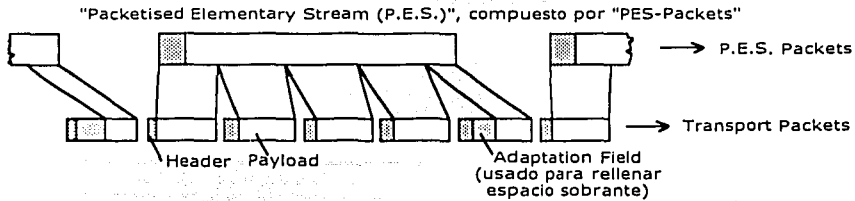
5.6 Flujo de programa

El flujo de "programa" de MPEG-2 se crea a partir de uno o varias PES que deben obligatoriamente compartir el mismo reloj de referencia. Este tipo de flujo está destinado a aplicaciones donde el dispositivo de almacenamiento o de transmisión sea susceptible de introducir muy pocos errores, 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).

5.7 Flujo de transporte (Transport stream)

El flujo 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; 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 con respecto a las diferentes tecnologías de transporte que existen como por ejemplo ATM. 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.

Los paquetes de transporte se forman a partir de los paquetes PES correspondientes a cada flujo elemental de señal (video, audio, datos), según se muestra en la figura 5.7.



TESIS CON FALLA DE ORIGEN

Figura 5.7.- División de un "P. E. S." en una serie de "paquetes de transporte"

5.7.1 Formación del "Flujo de transporte"

Los paquetes de transporte resultantes del proceso anterior, aplicado a cada uno de los flujos elementales de señal pertenecientes a su vez a varios programas audiovisuales, se disponen secuencialmente para configurar el "Flujo de Transporte MPEG-2 (TS)".

No existen condiciones en cuanto al orden en que los paquetes de transporte deben aparecer en el multiplexado tipo TS; tan sólo debe respetarse el orden cronológico de los paquetes de transporte pertenecientes a un mismo flujo elemental.

Es importante destacar que en el TS, además de los paquetes de transporte asociados a los flujos elementales de señal, es necesario incorporar paquetes de transporte que contienen información sobre el servicio, así como paquetes de transporte "nullos" que se emplean para absorber eventuales reservas de capacidad del múltiplex.

La figura 5.8 ilustra de forma simplificada el proceso de conformación del flujo de transporte MPEG-2.

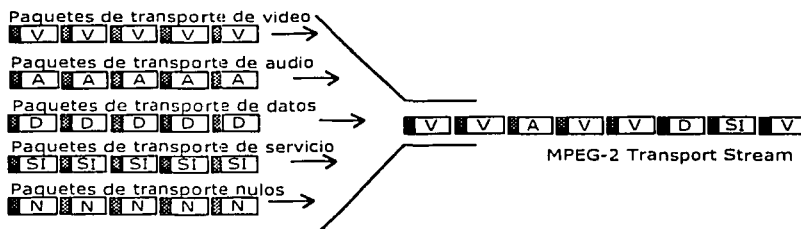


Figura 5.8.- Conformación del flujo de transporte MPEG-2, a partir de un número prácticamente ilimitado de flujos elementales de señal.

5.7.2 Paquetes de transporte

La figura 5.9 ilustra la estructura de un "paquete de transporte" o "transport packet".

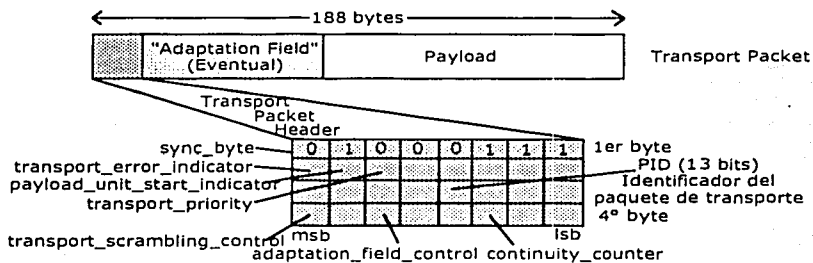


Figura 5.9 Estructura de un "transport packet"

5.7.3 Cabecera del paquete de transporte

A diferencia de los paquetes PES, cuya cabecera tiene longitud variable, la cabecera de un paquete de transporte tiene siempre una longitud fija de 4 bytes. El formato de la cabecera del paquete de transporte se detalla en las figuras 5.10, 5.11 y la tabla 5.2.

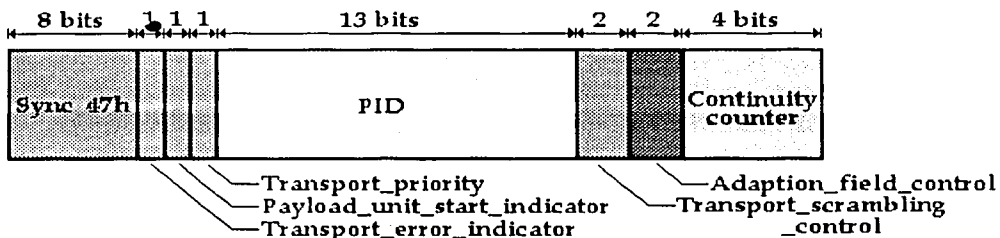


Figura 5.10 Detalle de la cabecera del paquete transporte

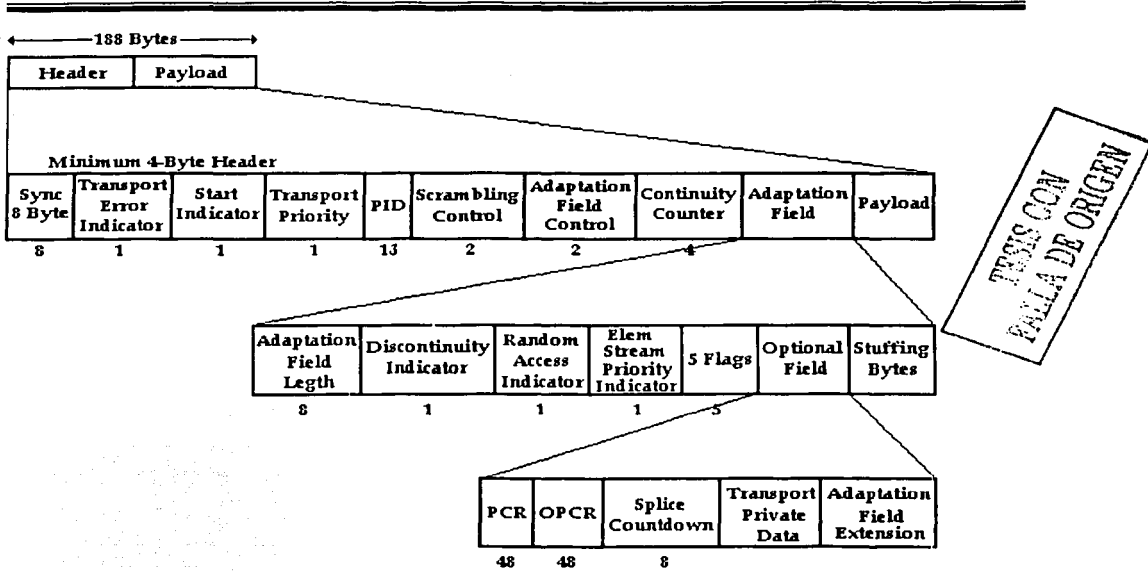


Figura 5.11 Estructura completa de un paquete de transporte

Campo	Definición (comentario)	Nº de bits
Sync byte	byte de sincronización 1000 0111 (47 hex)	8
Ei	transport_error_indicator (indica un error detectado mas atrás)	1
Pusi	payload_unit_start_indicator (inicio de PES en el paquete)	1
Tpr	transport_priority (indicador de prioridad)	1
PID	Packet Identifier (identificación del paquete)	13
Scr flags	transport_scrambling_flags (tipo de cifrado de transporte)	2
Af	adaptation_field_flag (campo de adaptación en el paquete)	1
Pf	payload_flag (datos útiles en el paquete)	1
Cc	continuity_counter (contador de continuidad entre trozos)	4

Tabla 5.2 Estructura de una cabecera de paquete de transporte MPEG-2

De los varios campos que contiene la cabecera de un paquete de transporte, hay 4 que son particularmente importantes:

❖ Sync Byte

Es el primer byte de la cabecera y siempre tiene el valor 47 (hexadecimal). Dicho valor no es único dentro de un paquete de transporte y puede ser tan frecuente, que aparezca en otros campos. Sin embargo, el hecho de que un byte de sincronía aparezca siempre cada 188 bytes dentro de un flujo de transporte hace fácil el proceso de su localización por los decodificadores para identificar el comienzo de cada paquete de transporte.

❖ Packet Identifier (PID)

Como se ha explicado, un TS puede contener muchos programas diferentes, cada uno de los cuales contiene a su vez muchos flujos elementales distribuidos en paquetes de transporte. El campo de 13 bits identificador del paquete (PID) se emplea para distinguir los paquetes de transporte asociados a un determinado flujo elemental, de entre todos los demás.

De los 2^{13} valores posibles para el PID, 17 están reservados para usos especiales. Esto deja 8175 posibles valores que pueden asignarse a los diferentes flujos elementales, por lo que esta cifra representa el máximo número de éstos que un TS puede contener. Es responsabilidad del multiplexor asegurar que cada flujo elemental tenga asignado un exclusivo valor de PID de entre todos los posibles. Por lo demás, MPEG no condiciona la asignación de valores concretos a los diferentes flujos elementales.

❖ Payload Unit Start Indicator

Se pone a 1 para indicar que el primer byte de la carga útil o "payload" del paquete de transporte, es también el primer byte de un paquete PES.

❖ Continuity count field

Se incrementa entre sucesivos paquetes de transporte pertenecientes al mismo flujo elemental. Esto permite al decodificador detectar la pérdida o ganancia de un paquete y así poder ocultar los errores que de otra forma podrían presentarse.

El proceso de formación según la norma ISO/IEC 13818-1, especifica de los paquetes de transporte está sujeto al cumplimiento de las dos condiciones fundamentales siguientes:

- El primer byte de cada PES-Packet debe ser el primer byte del "payload" de un transport packet.
- Un transport packet solamente puede contener datos tomados de un PES-Packet.

Este desperdicio de espacio puede minimizarse mediante una elección cuidadosa de la longitud de los paquetes PES. En principio, paquetes muy largos asegurarían una mayor proporción de paquetes de transporte completamente llenos de datos útiles.

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

Es improbable que un paquete PES rellene las “cargas útiles” de un número entero de paquetes de transporte de forma exacta. Como la longitud de los paquetes PES en general no es múltiplo de 184 bytes, con frecuencia se dará el caso de que, para no contravenir las dos condiciones anteriores, el espacio sobrante del último paquete de transporte correspondiente a un paquete PES, se rellene deliberadamente mediante un “campo de adaptación” cuya longitud será el complemento a 184 del número de bytes que queden por transmitir para terminar este último paquete PES, como se muestra en la figura 5.12

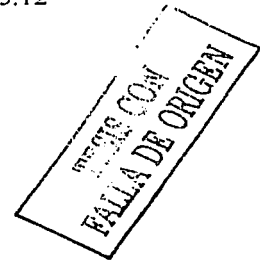
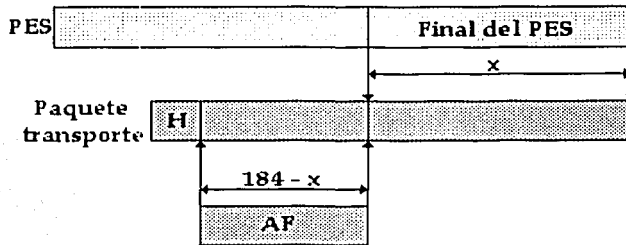


Figura 5.12 Constitución del último paquete de transporte de un PES

Además de esta función de relleno, el campo de adaptación se utilizara también para la transmisión del reloj de referencia del programa (PCR) que es un reloj con el cual se sincroniza todo un mismo programa, por lo que todo elemento perteneciente a dicho programa deberá contar forzosamente con el mismo valor de este reloj, cuyo ritmo de repetición mínimo es de 10 por segundo, así como diversos tipos de datos opcionales. Un paquete de transporte eventualmente podrá estar constituido únicamente por un campo de adaptación, de 184 bytes en este caso concreto (transporte de datos privados, PCR, etc.).

La figura 5.13 ilustra el formato general del campo de adaptación, cuyo contenido se especifica en la tabla 5.3

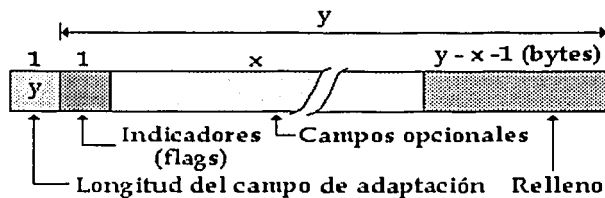


Figura 5.13 Detalle del campo de adaptación

Campo	Definición	Nº de bits
Adaptation field length	longitud total en bytes, menos 1 (y)	8
Flags	marcadores que indican la información transportada	8
optional fields	campos opcionales de datos (x bytes)	
Stuffing	campo de relleno (y-1-x bytes de valor FF hex)	

Tabla 5.3 Constitución del campo de adaptación MPEG-2

5.8 Información Específica de los Programas (PSI)

Según se ha visto, el flujo de transporte MPEG-2 puede contener varios programas audiovisuales, cada uno de los cuales está compuesto por uno o varios flujos elementales PES distribuidos en paquetes de transporte. Estos paquetes a su vez están marcados con un PID que identifica a qué flujo elemental pertenecen.

Sin embargo, para que el decodificador pueda recuperar completamente un programa a través de los valores de los PID de los paquetes correspondientes, es necesario incluir información adicional dentro del flujo de transporte que relacione estos PID con los programas a que pertenecen. Tal información se denomina "*Información Específica de los Programas (PSI)*".

Esta "*Información Específica de los Programas (PSI)*", definida por MPEG-2 para la Capa de Sistema (ISO/IEC 13818-1), comprende la inclusión dentro del flujo de transporte, de 4 tipos de tablas:

- *Tabla de asociación de programa (PAT)*
- *Tabla de acceso condicional (CAT)*
- *Tabla del mapa del programa (PMT)*
- *Tablas privadas*

Las Tablas que se adhieren en ATSC son:

- **Tabla de canal virtual terrestre (TVCT)**. Define el número mínimo de programas MPEG-2 dentro del TS.
- **Tabla guía maestra**. Define el tipo, los identificadores de paquete y las versiones para las otras tablas de información en el TS, excepto para la tabla del tiempo del sistema (STT).
- **Tabla de región de audiencia (RRT)**. Define el sistema de lineamientos de TV asociada, es referenciada por cualquier descriptor de ayuda transportado en el TS.

- **Tabla del tiempo del sistema (STT).** Define fecha y hora del día en curso.
- **Tabla de información de evento (EIT).** Son 4 tablas que describen 12 horas de programación televisivas, cada una con una cobertura de 3 horas incluyendo los canales virtuales contenidos en la tabla TVCT.

Cada tabla está constituida, según su importancia, por una o varias secciones (256 como máximo, con una longitud máxima de 1024 bytes, salvo para la tabla privada que puede alcanzar los 4026 bytes). Dichas secciones están distribuidas a lo largo de una serie de paquetes de transporte identificados con un PID común.

5.8.1 Tabla de asociación de programa (PAT)

Esta tabla, de inclusión obligatoria, es transportada por los paquetes con PID=0x0000 y contiene una lista completa de todos los programas disponibles en el flujo de transporte. Cada programa aparece junto con el valor del PID de los paquetes que a su vez contienen la tabla con los datos que identifican a dicho programa (PMT).

La PAT debe transmitirse sin cifrar aunque todos los demás programas lo estén.

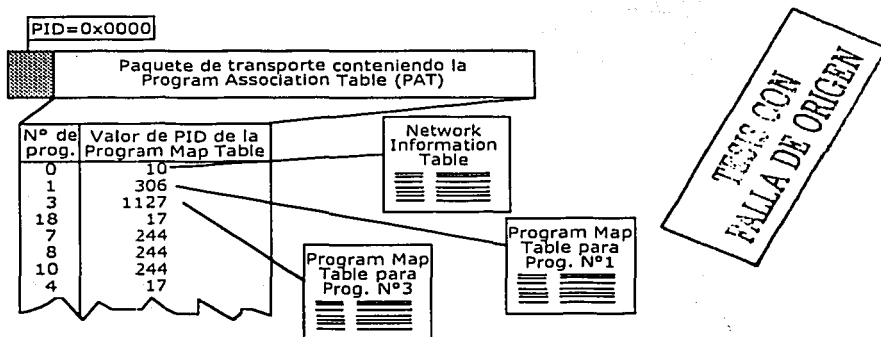


Figura 5.14 Representación de una tabla PAT

Como se muestra a modo de ejemplo en la figura 5.14, una sola tabla de asociación de programa (la nº 244) puede contener los detalles de varios programas (los numerados como 7, 8 y 10).

5.8.2 Tabla de acceso condicional (CAT)

Esta tabla debe estar presente si al menos un programa del múltiplex es de acceso condicional. Se transporta por los paquetes con PID=0x0001, y proporciona detalles de los sistemas de cifrado empleados, así como los valores de los PID de los paquetes de transporte que contienen la información del control de acceso condicional.

Los datos para el acceso condicional se envían en forma de mensajes de derecho de dirección (*EMM*). En estos "*EMM*" se especifican los niveles de autorización o los servicios a que pueden acceder determinados decodificadores, y pueden ir dirigidos a decodificadores individuales o a grupos de ellos.

El formato de esta información no está especificado en MPEG-2, puesto que depende del tipo de sistema de cifrado empleado.

5.8.3 Tabla del mapa del programa (PMT)

Cada programa audiovisual incluido en un flujo de transporte tiene una tabla *PMT* asociada con él. Dicha tabla proporciona detalles acerca del programa y de los flujos elementales que comprende.

Según MPEG-2, las tablas *PMT* pueden ser transportadas por paquetes con valores de PID arbitrarios, exceptuando los valores 0x0000, reservado para *PAT*, y 0x0001 reservado para *CAT*.

En la figura 5.15 se muestra un ejemplo de tabla *PMT*. Mediante ella, un decodificador puede determinar que el flujo elemental codificado de video correspondiente al programa nº 3 se encuentra en los paquetes de transporte identificados mediante el PID=726 y que el flujo elemental cuyos paquetes están etiquetados con PID=57 lleva el audio correspondiente en inglés.

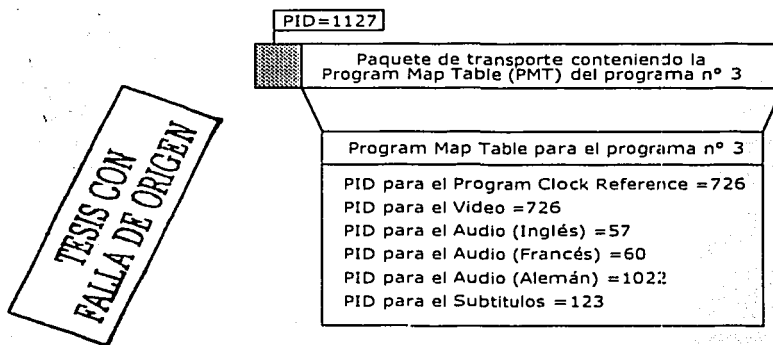


Figura 5.15 Ejemplo de una tabla *PMT*

En las tablas *PMT* los datos de los PID no pueden estar cifrados. Sin embargo, dichas tablas *PMT* también pueden contener información *privada* relativa al programa, que eventualmente si puede estar cifrada (por ejemplo datos para el control de acceso).

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

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

Las tablas 5.4 a la 5.7 muestran la estructura y los diferentes campos de las secciones de las 4 categorías de tablas (PAT, CAT, PMT Y PRIVATE) definidas por la norma MPEG-2.

Campo (PAT)	Comentario	Nº de bits
table_id (00)	siempre 0 para Program Allocation Table	8
section_syntax_indicator	siempre "1"	1
"0"		1
reserved		2
section_length	2 bits de mayor peso a "0", valor máx. 1021	12
transport_stream_id	identificación del tren en una red	16
reserved		2
version_number	incrementando cada modificación del PAT	5
current_next_indicator	"1"=PAT actual, "0"=próximo	1
section_number	número de la sección en curso (1ª=00)	8
last_section_number	número de la última sección (Ntot-1)	8
program_number 0 ⁽¹⁾	Prg 0= tabla de información de red (NIT)	16
reserved		3
network_PID	PID de tabla de información de red	13
program_number 1	número de programa (1 a 65535)	16
reserved		3
program_map_PID	PID de tabla de identificación de programa (PMT)	13
....etc	4 bytes por programa suplementario	...
CRC_32	CRC sobre 32 bits	32
(1) el programa 0 (tabla de información de red, o NIT) es opcional		

Tabla 5.4 Descripción de una sección de la tabla de asignación de programas (PAT)

Capítulo # 5

Campo (CAT)	Comentario	Nº de bits
table_id (01)	siempre 01 para Control Access Table	8
section_syntax_indicator	siempre "1"	1
"0"		1
reserved		2
section_length	2 bits de mayor peso a "0", valor máx. 1021	12
reserved		16
reserved		2
version_number	incrementando con cada modificación de CAT	5
current_next_indicator	"1"=CAT actual, "0"=próximo	1
section_number	número de sección en curso (1ª =00)	8
last_section_number	número de la última sección (Ntot-1)	8
descriptors	datos de control de acceso	máx. 1012 bytes
CRC 32	CRC sobre 32 bits	32

Tabla 5.5 Descripción de una sección de la tabla de control de acceso (CAT)

Campo (PMT)	Comentario	Nº de bits
Table_id (02)	siempre 02 para Program Map Table	8
Section_syntax_indicator	Siempre "1"	1
"0"		1
Reserved		2
Section_length	2 bits de mayor peso a "0", valor máx. 1021	12
Program_number	especifica el programa descrito a continuación	16
Reserved		2
Version_number	incrementando cada modificación de PMT	5
Current_next_indicator	"1"=MPT actual, "0"=próximo	1
Section_number	siempre 0 (una sola sección, identificada .../...)	8
last_section_number	.../... por el número de programa)	8
Reserved		3
PCR_PID	PID de la Program Clock Reference (PCR)	13
Reserved		4
Program_info_length	longitud total de los datos útiles (bytes)	12
Stream_type ₁	naturaleza del tren elemental N° 1	8
Reserved		3

Elementary PID ₁	PID del tren elemental N° 1	13
Reserved		4
ES_info_length ₁	N ₁ = N° de bytes de los descriptores siguientes	12
Descriptors ₁	datos complementarios	N ₁ bytes
Stream_type ₂	naturaleza del tren elemental N° 2	8
Reserved		3
Elementary PID ₂	PID del tren elemental N° 2	13
Reserved		4
ES_info_length ₂	N ₂ = N° de bytes de los descriptores siguientes	12
Descriptors ₂	datos complementarios	N ₂ bytes
..., etc. (ES N° x)		
CRC 32	CRC sobre 32 bits	32

Tabla 5.6 Descripción de una sección de la tabla de identificación de programación (PMT)

Campo (private)	Comentario	N° de bits
table_id	libre (excepto de 00h a 3Fh y FFh)	8
section_syntax_indicator	0=formato libre después de length, estándar	1
private_indicator	marcador (uso definido por el usuario)	1
reserved		2
private_section_length	2 bits de mayor peso a "1", valor max. 4093	12
private_data_byte(1)	datos privados	max. 4093 bytes

(1) aquí, la zona de datos privados está representada en formato libre (syntax_indicator=0). Si syntax_indicator=1, el formato de datos privados es parecido al de la CAT (excepto en la longitud).

Tabla 5.7 Descripción de una sección de datos privados (private section)

La figura 5.16 muestra, de qué forma se utilizan las tablas descritas anteriormente, constituyendo una PSI.

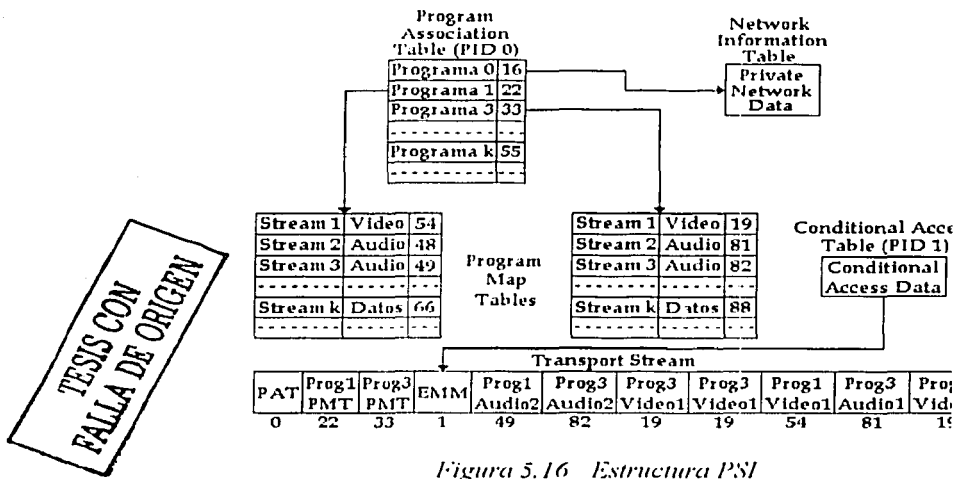


Figura 5.16 Estructura PSI

5.9 Inserción de las secciones en los paquetes de transporte

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

La posición de comienzo de una nueva sección en un paquete TS está indicada por el "Pointer_field", que debe ser el primer byte de la carga útil del citado paquete y proporciona el desplazamiento (offset), en bytes, desde el final del "pointer_field" hasta el primer byte de dicha sección. Así, un valor de 0x00 indica que la sección comienza inmediatamente después del "pointer_field". Si ninguna sección tiene su comienzo dentro del paquete TS, no habrá "pointer_field" y el "payload_unit_start_indicator" se pondrá a 0. Solamente hay un "pointer_field" en un paquete TS para marcar la posición de la primera sección que tenga su comienzo dentro de él, puesto que el comienzo de cualquier otra sección de una tabla se puede identificar fácilmente contando la longitud de las secciones precedentes, ya que la sintaxis no permite que existan huecos entre ellas.

La figura 5.17 ilustra esta situación en el caso general.

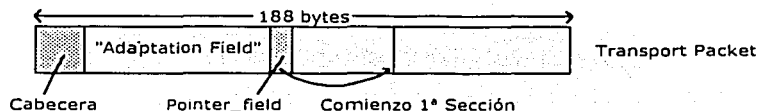


Figura 5.17 Primer comienzo de sección dentro de un paquete TS

Si una sección termina antes del final de un paquete de transporte, pero no es conveniente abrir otra dentro del mismo paquete, se debe emplear un mecanismo de relleno que emplee para cada byte sobrante el valor 0xFF. Consecuentemente, este valor no debe usarse para el byte que identifica el tipo de tabla: "*table_id*". No obstante, también puede usarse como método de relleno de los paquetes de transporte el mecanismo "Adaptation_field".

5.9.1 Composición de las secciones

Las secciones pueden tener longitud variable. Dentro de cada tabla, la longitud de las secciones está limitada a 1K (1024 bytes), excepto en el caso de las secciones de las tablas privadas cuya longitud está limitada a 4K (4096 bytes).

Cada sección se identifica por la combinación, entre otros, de los siguientes elementos:

table_id: El campo "*table_id*" identifica a qué tipo de tabla pertenece la sección.

Algunos valores de "*table_id*" han sido definidos por ISO y otros por ETSI. También el usuario puede asignar otros valores para uso privado.

table_id_extension: Se usa para identificar una sub_tabla.

version_number: Cuando las características descritas del flujo de transporte (TS) cambian; por ejemplo: comienzo de nuevos acontecimientos, diferente composición de los flujos elementales para un servicio dado, etc, deben transmitirse los nuevos datos de la información del servicio (SI) conteniendo la información actualizada. Esto se hace enviando una sub_tabla con los mismos identificadores que en las sub_tablas previas que contienen los datos relevantes, pero incrementando en 1 el valor previo del número de versión.

current_next_indicator: Cada sección debe numerarse como válida "ahora" o válida en un inmediato futuro. Esto permite la transmisión de una versión futura de la misma antes del cambio, dando al decodificador la oportunidad de prepararse para ello. No hay, sin embargo, ningún requerimiento para transmitir por adelantado la siguiente versión de una sección, pero si se hace, deberá ser la versión correcta inmediata de dicha sección.

section_number: Este campo (que se incrementa en 1 con cada sección adicional), permite que las diferentes secciones de una sub-tabla particular sean reensambladas en su orden original por el decodificador. No obstante, se recomienda que las secciones se transmitan en orden numérico, a menos que se quieran transmitir algunas secciones de la sub-tabla con más frecuencia que otras (por ejemplo debido a consideraciones de acceso aleatorio).

5.10 Cómo se realiza el demultiplexado MPEG-2

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

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

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

Después de elegir el programa:

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

Una vez hecha esta nueva elección.

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

La parte visible por el usuario de este proceso es la presentación interactiva de la "guía electrónica de programa (EPG)", para permitirle navegar fácilmente por los distintos programas y servicios que se le ofrecen.

5.11 Errores

La función principal de las técnicas de control de errores es introducir redundancia controlada permitiendo que los mensajes transmitidos que hayan sido corrompidos sean corregidos antes de ser procesados. Con esta redundancia controlada sólo un subconjunto de todos los posibles mensajes (secuencias de bits) contiene mensajes válidos. Esto quiere decir que si formamos secuencias de 7 bits, podremos recibir $2^7=128$ secuencias distintas, pero los posibles mensajes transmitidos no serán 128 sino un subconjunto de ellos, luego si no hemos recibido una secuencia de 7 bits que coincida con uno de los posibles mensajes transmitidos, habremos recibido una secuencia con errores.

Este subconjunto de posibles mensajes transmitidos se llama código y a los mensajes válidos que lo forman los llamamos codewords o codevectors.

Pues bien, un buen código es aquel en el que los codewords están tan separados que la probabilidad de que los errores conviertan al codeword deseado en otro codeword es muy pequeña.

Entonces la detección de errores se limita a responder a la pregunta: "¿Es el mensaje recibido un codeword o no?" Si es un codeword, uno supone que no hay errores, de esta manera la probabilidad de no detectar un error es la probabilidad de que los posibles errores conviertan un codeword en otro.

Cuando se detecta un error existen dos métodos para corregirlo:

1°.- El receptor rechaza el mensaje recibido y solicita al transmisor que retransmita el mensaje.

2°.- El receptor corrige el error al encontrar el codeword válido más próximo al mensaje recibido. Este procedimiento se llama *forward error correction* (FEC).

5.11.1 Distancia Hamming y distancia de código

Sean los mensajes secuencias binarias de n bits, esto hace que los mensajes o vectores de n bits formen un conjunto de 2^n posibles.

La distancia Hamming o distancia entre dos vectores es el número de posiciones de bits en los cuales dichos vectores toman valores diferentes.

Por ejemplo, sean 2 vectores de 4 bits:

$$c_1 = 0101 \text{ y } c_2 = 1100$$

la distancia entre c_1 y c_2 es 2.

Supongamos que de entre los 2^n posibles vectores elegimos un subgrupo que constituyen nuestro código. Pues la distancia del código será la mínima distancia Hamming entre todas las posibles combinaciones de codewords. Se demuestra que si la distancia de un código es d , entonces podemos detectar y corregir los errores que afecten a t bits de cada codeword, siendo:

$$d = 2*t + 1$$

Así si la distancia de un código es 5 entonces podemos detectar y corregir errores en 2 bits.

5.11.2 Códigos lineales y corrección de errores mediante el método del síndrome

Sea nuestro código formado por M codewords de los 2^n posibles.

Un código será lineal si dados dos codewords c_1 y c_2 cualesquiera de los M posibles, entonces $c = \alpha_1 c_1 + \alpha_2 c_2$ es un codeword, donde α_1 y α_2 pueden tomar valores 0 ó 1 verificándose las siguientes operaciones:

γ	0+0=0
γ	0+1=1+0=1
γ	1+1=0
γ	1* 1=1
γ	0*0=0
γ	1*0=0*1=0

Además, el vector formado por todos sus componentes 0 debe formar parte del código. Veamos un ejemplo :

Tomemos vectores de 8 bits y formaremos un código con 4 codewords que sea lineal :

(0 0 0 0 0 0 0 0) (0 1 1 1 1 1 0 0) (1 0 0 1 1 1 1 1) (1 1 1 0 0 0 1 1)

Se puede verificar que la combinación lineal de dos de ellos da otro :

$$1*(1 0 0 1 1 1 1 1) + 1*(1 1 1 0 0 0 1 1) = (0 1 1 1 1 1 0 0) \in \text{al código.}$$

Otra de las condiciones que cumple un código lineal es que :

$$\frac{2^k}{M} = \alpha \in \mathbb{N} - (0) \Rightarrow M = 2^k ; k \in \mathbb{N} - (0)$$

Luego esto hace pensar que realmente sólo hacen falta k bits para obtener M vectores distintos, pero se forman vectores con n bits para introducir la redundancia controlada de la que se hablo anteriormente.

Así que de los n bits, se puede decir que k pertenecen al mensaje que realmente queremos enviar y n-k bits forman los dígitos de comprobación. Un código así se llama código (n,k).

Todos los codewords se pueden obtener a partir de una matriz generatriz **G** :

$$c = m * G$$

donde **c** es el codeword de n bits del código (n,k), **m** es el mensaje que realmente nos interesa transmitir de los M posibles y como $M = 2^k$ implica que **m** se puede codificar con k bits (matriz 1xk) y **G** es una matriz con k filas y n columnas que transforma (codifica) nuestro mensaje **m** de k bits en un codeword de longitud n que será transmitido o almacenado.

Como **G** es una matriz (k x n) con k vectores linealmente independientes, aplicamos las reglas del álgebra y la escribimos de la forma $G = (I, P)$ donde **I** es la matriz identidad de (k x k) y **P** la matriz paridad de (k x (n-k)).

$$G = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & a_{k+1} & a_{k+2} & \dots & a_n \\ 0 & 1 & 0 & \dots & 0 & b_{k+1} & b_{k+2} & \dots & b_n \\ \dots & \dots & \dots & \dots & 0 & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & \dots & \dots & \dots & \dots \end{pmatrix}$$

entonces :

$$c = (m_1 \ m_2 \ \dots \ m_k) * G = (m_1 \ m_2 \ \dots \ m_k \ q_{k+1} \ q_{k+2} \ \dots \ q_n)$$

Luego se evidencia que el codeword está formado por k bits de nuestro mensaje y n-k bits.

Sea una matriz generadora :

$$G = \left[\begin{array}{cccccccc|c} 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{array} \right]$$

donde se observa que I es de 2 x 2 luego k=2 y n=8, con esto M=2^k=2²=4 codewords que forman el código.

Para obtenerlo probamos con todas las combinaciones posibles :

$$m = (00) \Rightarrow c_0 = (00000000)$$

$$m = (01) \Rightarrow c_1 = (01111100)$$

$$m = (10) \Rightarrow c_2 = (10011111)$$

$$m = (11) \Rightarrow c_3 = (11100011)$$

Se puede observar que la distancia de este código es d=5 y por tanto d=2*t+1 => t=2, luego se pueden detectar y corregir 2 bits erróneos.

Veamos cómo se detectan y corrigen estos errores. Para ello obtenemos la matriz nula de la siguiente manera :

Si $G = (I_k, P)$ entonces $H = (P^T, I_{n-k})$ de forma que si P es de $(k \times (n-k))$ I_k de $(k \times k)$, entonces P^T es de $((n-k) \times k)$ obligando a I_{n-k} a ser de $((n-k) \times (n-k))$ y por tanto H es de $((n-k) \times n)$. Pues bien se puede comprobar que $G * H^T = 0$.

Llamemos síndrome del codeword c a :

$$s = c * H^T$$

pero teniendo en cuenta que $c = m * G$ entonces :

$$s = m * G * H^T$$

pero como $\mathbf{G} * \mathbf{H}^T = \mathbf{0}$ entonces :

$$\mathbf{s} = \mathbf{m} * \mathbf{0} = \mathbf{0}$$

luego se observa que el síndrome de un codeword es nulo.

Imaginemos ahora que el vector transmitido es el codeword \mathbf{c} y el recibido es \mathbf{c} más un error \mathbf{e} , o sea $\mathbf{r} = \mathbf{c} + \mathbf{e}$, calculemos el síndrome de \mathbf{r} :

$$\mathbf{s} = \mathbf{r} * \mathbf{H}^T = (\mathbf{c} + \mathbf{e}) * \mathbf{H}^T = \mathbf{c} * \mathbf{H}^T + \mathbf{e} * \mathbf{H}^T = \mathbf{e} * \mathbf{H}^T + \mathbf{0}$$

Luego, cuando el síndrome de un vector recibido no es nulo, entonces, hemos detectado un error.

Una vez detectado, la corrección se obtiene haciendo :

$$\mathbf{c} = \mathbf{r} + \mathbf{e}$$

como sabemos que podemos corregir t bits erróneos del codeword como máximo, conocemos el tipo de errores que podemos corregir y los síndromes que darían a priori por lo que podemos identificar un síndrome con un error. El síndrome lo obtenemos de $\mathbf{s} = \mathbf{r} * \mathbf{H}^T$ miramos en una tabla a qué error corresponde y corregimos el vector mediante $\mathbf{c} = \mathbf{r} + \mathbf{e}$.

5.11.3 Bloque FEC de ATSC

Para la norma ATSC, se utiliza la corrección de errores FEC, por lo que a continuación se estudiara a fondo dicho concepto.

Los cinco diferentes FEC, son códigos que varían según el código de redundancia que queramos añadir a los datos que transporta el flujo MPEG-2.

Así un operador que utiliza normalmente un FEC de 5/6, puede que opte para la emisión de un importante encuentro de fútbol en que gran cantidad de personas estarán atentos al mismo por un código de corrección de errores más robusto de por ejemplo de 2/3 y así poder solventar problemas en su recepción debido a una fuerte lluvia en la zona. Así pues el operador depende de la importancia de la información que emita, tendrá que asegurarse que la información redundante es suficientemente alta, para no tener problemas en la transmisión.

Como ejemplo 3/4 significa que el 25% de los bits corresponden a información redundante para evitar que haya errores en la transmisión. Normalmente esta corrección, se adopta dependiendo de las condiciones hidrometeorológicas de la área de recepción. El emisor puede optar por un código de corrección que solo le produzca unos minutos de pérdidas de imagen en un año, causados por una fuerte lluvia, cosa que puede ser aceptable por el televidente.

Esto supone que al aumentar la información para corregir errores en una emisión nuestro ancho de banda aumente y por lo tanto el disponible disminuye.

El bloque FEC en ATSC se compone de 4 bloques básicamente, los cuales se muestran en la figura 5.18:

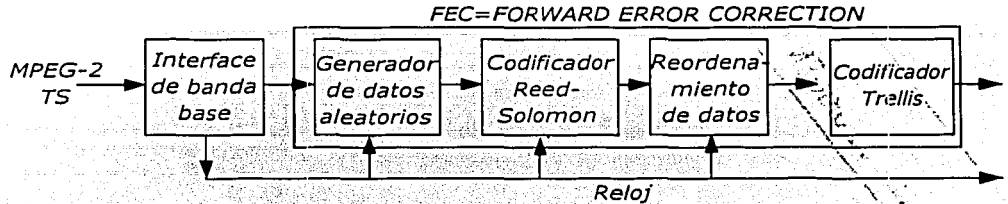


Figura 5.18 Bloque FEC en ATSC

A continuación se describirá todas las características de los procesos que siguen los paquetes de transporte al salir del codificador MPEG-2:

5.11.4 Interfase en banda base

Se refiere a los paquetes de transporte obtenidos de la codificación MPEG-2, que a lo largo de este capítulo se estudiaron.

COPIA CON
SELLA DE ORIGEN

5.11.5 Generador de datos aleatorios

Los datos son alimentados al generador de datos aleatorios para fragmentar largas secuencias de 1s o 0s que quizá puede contener los paquetes de transporte, sin incluir el bit de sincronización de cada paquete. Los datos son aleatorizados por medio de compuertas XORs, realizando dicha operación entre los datos y una secuencia binaria pseudo aleatoria (PRBS), la cual consiste en un registro de cambio con realimentación de 16 bits; dichos bits son conocidos como la palabra inicial que se define en el intervalo de sincronización.

El propósito de esta técnica es distribuir de mejor manera la energía presente en el espectro disponible para cada canal. Si se presentan patrones definidos en el tren de bits, entonces algunos puntos discretos del espectro de frecuencia tendrán más amplitud que otros, lo cual implica que la energía no se distribuya de manera uniforme en el espectro, creando picos de energía que exceden los límites permitidos además de generar una posible pérdida de sincronía.

De forma grafica, se presenta la figura 5.19:

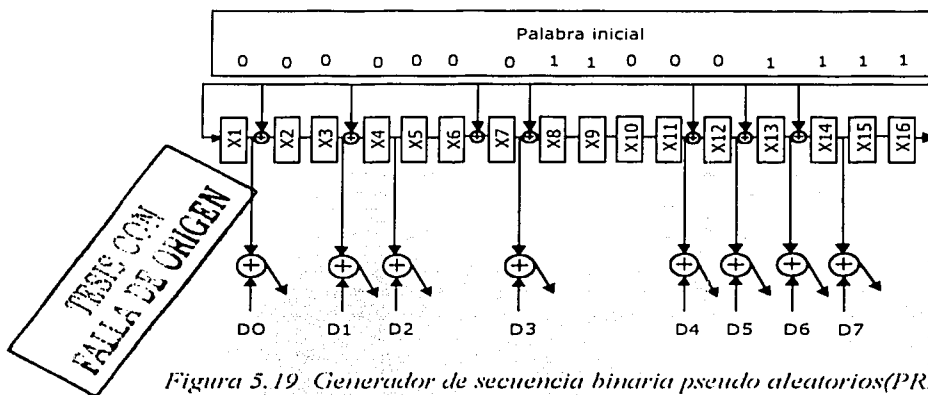


Figura 5.19 Generador de secuencia binaria pseudo aleatorios(PRBS)

5.11.6 Codificador Reed-Solomon

El codificador Reed-Solomon toma un bloque de información digital y añade bits redundantes. Los errores pueden ocurrir durante la transmisión o almacenamiento de información por varios motivos (p. Ej. Ruido o interferencia, ralladuras en los discos compactos etc.):

El decodificador Reed-Solomon procesa cada bloque e intenta corregir los errores y recuperar la información original. El número y tipo de errores que pueden ser corregidos depende de las características del código Reed-Solomon.

Un código binario natural de 3 bits asigna los siguientes pesos decimales $(2^2, 2^1, 2^0) = (4, 2, 1)$ así el vector $(1, 0, 1)$ se representa como 2^2+1 , $(1, 1, 0)$ como 2^2+2 , pues bien, si en vez de emplear como base el 2 empleamos α , entonces los pesos serán (α^2, α^1) así $(1, 0, 1)$ se representa por α^2+1 , $(1, 1, 0)$ será $\alpha^2+\alpha$, así que podemos representar cualquier secuencia de tres bits con $(\alpha^2, \alpha, 1)$.

Por otro lado, en un campo finito o de Galois⁽¹⁾ (GF) se definen las siguientes operaciones:

$$(\text{elemento del campo}) + (\text{el mismo elemento}) = 0$$

lo que significa que si α^3 es un elemento del campo, entonces $\alpha^3 + \alpha^3 = 0 \Rightarrow \alpha^3 = -\alpha^3$ por lo que en las operaciones con elementos del campo finito, las operaciones que impliquen un cambio de signo, las sustituimos por los mismos operandos con signos positivos.

Si en dicho campo se definen α y un polinomio $p(x)$ donde α es la raíz de $p(x)$, se pueden obtener valores de α^i con dicho polinomio aplicando las reglas anteriores, por ejemplo sea:

$$p(x) = x^2 + x + 1$$

y α una raíz suya, entonces :

$$p(\alpha) = \alpha^2 + \alpha + 1 = 0$$

$$-\alpha^2 = \alpha^2 = \alpha + 1$$

multiplicamos por α

$$\alpha^3 = \alpha^2 + \alpha = \alpha + 1 + \alpha_{\alpha+\alpha=0} = 1$$

multiplicamos por:

$$\alpha^4 = \alpha$$

multiplicamos por α :

$$\alpha^5 = \alpha^2 = \alpha + 1$$

y a partir de aquí se repite :

$$\alpha^6 = \alpha^2 = \alpha + 1$$

Si un polinomio tiene por raíces por ejemplo 1, α , α^2 éste será :

$$p(x) = (x+1)(x+\alpha)(x+\alpha^2)$$

Este tipo de códigos emplea algo de todo lo visto hasta ahora para llevar a cabo la detección y corrección de errores.

Sus elementos forman parte de un campo finito donde rigen reglas como las vistas en el apartado anterior. Parten de una raíz α , un polinomio y un campo finito de dimensión $q=2^m$, o sea $GF(q) = GF(2^m)$, en este caso codificamos los bits según los pesos ($\alpha^{m-1} \alpha^{m-2} \dots \alpha^1$).

Veamos cómo se codifican y decodifican códigos Reed-Solomon con un ejemplo.

Supongamos un campo finito de 8 elementos, o sea $GF(2^3)$ luego los pesos serán ($\alpha^2 \alpha^1$) una tripleta de este tipo se considera un símbolo para distinguirlo de un bit. Por tanto, en nuestro ejemplo un símbolo es de 3 bits, así que si tenemos los siguientes símbolos ($\alpha^1 0 \alpha^2$) realmente tendríamos la siguiente secuencia de bits (010 001 000 100) por tanto 4 símbolos de nuestro ejemplo suponen $4 \times 3 = 12$ bits. Pues bien, los códigos Reed-Solomon detectan símbolos erróneos y corrigen símbolos, lo que significa que en este caso, se pueden corregir dos símbolos, o 6 bits erróneos, pero no nos confundamos, no 6 bits cualesquiera, tendrían que pertenecer a dos símbolos diferentes ya que podría darse el caso de 6 bits erróneos pero cada uno en un símbolo lo que supondría corregir 6 símbolos. Para la secuencia de 4 símbolos anterior, dos símbolos erróneos podrían ser (001 001 000 010) donde el primer símbolo ha pasado de α a 1 y el último de α^2 a α , se observa que han cambiado 4 bits pero en 2 símbolos. Entonces es posible corregir símbolos y no bits.

En los códigos Reed-Solomon la distancia de código viene dada por la fórmula $d=m+1=n-k+1$ donde m es el grado de un polinomio generador $g(x)$, k viene dada por el exponente de la base 2 en las dimensiones del campo de Galois, o sea $GF(2^k)$. Así que si nos dicen que $g(x)$ tiene grado $m=4$ y los elementos del código se toman de un campo finito

Capítulo # 5

$GF(2^3)$ entonces $k=3$ y como $m=n-k$, obtenemos $n=7$, por tanto el código será $(7,3)$. Que significa que empleamos 7 símbolos y 3 son para el mensaje que queremos transmitir o almacenar. Para obtener con cuantos bits codificamos cada símbolo empleamos otro polinomio y una raíz de él.

Con el polinomio generatriz $g(x)$ se puede obtener la matriz generatriz G , veamos cómo con un ejemplo. Sea $g(x) = x^4 + x^3 + x^2 + 1$, suponiendo un código $(7,3)$, las filas de G vienen dadas por la fórmula :

$$\text{fila}(j) = x^{n-j} + r_{n-j}(x) \quad 1 \leq j \leq k$$

donde $r_j(x)$ es el residuo de dividir x^j entre $g(x)$.

Para este ejemplo :

$$\text{fila}(1) = x^{7-1} + r_{7-1}(x) = x^6 + r_6(x)$$

$$r_6(x) = x^3 + x^2 + x$$

La fila (1) = $x^6 + x^3 + x^2 + x$ se traduce en :

$$\text{fila}(1) = \begin{pmatrix} x^6 & x^5 & x^4 & x^3 & x^2 & x^1 & x^0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\text{fila}(2) = x^5 + r_5(x)$$

$$r_5(x) = x^2 + x + 1$$

La fila (2) = $x^5 + x^2 + x + 1$ se traduce en :

$$\text{fila}(2) = (0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1)$$

$$\text{fila}(3) = x^4 + r_4(x)$$

$$r_4(x) = x^3 + x^2 + 1$$

La fila (3) = $x^4 + x^3 + x^2 + 1$ se traduce en :

$$\text{fila}(3) = (0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1)$$

Luego G queda :

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Siendo

$$P = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

Entonces se puede obtener H

$$H = (P^T, I) = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

En el caso de códigos Reed-Solomon el polinomio generatriz $g(x)$ es del tipo :

$$g(x) = \prod_{i=0}^{m-1} (x - \alpha^i)$$

Sean las raíces del polinomio generatriz $g(x)$ $1, \alpha, \alpha^2, \alpha^3$, cuyos coeficientes están definidos en un campo finito $GF(2^3)$ por lo que $k=3$; y sea el polinomio que define la base x^3+x+1 cuya raíz primitiva es α . Veamos el tipo de código que definen estos y como se codifican, detectan y corrigen errores :

El polinomio x^3+x+1 tiene por raíz α por lo que $\alpha^3+\alpha+1=0$. La base llegará como máximo hasta α^2 , ya que α^3 se puede poner en función de elementos de menor grado porque $-\alpha^3 = \alpha^3 = \alpha + 1$, luego obtenemos que los pesos son $(\alpha^2 \ \alpha \ 1)$ que nos indica que codificamos los símbolos con tres bits.

$$\alpha^3 = \alpha + 1$$

multiplicando por α :

$$\alpha^4 = \alpha^2 + \alpha$$

multiplicando por α y realizando operaciones :

$$\alpha^5 = \alpha^3 + \alpha^2 = \alpha^2 + \alpha + 1$$

Capítulo # 5

multiplicando por α y realizando operaciones :

$$\alpha^6 = \alpha^3 + \alpha^2 + \alpha = \alpha + 1 + \alpha^2 + \alpha = \alpha^2 + 1$$

multiplicando por α :

$$\alpha^7 = \alpha^3 + \alpha = \alpha + 1 + \alpha = 1$$

multiplicando por α :

$$\alpha^8 = \alpha$$

multiplicando por α :

$$\alpha^9 = \alpha^2$$

multiplicando por α :

$$\alpha^{10} = \alpha^3 = \alpha + 1$$

y así sucesivamente.

$$g(x) = \prod_{i=0}^{m-1} (x + \alpha^i) = (x+1) * (x+\alpha) * (x+\alpha^2) * (x+\alpha^3) = (x^2 + (\alpha+1)x + \alpha) * (x^2 + (\alpha^3 + \alpha^2)x + \alpha^5)$$

Para obtener $g(x)$:

$$\alpha + 1 = \alpha^3; \alpha^3 + \alpha^2 = \alpha + 1 + \alpha^2 = \alpha^5$$

$$g(x) = (x^2 + \alpha^3 x + \alpha) (x^2 + \alpha^5 x + \alpha^5) = x^4 + \alpha^5 x^3 + \alpha^5 x^2 + \alpha^3 x^3 + \alpha^8 x^2 + \alpha^8 x + \alpha x^2 + \alpha^6 x + \alpha^6$$

$$g(x) = x^4 + (\alpha^5 + \alpha^3) x^3 + (\alpha^5 + \alpha^8 + \alpha) x^2 + (\alpha^8 + \alpha^6) x + \alpha^6$$

$$\alpha^5 + \alpha^3 = \alpha^2 + \alpha + 1 + \alpha + 1 = \alpha^2$$

$$\alpha^5 + \alpha^8 + \alpha = \alpha^5 + \alpha + \alpha = \alpha^5$$

$$\alpha^8 + \alpha^6 = \alpha + \alpha^2 + 1 = \alpha^5$$

$$g(x) = x^4 + \alpha^2 x^3 + \alpha^5 x^2 + \alpha^5 x + \alpha^6$$

bueno ya tenemos nuestro polinomio generatriz de grado $m=4$ por tanto $m=n-k \Rightarrow n=7$ y la distancia del código será $d=m+1=4+1=5$. Se trata de un código (7,3) pero ahora son símbolos y no bits, esto es, cada codeword es de 7 símbolos, de los cuales 3 símbolos son el mensaje que interesa codificar, cada símbolo es de 3 bits según se obtiene de los pesos ($\alpha^2 \alpha 1$) por tanto, cada codeword consta de $7 \times 3 = 21$ bits.

Obtengamos la matriz generatriz según el procedimiento ya descrito :

$$\text{fila (1)} = x^6 + r_6(x)$$

$$\text{fila(1)} = x^6 + \alpha x^3 + \alpha^3 x^2 + \alpha^6 x + \alpha^6$$

$$\text{fila(1)} = (1, 0, 0, \alpha, \alpha^3, \alpha^6, \alpha^6)$$

$$\text{fila (2)} = x^5 + r_5(x)$$

$$\text{fila (2)} = x^5 + x^3 + \alpha^4 x^2 + \alpha^2 x + \alpha$$

$$\text{fila (2)} = (0 \ 1 \ 0 \ 1 \ \alpha^4 \ \alpha^2 \ \alpha)$$

$$\text{fila(3)} = x^4 + r_4(x)$$

$$\text{fila (3)} = x^4 + \alpha^2 x^3 + \alpha^5 x^2 + \alpha^5 x + \alpha^6$$

$$\text{fila (3)} = (0 \ 0 \ 1 \ \alpha^2 \ \alpha^5 \ \alpha^5 \ \alpha^6)$$

Así la matriz generatriz es :

$$G = \begin{vmatrix} 1 & 0 & 0 & \alpha & \alpha^3 & \alpha^6 & \alpha^6 \\ 0 & 1 & 0 & 1 & \alpha^4 & \alpha^2 & \alpha \\ 0 & 0 & 1 & \alpha^2 & \alpha^5 & \alpha^5 & \alpha^6 \end{vmatrix}$$

y el código se obtiene de $c = m * G$ donde m es el mensaje que se quiere transmitir o almacenar con $k=3$ símbolos, esto es $3 \times 3 = 9$ bits. Si queremos codificar el siguiente mensaje :

$$m = (001 \ 101 \ 111)$$

que con símbolos es :

$$001 = 1$$

$$101 = \alpha^2 + 1 = \alpha^3$$

$$111 = \alpha^2 + \alpha + 1 = \alpha^5$$

$$m = (1 \ \alpha^3 \ \alpha^5)$$

da como resultado el siguiente codeword :

$$c = m * G = [1, \alpha^3, \alpha^5] * \begin{vmatrix} 1 & 0 & 0 & \alpha & \alpha^3 & \alpha^6 & \alpha^6 \\ 0 & 1 & 0 & 1 & \alpha^4 & \alpha^2 & \alpha \\ 0 & 0 & 1 & \alpha^2 & \alpha^5 & \alpha^5 & \alpha^6 \end{vmatrix} = [1, \alpha^3, \alpha^5, 0, 1, 1, \alpha^6]$$

Éste sería el codeword transmitido o almacenado.

Para obtener la matriz nula H :

$$H = \begin{pmatrix} \alpha & 1 & \alpha^2 & 1 & 0 & 0 & 0 \\ \alpha^3 & \alpha^4 & \alpha^5 & 0 & 1 & 0 & 0 \\ \alpha^6 & \alpha^2 & \alpha^5 & 0 & 0 & 1 & 0 \\ \alpha^6 & \alpha & \alpha^6 & 0 & 0 & 0 & 1 \end{pmatrix}$$

el número de símbolos que se pueden corregir con este código se puede calcular con:

$$d=m+1=2t+1=n-k+1$$

$$2t=4$$

$t=2$ símbolos como máximo se pueden corregir.

Suponiendo que el codeword transmitido o almacenado es

$$c = (1 \alpha^3 \alpha^5 0 1 1 \alpha^6) = (001 011 111 000 001 001 101)$$

y se tiene un error en el segundo símbolo y otro en el sexto símbolo, de manera que se recibe :

$$r = (001 100 111 000 001 100 101)$$

$$r = (1 \alpha^2 \alpha^5 0 1 \alpha^2 \alpha^6)$$

Calculando el síndrome $s = (S_1 S_2 S_3 S_4) = r * H^T$, la matriz nula que se empleará es la obtenida anteriormente pero modificada de acuerdo con las reglas del álgebra matricial, convirtiéndola en otra más adecuada para hacer los cálculos. También se puede obtener la forma de la matriz nula directamente mediante la fórmula :

$$H = [(\alpha^i)^{j-1}, (\alpha^i)^{j-2}, \dots, (\alpha^i)^2, (\alpha^i), 1]$$

donde α^i son todas y cada una de las raíces de $g(x)$, así que para nuestro ejemplo, las raíces son $1, \alpha, \alpha^2, \alpha^3$ con lo cual H queda transformada en :

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha^6 & \alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 & \alpha & 1 \\ \alpha^{12} & \alpha^{10} & \alpha^8 & \alpha^6 & \alpha^4 & \alpha^2 & 1 \\ \alpha^{18} & \alpha^{15} & \alpha^{12} & \alpha^9 & \alpha^6 & \alpha^3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha^6 & \alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 & \alpha & 1 \\ \alpha^5 & \alpha^3 & \alpha & \alpha^6 & \alpha^4 & \alpha^2 & 1 \\ \alpha^4 & \alpha & \alpha^5 & \alpha^2 & \alpha^6 & \alpha^3 & 1 \end{pmatrix}$$

De esta manera y empleando esta forma de la matriz nula calculamos los componentes del síndrome del vector recibido, obteniendo :

$$S_1 = 1 + \alpha^2 + \alpha^5 + 1 + \alpha^2 + \alpha^6 = \alpha$$

$$S_2 = \alpha^6 + \alpha^7 + \alpha^9 + \alpha^2 + \alpha^3 + \alpha^6 = 1 + \alpha^2 + \alpha^2 + \alpha^3 + \alpha^6 = 1 + \alpha^2 + \alpha^2 + \alpha^3 = 1 + \alpha + 1 = \alpha$$

$$S_3 = \alpha^5 + \alpha^5 + \alpha^6 + \alpha^4 + \alpha^4 + \alpha^6 = 0$$

$$S_4 = \alpha^4 + \alpha^3 + \alpha^{10} + \alpha^6 + \alpha^5 + \alpha^6 = \alpha^2 + \alpha + \alpha^2 + \alpha + 1 = 1$$

Ahora se localizan los símbolos erróneos, para ello se sigue un procedimiento que implica resolver uno sistemas de ecuaciones, el primero de ellos sigue la fórmula general :

$$S_{t+j} = f_1 * S_{t+j-1} + \dots + f_{t-1} * S_{j+1} + f_t * S_j$$

donde $1 \leq j \leq t$ los f_i son las incógnitas y S_i los componentes del síndrome calculado.

Para este caso particular $t=2$, $1 \leq j \leq 2$ por lo que el sistema de ecuaciones queda :

$$S_3 = f_1 * S_2 + f_2 * S_1$$

$$S_4 = f_1 * S_3 + f_2 * S_2$$

resolviendo por Cramer^[2], teniendo en cuenta la aritmética en un campo finito.

$$\begin{vmatrix} s_2 & s_1 \\ s_3 & s_2 \end{vmatrix} = s_2 * s_2 + s_1 * s_3 = \alpha * \alpha + \alpha * 0 = \alpha^2$$

$$f_1 = \frac{\begin{vmatrix} s_3 & s_1 \\ s_4 & s_2 \end{vmatrix}}{\alpha^2} = \frac{0 * \alpha + \alpha * 1}{\alpha^2} = \frac{\alpha}{\alpha^2} = \frac{\alpha^8}{\alpha^2} = \alpha^6$$

$$f_2 = \frac{\begin{vmatrix} s_2 & s_3 \\ s_3 & s_4 \end{vmatrix}}{\alpha^2} = \frac{\alpha * 1 + 0}{\alpha^2} = \frac{\alpha}{\alpha^2} = \frac{\alpha^8}{\alpha^2} = \alpha^6$$

La posición en la que están los errores se obtiene de los exponentes de las raíces de un polinomio $f(x)$ cuya fórmula general es :

$$f(x) = x^t + f_1 x^{t-1} + \dots + f_{t-1} x + f_t$$

que en este caso particular, se obtiene :

$$f(x) = x^2 + f_1 x + f_2$$

sustituyendo valores

$$f(x) = x^2 + \alpha^6 x + \alpha^6$$

luego se obtienen las raíces de $f(x)$ por tanteo :

Capítulo # 5

para $x=\alpha$; $f(\alpha) = \alpha^2 + \alpha^6 \alpha + \alpha^6 = \alpha^2 + \alpha^7 + \alpha^6 = \alpha^2 + 1 + \alpha^2 + 1 = 0$ luego $x=\alpha$ es una raíz de $f(x)$

para $x=\alpha^2$; $f(\alpha^2) = \alpha^4 + \alpha^6 \alpha^2 + \alpha^6 = \alpha^4 + \alpha^8 + \alpha^6 = 1 \alpha 0$

para $x=\alpha^3$; $f(\alpha^3) = \alpha^6 + \alpha^6 \alpha^3 + \alpha^6 = \alpha^2 \alpha 0$

para $x=\alpha^4$; $f(\alpha^4) = \alpha^8 + \alpha^6 \alpha^4 + \alpha^6 = \alpha + \alpha^{10} + \alpha^6 = \alpha + \alpha^3 + \alpha^6 = \alpha^2 \alpha 0$

para $x=\alpha^5$; $f(\alpha^5) = \alpha^{10} + \alpha^6 \alpha^5 + \alpha^6 = \alpha^3 + \alpha^{11} + \alpha^6 = \alpha^3 + \alpha^4 + \alpha^6 = \alpha + 1 + \alpha^2 + \alpha + \alpha^2 + 1 = 0$ por tanto la otra raíz es $x=\alpha^5$.

Pues bien, los símbolos erróneos son arrojados por los exponentes de las raíces de la siguiente manera :

$x=\alpha \Rightarrow (n - \text{exponente de la raíz}) = (7 - 1) = 6 \Rightarrow$ el sexto símbolo desde la izquierda es erróneo

donde

$$(1^{\circ} 2^{\circ} 3^{\circ} 4^{\circ} 5^{\circ} 6^{\circ} 7^{\circ})$$

$$(1 \alpha^2 \alpha^5 0 1 \alpha^2 \alpha^6)$$

$x=\alpha^5 \Rightarrow (7 - 5) = 2 \Rightarrow$ el segundo símbolo desde la izquierda es erróneo.

Ahora se sabe que los dos símbolos α^2 son erróneos, ¿pero cuál es su valor correcto ?

Los pasos que hay que seguir son los siguientes :

Se resuelve un sistema de ecuaciones con tantas incógnitas como errores hayamos detectado y que tiene la siguiente forma para este ejemplo :

$$Y_1 + Y_2 = S_1$$

$$x_1 Y_1 + x_2 Y_2 = S_2$$

donde Y_1 indica el valor del error en el símbolo dado por ($n - \text{exponente de } x_1$), x_1 es la raíz de $f(x)$, Y_2 indica el valor del error en el símbolo dado por ($n - \text{exponente de } x_2$), x_2 es la otra raíz de $f(x)$.

Así

$$Y_1 + Y_2 = \alpha$$

$$\alpha^5 Y_1 + \alpha Y_2 = \alpha$$

y se resuelve

$$\begin{vmatrix} 1 & 1 \\ \alpha^5 & \alpha \end{vmatrix} = \alpha + \alpha^5 = \alpha + \alpha^2 + \alpha + 1 = \alpha^2 + 1 = \alpha^6$$

$$Y_1 = \frac{\begin{vmatrix} \alpha & 1 \\ \alpha & \alpha \end{vmatrix}}{\alpha^6} = \frac{\alpha^2 + \alpha}{\alpha^6} = \frac{\alpha^4}{\alpha^6} = \frac{\alpha^{11}}{\alpha^6} = \alpha^5$$

$$Y_2 = \frac{\begin{vmatrix} 1 & \alpha \\ \alpha^5 & \alpha \end{vmatrix}}{\alpha^6} = \frac{\alpha + \alpha^6}{\alpha^6} = \frac{\alpha^5}{\alpha^6} = \frac{\alpha^{12}}{\alpha^6} = \alpha^6$$

Ahora que ya se tiene el error cometido en los símbolos, el vector error será :

$$e = (0 \alpha^5 0 0 0 \alpha^6 0)$$

y el codeword recuperado será :

$$c = e + r$$

$$(1 \alpha^2 \alpha^5 0 1 \alpha^2 \alpha^6) + (0 \alpha^5 0 0 0 \alpha^6 0)$$

$$c = (1 (\alpha^2 + \alpha^5) \alpha^5 0 1 (\alpha^2 + \alpha^6) \alpha^6) = (1 \alpha^3 \alpha^5 0 1 1 \alpha^6)$$

que efectivamente se corresponde con nuestro codeword original.

5.11.7 Reordenamiento de datos

El reordenamiento se usa para hacer mas efectiva la técnica de codificación Reed-Solomon, ya que el objetivo es esparcir los errores a otros bloques de datos, que se pueden provocar por una ráfaga grande de errores debido a una fuente de ruido.

El reordenamiento que se emplea en la transmisión VSB tiene 52 registros de almacenamiento. El reordenamiento debe tener una duración de aproximadamente 1/6 del campo de datos (4ms. de duración) Solo los bytes de datos deben de estar reordenados. El reordenador está sincronizado con el primer byte de cada campo. El objetivo es lograr que cada byte tenga un cierto retardo de acuerdo a su posición, cada secuencia de bytes es almacenada temporalmente hasta que llega su turno de ser transmitido en la secuencia ciclica. Al final el número de bytes en cada paquete se mantiene constante.

Está técnica es el preámbulo de la Codificación Trellis. La figura 5.20 muestra de manera gráfica todo el proceso

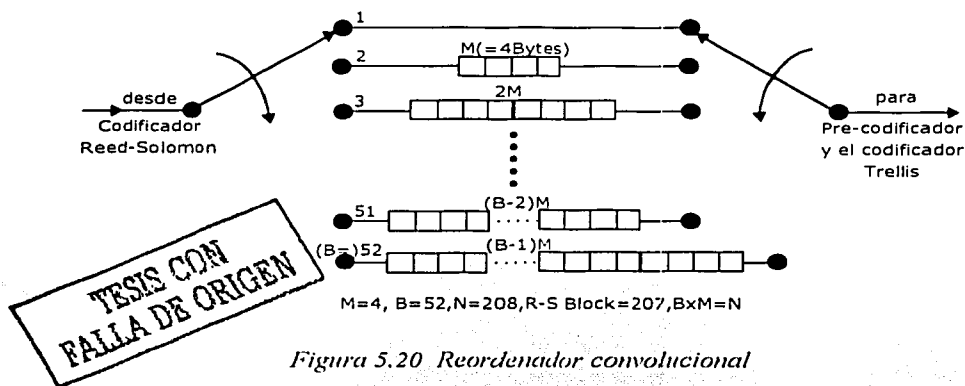


Figura 5.20 Reordenador convolucional

Los bytes de entrada son retrasados 4, 8, 12, ... 204 bytes, dependiendo del índice de cada byte. El primer byte de sincronía no es retrasado porque pasa directamente, con el fin de mantener el orden de dichos bytes y no perder la sincronía.

5.11.8 Codificador Trellis

La finalidad de este código es la introducción de un bit de redundancia por cada dos bits de información, o sea una codificación 2/3, para que así se fortalezca la función de corrección de errores. Básicamente se tiene el fundamento en el codificador convolucional^[3] ya que añade una estructura a la secuencia de bits. Incluso aunque la entrada sea totalmente aleatoria, se fuerza a que la salida siga unas determinadas secuencias. Esta restricción es la que da la capacidad correctora a los códigos convolucionales. El procedimiento de decodificación es equivalente a comparar la secuencia recibida con todas las posibles secuencias que pueden obtenerse con el correspondiente codificador y seleccionando la secuencia que está más próxima a la secuencia recibida.

Este codificador realiza las siguientes funciones:

Entrada de dos bits en paralelo. Se codifica en forma diferencial una de las líneas de datos. La idea es codificar sólo la diferencia entre el valor actual de la señal y una predicción del valor (señal diferencial). Esta codificación es adecuada para señales en las que los valores sucesivos difieren poco unos de otros, ya que esto hace posible detectar mejor la presencia de ruido. Ver figura 5.21.

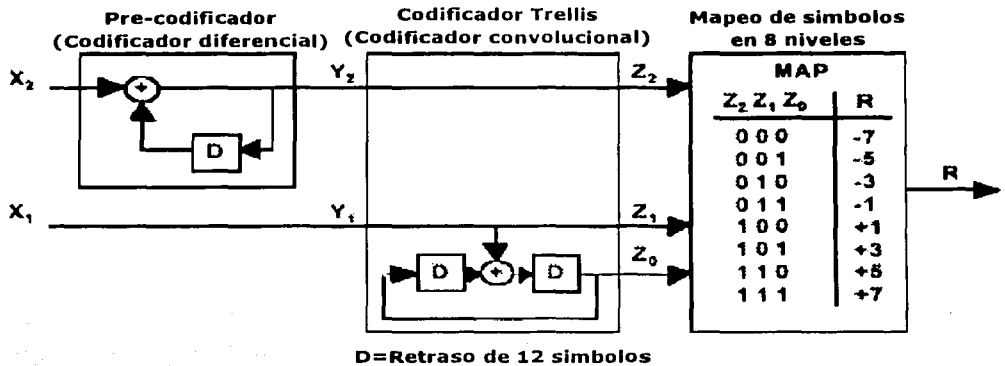


Figura 5.21 Codificador Trellis, precodificador y mapeador de símbolos

Se almacena el bit anterior siendo este la predicción del valor. Cuando llega un nuevo bit, solo realiza la diferencia entre estos dos, mandando el resultado al mapeador^[4] y al registro T para almacenarlo y así poderlo comparar con el siguiente bit.

La otra línea de datos permite obtener la codificación convolucional por medio de un código convolucional de razón $\frac{1}{2}$, el cual significa que se introducirá un bit y saldrán dos. (Ver figura 5.21). Los códigos convolucionales son adecuados para usar sobre canales con mucho ruido (alta probabilidad de error).

Los códigos convolucionales son códigos lineales, donde la suma de dos palabras de código cualesquiera también es una palabra de código. El sistema tiene memoria: la codificación actual depende de los datos que se envían ahora y los que se enviaron en el pasado.

Un código convolucional queda especificado por tres parámetros (n,k,m):

- n es el número de bits de la palabra codificada
- k es el número de bits de la palabra de datos
- m es la memoria del código o longitud restringida

TESIS CON FALLA DE ORIGEN

El proceso de codificación de estos códigos se realiza utilizando un dispositivo lógico en el codificador.

Para ejemplificar lo anterior se tiene un codificador convolucional (4,3,5). Ver figura 5.22

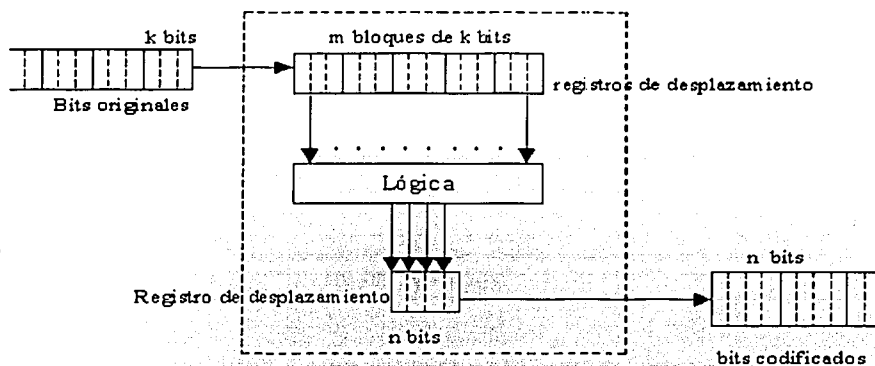


Figura 5.22 Ejemplo gráfico del codificador convolucional (4,3,5)

La palabra codificada se obtendría como el resultado de realizar una serie de operaciones lógicas entre determinados bits que están almacenados en los registros intermedios.

Con un codificador convolucional (2,1,3) se tiene la figura 5.23.

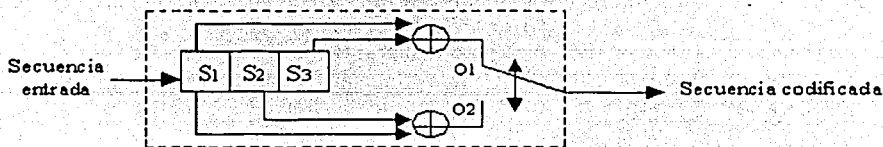


Figura 5.23 Código convolucional (2,1,3)

- El conmutador con las dos entradas hace el papel de un registro de desplazamiento de dos estados.
- El código convolucional es generado introduciendo un bit de datos y dando una revolución completa al conmutador.
- Inicialmente se supone que los registros intermedios contienen ceros.

En este ejemplo la palabra codificada se obtiene como resultado de sumas módulo-2 entre los bits indicados que están almacenados en los registros intermedios.

Como ejemplo del funcionamiento de este codificador, suponiendo que se quiere enviar la secuencia de bits 0101 (donde los bits más a la derecha son los más antiguos). El proceso de codificación es el siguiente:

- Se introduce el primer bit de la secuencia en el codificador (Ver figura 5.24):

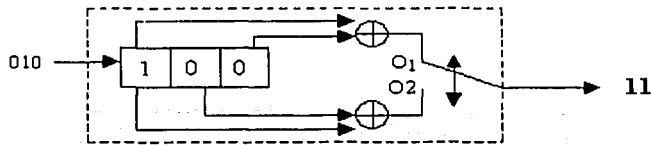


Figura 5.24 Primer bit de la secuencia 0101.

- Se introduce el segundo bit de la secuencia en el codificador (Ver figura 5.25):

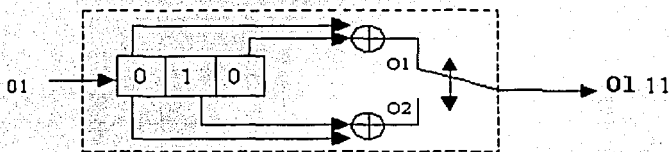


Figura 5.25 Segundo bit de la secuencia 0101.

- Se introduce el tercer bit de la secuencia en el codificador (Ver figura 5.26):

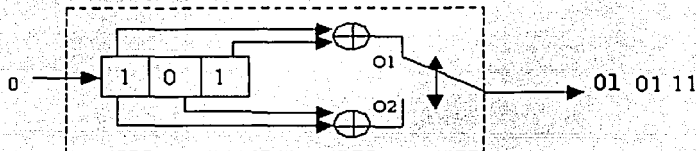


Figura 5.26 Tercer bit de la secuencia 0101.

- Se introduce el cuarto bit de la secuencia en el codificador (Ver figura 5.27):

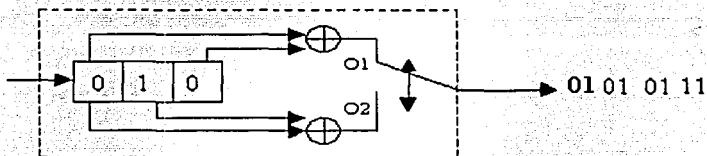


Figura 5.27 Cuarto bit de la secuencia 0101.

TESIS CON FALLA DE ORIGEN

Al final del proceso de codificación obtenemos que la secuencia codificada es 01 01 01 11.

El resultado de las codificaciones anteriores son 3 datos en paralelo: FEC-2/3.

Ya teniendo los datos anteriores son mapeados en un símbolo de 8 estados (000= -7, 001= -5, , 111= +7), ya que posteriormente llegarán a un modulador de 8 estados.

Cabe mencionar que el codificador Trellis tiene incorporado un reordenador de datos interno a través de 12 codificadores y precodificadores idénticos, los cuales reciben 12 grupos en paralelo de forma que en la salida extrae un símbolo de cada uno de ellos, por lo que en cada codificador y precodificador tendrá un retardo de 12 símbolos en cada uno de ellos. Este reordenamiento se realiza para evitar las ráfagas de errores que puedan ocurrir. Ver figura 5.28.

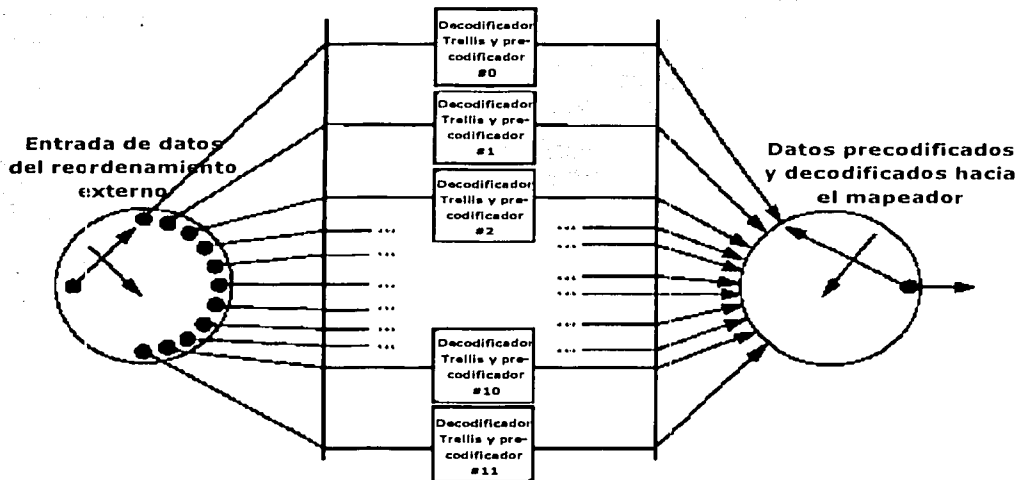


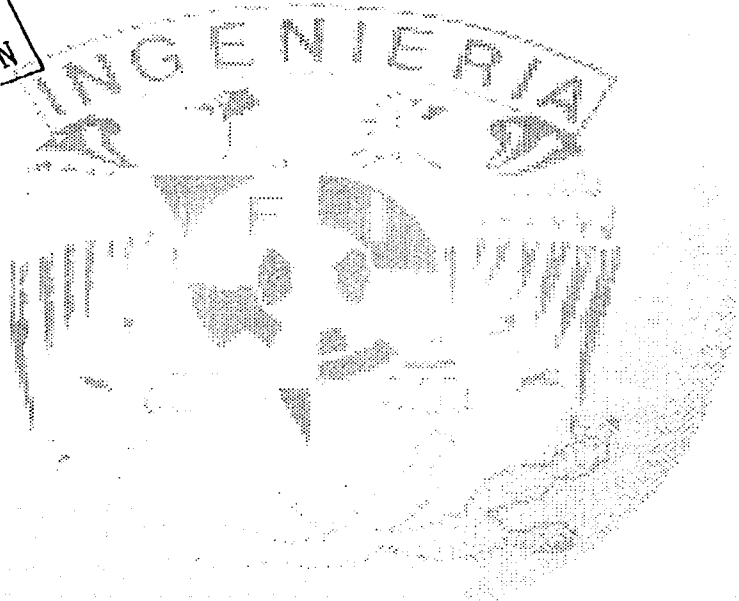
Figura 5.28 Reordenador Trellis

TESIS CON
FALLA DE ORIGEN

REFERENCIAS

- [1] Publicación de artículos por Juan Manuel Grijalvo. Sitio web <http://www.grijalvo.com>
- [2] Universidad Autónoma de Madrid, España. Sitio web <http://www.uam.es>
- [3] Universidad de Concepción, Chile, División de ingeniería eléctrica. Sitio Web <http://www.die.udec.cl>
- [4] Foro del Comité de Servicios de Televisión Avanzada (ATSC), organización internacional no lucrativa que desarrolla estándares voluntarios para la TV Digital. Sitio web <http://www.atscforum.org>

TESIS CON
FALLA DE ORIGEN



271

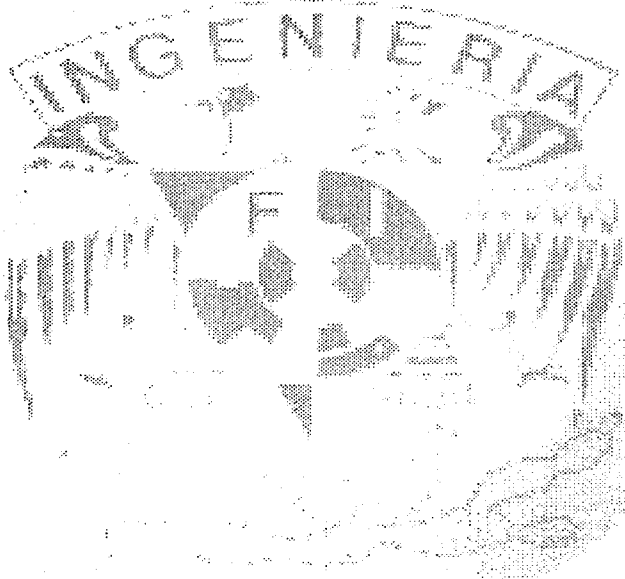
TESIS CON
FALLA DE ORIGEN



Capítulo # 6

*Modulación
8-VSB*

1974



TESIS CON
FALLA DE ORIGEN

174

6

6.1. ¿Que es 8-VSB?

8-VSB (Banda lateral residual de 8 niveles) es el formato de modulación de radio frecuencia escogido y desarrollado por la ATSC (Advanced Television Systems Committee) para la transmisión de televisión digital. VSB es un tipo de modulación basado en técnicas de modulación en amplitud (AM), la cual codifica datos variando la amplitud de una sola frecuencia portadora. En VSB las porciones de una de las bandas laterales redundantes son removidas para formar una banda lateral residual, banda así llamada por que es un residuo de la banda lateral. En AM la portadora no varía en amplitud, en cambio la modulación de datos se ve en el dominio de la frecuencia como la aparición de nuevas componentes espectrales, llamadas bandas laterales, que están en función de los cambios de la señal original, estas se encuentran en frecuencias de $F_{\text{portadora}} \pm F_{\text{mensaje}}$, la banda lateral izquierda o inferior (LSB) aparece en una frecuencia menor que la portadora y la banda lateral derecha o superior (USB) aparece en un frecuencia más alta que la portadora, ambas bandas laterales contienen la misma información.

La transmisión VSB es similar a la de banda lateral única (SSB), en la cual una de las bandas laterales es completamente suprimida, sin embargo en transmisión VSB la segunda banda no es completamente suprimida, solamente es filtrada para remover el rango de frecuencias deseado lo que provoca que queden residuos de ésta. El modo 8-VSB incluye 8 niveles ($\pm 1, \pm 3, \pm 5, \pm 7$) de amplitud para transportar datos a una velocidad de 19.3 Mbps en un solo canal de 6 Mhz. 8-VSB es considerado efectivo para la transmisión simultánea de más de un programa de DTV (Televisión Digital multicasting) y la transmisión de datos junto con programas de televisión (datacasting) por que este soporta un gran numero de carga útil de datos. El ATSC^[1] adoptó el sistema de transmisión VSB por que tiene un gran ancho de banda, el cual es necesario para la transmisión de HDTV (Televisión digital de Alta definición).

6.2. Descripción de un transmisor VSB.

En la figura 6.1 se presenta el diagrama a bloques de un transmisor de 8-VSB típico, como observaremos es en esta etapa en donde se encuentra la mayor parte del proceso de la transmisión.

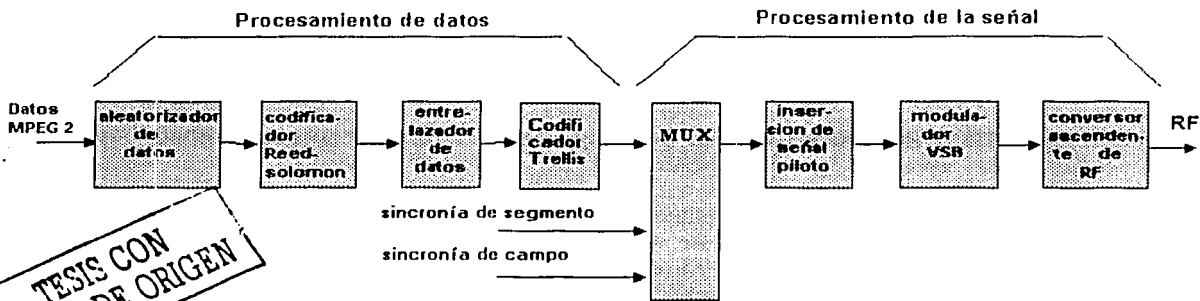


Figura 6.1 Diagrama a bloques de un transmisor VSB.

6.3 Procesamiento de datos.

En este bloque del transmisor VSB se tienen 4 etapas en las cuales se realizan una serie de alteraciones y procesos a los paquetes de información que serán enviados, estos paquetes llegan a esta etapa una vez que han pasado previamente por el proceso de compresión.

Lo primero que debe tenerse en cuenta es que los paquetes de MPEG-2^[2] (Moving Picture Expert Group) deben estar sincronizados con los circuitos internos del transmisor en la entrada de la señal a este. Antes de que algún procesamiento de señal pueda ocurrir el transmisor 8-VSB debe identificar correctamente el inicio y el fin de cada paquete de datos MPEG-2. Esto se realiza utilizando el byte de sincronía MPEG-2. Los paquetes de MPEG-2 son de 188 bytes de longitud, en donde siempre el primer byte al inicio del paquete es de sincronía.

6.3.1 Aleatorizador de datos.

El aleatorizador de datos permite distribuir, en forma homogénea, la energía de la señal que se transmitirá en el ancho de banda del canal de transmisión. Esto se logra generando una secuencia pseudoaleatoria de datos y su objetivo es hacer un uso más eficiente de la energía de la señal. La carga de bits 8-VSB debe tener una completa aleatorización, como ruido natural. Esto es por que la respuesta en frecuencia de la señal transmitida debe tener un espectro liso como ruido para usar el canal de RF (Radio Frecuencia) asignado con la máxima eficiencia.

Si los datos contienen patrones repetitivos, el recurrente ritmo de estos patrones puede causar que la energía de RF contenida en la señal de transmisión se “amontone” en ciertos puntos discretos en el espectro de frecuencias, por lo que daría lugar a la existencia de “hoyos” es decir frecuencias en las cuales no hay energía de la señal. Esto implica que ciertas partes del canal de 6 MHz estén sobre-ocupadas y otras partes muy poco ocupadas por la energía de la señal. En el aleatorizador de datos, el valor de cada byte es cambiado de

acuerdo a un patrón conocido de generación de números pseudoaleatorios. Este proceso se realiza de forma inversa en el receptor para recuperar la señal original.

6.3.2 Codificador Reed-Solomon.

La información a la salida del aleatorizador ingresa a un bloque de codificación Reed Solomon que añade 20 bytes de redundancia a los paquetes MPEG-2 de 188 bytes de información. La codificación Reed-Solomon consiste en un método de corrección directa de errores (FEC, Forward Error Correction) "sin memoria", por lo que su salida no se ve afectada por los símbolos anteriores que entran y salen del codificador, los códigos Reed-Solomon permiten la corrección de hasta 10 bytes de información errónea.

El Forward Error Correction es generalmente un término usado para describir una variedad de técnicas que pueden usarse para corregir errores en los bits que ocurren durante la transmisión. El ruido atmosférico, la multitrayectoria, el desvanecimiento de la señal y los transmisores no lineales pueden crear bits erróneos en la recepción. El FEC puede detectar y corregir estos errores, dentro de un límite razonable.

El codificador Reed-Solomon toma los 187 bytes de un paquete MPEG-2 de entrada (el byte de sincronía ha sido removido del paquete) y matemáticamente manipula estos como un bloque para crear un tipo de "identificador digital (ID)", este "identificador ID" ocupa 20 bytes adicionales los cuales son colocados al final del paquete original de 187 bytes codificado. Estos 20 bytes son conocidos como bytes de paridad Reed-Solomon. El esquema de codificación Reed-Solomon usado en DTV puede soportar máximo 10 bytes en error por paquete.

6.3.3. Entrelazado de datos.

El bloque que sigue a continuación es el de entrelazado externo de datos, que combina los bytes codificados utilizando 12 niveles de entrelazado. Este proceso puede entenderse como una dispersión de los bytes de los diferentes paquetes MPEG-2, los cuales son combinados para dar origen a nuevos paquetes del mismo tamaño que los originales, pero combinados y repartidos de acuerdo al patrón de 12 niveles. El entrelazado permite incrementar la confiabilidad del sistema, pues en caso de perderse un paquete, únicamente una porción de la información correspondiente a cada paquete se extraviaría, misma que puede recuperarse gracias a la codificación. Este tipo de codificación protege a la información de un tipo de ruido conocido como "impulsivo".

El entrelazador de datos perturba el orden secuencial del flujo de datos y dispersa los paquetes de datos MPEG-2 en todo el tiempo (en un rango de alrededor de 4.5 ms por medio del uso de buffers de memoria), a fin de minimizar la sensibilidad de la señal transmitida a la interferencia. El entrelazador de datos entonces ensambla nuevos paquetes de datos incorporando pequeños fragmentos de diferentes paquetes MPEG-2. Los paquetes de datos reconstruidos son de la misma longitud que los paquetes originales de MPEG-2: 207 bytes (después de codificación Reed-Solomon). Por ejemplo, esto es equivalente a "repartir todas tus manzanas (bytes) en muchas diferentes canastas (tiempo)". Si el ruido

impulsivo le pega a la señal durante la propagación y “una canasta” es perdida (varios milisegundos), muchos diferentes paquetes MPEG-2 perdieron “una manzana” en lugar de que un paquete MPEG-2 perdiera todas sus “manzanas”, esto es conocido como diversidad temporal, si cada paquete solo pierde un pequeño número de bytes, el decodificador Reed-Solomon en el receptor DTV es capaz de corregir los errores y recuperar los datos.

6.3.4. Codificador Trellis.

El Codificador Trellis es otra forma de FEC diferente del código Reed-Solomon, el cual trata el paquete entero MPEG-2 como un bloque. El código Reed-Solomon es conocido como un código de bloque, mientras que el código Trellis es un código convolucional. Para la Codificación Trellis cada byte es dividido en 4 palabras de 2 bits. En el Codificador Trellis cada palabra de 2 bits que llega es comparada con la palabra pasada de 2 bits, un código binario de 3 bits es generado para describir la transición de la palabra de 2 bits previa con la actual. Las palabras originales de 2 bits son sustituidas por estos códigos de 3 bits y transmitidas a través del espacio como los 8 niveles de símbolos de 8-VSB ($3 \text{ bits} = 2^3 = 8$ combinaciones o niveles). Por cada 2 bits que entran al Codificador Trellis salen 3 bits.

6.4. Procesamiento de la señal.

Ahora la siguiente etapa que tiene lugar en el transmisor es la de procesar la señal obtenida a la salida de la Codificación Trellis. En esta etapa se prepara la señal para ser enviada por la antena hacia el espacio.

6.4.1 Sincronía e inserción de la señal piloto.

Por medio del multiplexor se insertan a los paquetes binarios que salen de Codificador Trellis, los paquetes de sincronía de campo y de frame. El siguiente paso en la cadena de procesamiento de la señal es la inserción de varias señales de ayuda, que auxilian al receptor de DTV para localizar y demodular la señal transmitida de RF con precisión. Estas son: la señal piloto ATSC, la señal de sincronía de segmento y la señal de sincronía de campo. Las señales de sincronía y piloto son insertadas después de la zona de aleatorización de datos y códigos de errores para no destruir el tiempo estable y la relación de amplitud que estas señales deben de tener para ser efectivas.

La primera señal de ayuda es el piloto ATSC. Solo antes de la modulación, un pequeño cambio en DC es aplicado a la señal en banda base de 8 niveles (la cual fue previamente centrada alrededor de 0 volts sin ninguna componente de DC). Esto causa una pequeña portadora residual que aparece en el punto de frecuencia cero del espectro modulado resultante, que es la señal piloto ATSC. Esto da a los circuitos PLL (Phase Lock Loops) de RF en el receptor de DTV la indicación de que es independiente de los datos transmitidos.

Las otras señales de ayuda son la sincronía de campo y de segmento ATSC. Un segmento de datos ATSC consta de 207 bytes de un paquete de datos entrelazado. Después de la Codificación Trellis el segmento de 207 bytes ha sido extendido en un flujo de 828 símbolos en 8 niveles. El segmento de sincronía es un pulso de 4 símbolos que es sumado

al frente de cada segmento de datos y reemplaza al primer byte perdido (byte de sincronía de paquete) del original paquete de datos MPEG-2. La sincronía de segmento aparece una vez cada 832 símbolos y siempre toma la forma de un pulso positivo-negativo-positivo oscilando entre niveles de +5 y -5 volts, ver figura 6.2.

Un segmento de datos es aproximadamente análogo a una línea de NTSC^[3] (National Television Systems Committee), sin embargo la sincronía de segmento ATSC es como la sincronía horizontal de NTSC. Su duración y frecuencia de repetición son, por su puesto, completamente diferentes, cada sincronía de segmento dura 0.37 μ seg, mientras que la sincronía de NTSC dura 4.7 μ seg. En ATSC el segmento de datos dura 77.3 μ seg y en NTSC la línea dura 63.6 μ seg.

Se combinan 313 segmentos de datos consecutivos para formar un campo de datos. En la figura 6.3 se muestra la composición de un campo de datos ATSC. La sincronía de campo ATSC es un segmento de datos completo que es repetido una vez por campo (24.2 mseg) y es aproximadamente análogo al intervalo vertical en NTSC. La sincronía de campo tiene un patrón conocido de símbolos, pulsos positivos-negativos y es usado en la recepción para eliminar los problemas de "fantasmas" ocasionados por una pobre recepción.

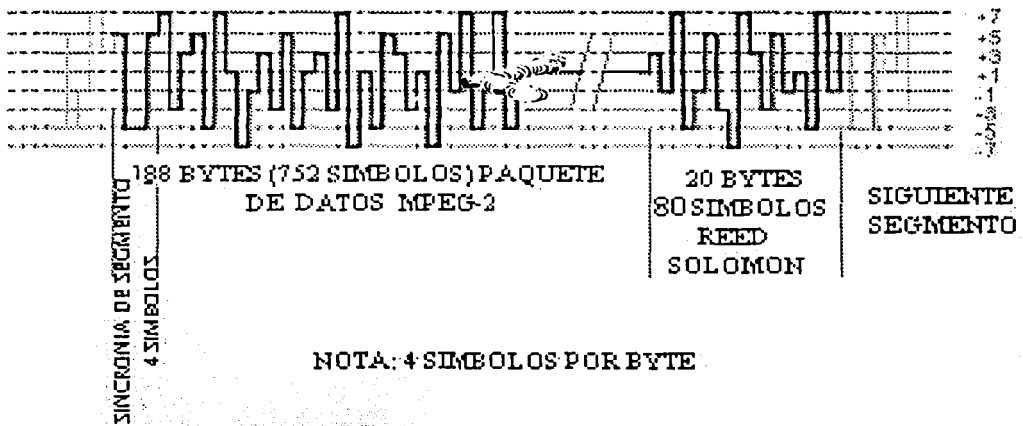


Figura 6.2 Segmento de datos en banda base ATSC

TESIS CON FALLA DE ORIGEN

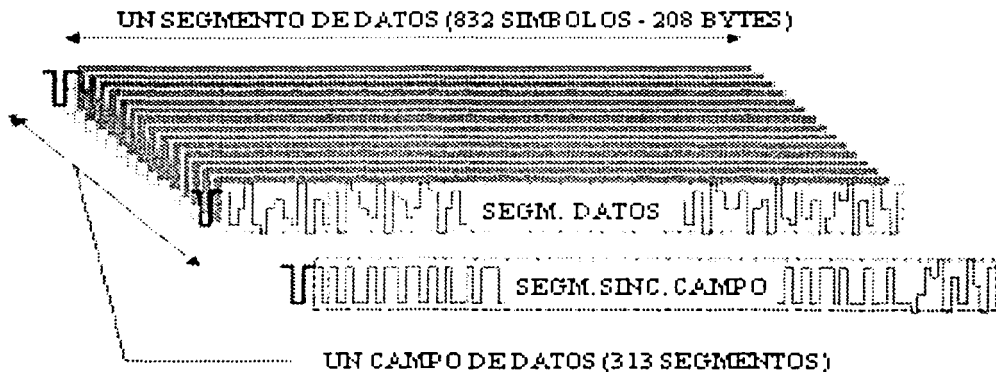


Figura 6.3 Campo de datos en banda base ATSC

En el final de cada segmento de sincronía de campo, los últimos 12 símbolos del último segmento de datos son repetidos con el propósito de reiniciar el Codificador Trellis en el receptor.

La robustez de la sincronía de segmento y campo permite una mayor exactitud en la recuperación del reloj en el receptor y cancela el efecto de "fantasmas", cuando la carga útil es completamente afectada por las condiciones de pobre recepción.

6.4.2 Modulador VSB.

El modulador de VSB tiene varias etapas que se describirán a continuación.

❖ Modulación AM^[4].

Los 8 niveles de banda base de señal de DTV, con los cambios sumados de sincronías y piloto de DC, es después modulada sobre una portadora de frecuencia intermedia (IF 44 MHz). Esto produce un amplio espectro de dobles bandas laterales de IF alrededor de la frecuencia de la portadora, como se muestra en la figura 6.5. El ancho de banda ocupado de esta señal de IF es demasiado grande para ser transmitido en el canal de RF de 6 MHz. Afortunadamente, hay "trucos" que pueden ser empleados para filtrar una gran parte de este espectro sin destruir ninguna información digital vital.

TESIS CON
FALLA DE ORIGEN

Una simple inspección a la figura 6.5 revela el alto grado de redundancia en la doble banda lateral del espectro de AM, el gran número de lóbulos laterales son simplemente incrustaciones de copias del espectro central, y la banda lateral inferior completa es una imagen espejo de la banda lateral superior. Por este hecho es posible descartar casi completamente la banda lateral inferior y todos los lóbulos laterales de la banda lateral superior sin ninguna pérdida de información. La señal restante (la mitad del lóbulo mayor en el centro del espectro) puede ser aun recortada a la mitad gracias al teorema de Nyquist, el cual declara que solo la mitad del ancho de banda en frecuencia es requerido para transmitir una señal digital en la velocidad de muestreo dada. El trabajo de reducir el espectro de doble banda lateral de IF corresponde ahora a la siguiente etapa, el filtro de Nyquist.

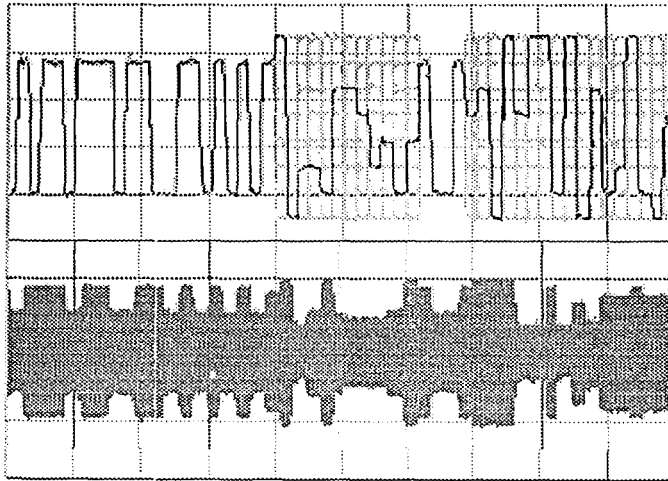


Figura 6.4 La señal en banda base de 8 niveles es modulada en AM sobre la portadora de IF.

Primera señal: La señal en banda base de 8 niveles en el final de un segmento de sincronía de campo. En las áreas ampliadas se muestra la señal original sin modular

Segunda señal: Un segmento de sincronía de campo después de la modulación AM (diferente) sobre la portadora IF.

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

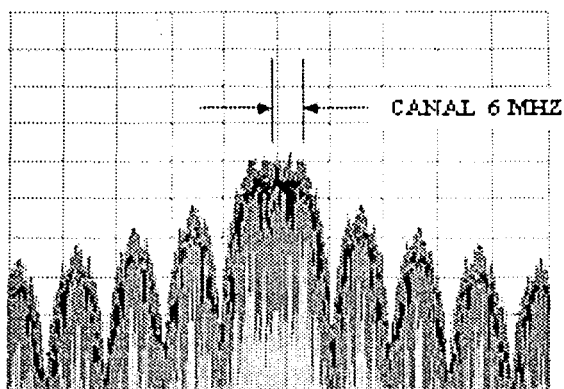


Figura 6.5 Espectro de doble banda lateral creado por la modulación AM

El espectro de doble banda lateral creado por la modulación AM de la señal en banda base sobre la portadora de IF es mucho más grande que el canal de 6 MHz de RF.

❖ Filtro Nyquist.

Como resultado de los datos sumados a la señal en forma de FEC y la inserción de sincronía, la velocidad de datos de la señal se incrementa de 19.39 Mbps en la entrada del transmisor a 32.28 Mbps en la salida del Codificador Trellis. Debido a que 3 bits son transmitidos en los 8 niveles, la velocidad de símbolo resultante es $32 \text{ Mb}/3 = 10.76$ millones de símbolos por segundo. Gracias al teorema de Nyquist es posible transmitir 10.76 millones de símbolos por segundo en una señal VSB con un mínimo ancho de banda de frecuencia $\frac{1}{2} * 10.76 \text{ MHz} = 5.38 \text{ MHz}$. Con el ancho de banda del canal asignado de 6 MHz, es posible con la ayuda del filtrado VSB disminuir la señal para el canal de 6 MHz. Se tienen 5.38 MHz (mínimo ancho de banda por Nyquist) que sumados a 620 KHz (ancho de banda de sobra) nos dan 6.0 MHz (ancho de banda del canal). El gran ancho de banda de sobra con el que se cuenta facilita la implementación de hardware, en términos de requerimiento de filtrado y costos de reloj para el muestreo.

La respuesta en frecuencia resultante después del filtro VSB Nyquist se muestra en la figura 6.6; note como el formato 8-VSB usa el método de banda lateral residual para conservar el espacio del espectro, 8-VSB toma este concepto para agrandar los extremos: la bandas laterales bajas de RF son casi completamente removidas.



Figura 6.6 Espectro en frecuencia de RF 8-VSB.

Note la presencia de la señal piloto ATSC en el extremo del canal. La banda lateral inferior (debajo de la frecuencia de la señal piloto) está casi completamente removida.

La virtual eliminación de la banda lateral inferior a través del filtrado de la banda superior crea cambios muy significativos en la forma de la onda de RF que es finalmente transmitida esto se muestra en la figura 6.7.

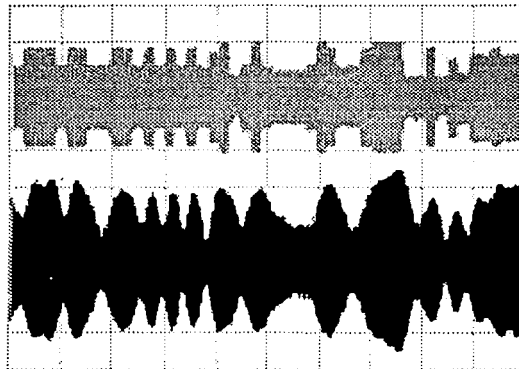


Figura 6.7 Efectos del filtro Nyquist sobre la envolvente de RF 8-VSB.

Primera imagen: envolvente de la doble banda lateral de IF antes del filtrado de Nyquist

Segunda imagen: la misma señal de IF después del filtrado de Nyquist. Las cuadraturas de las transiciones son perdidas y la envolvente adquiere la apariencia como de ruido.

Al observar la figura 6.7 es natural que pensemos que se han perdido los 8 niveles de la señal y toda la información contenida en ellos, pero no es así.

❖ Diagrama de ojo 8-VSB.

El diagrama de ojo es la sobreposición de muchas trazas de la amplitud de la señal de RF en el receptor en el instante del muestreo. Aunque la señal de RF a la hora de ser muestreada en un intervalo de tiempo puede tomar cualquiera de los ocho niveles posibles, la convergencia de muchas trazas de la señal forman 7 "ojos" que coinciden con la ocurrencia de los pulsos de reloj en el receptor, como se muestra en la figura 6.8. Este diagrama es muy utilizado en aparatos de medición y prueba, si la señal es dañada durante la transmisión, estos "ojos" podrían cerrarse y desaparecer. Si esto sucede indica que la señal de RF no tiene la amplitud correcta en el instante correcto.

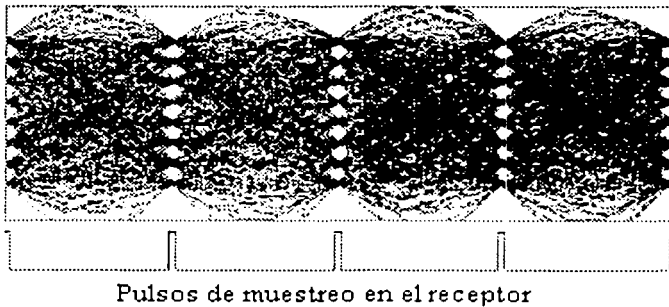


Figura 6.8 Diagrama de ojo 8-VSB.

❖ Constelación de la señal 8-VSB.

Otra representación de la señal 8-VSB muy popular en equipos de prueba es la constelación 8-VSB. Esta es una representación gráfica en dos dimensiones de la señal de RF en amplitud de la portadora y fase de cada muestra en el tiempo. En VSB solo se envía en diferentes amplitudes la información y la fase no es de interés como lo es en un sistema de QAM (Modulación en amplitud por multiplexaje en cuadratura). La comparación entre 8-VSB y QAM se muestra en la figura 6.9. En QAM los 8 niveles son recuperados por un detector de fase, y la constelación de la señal 8-VSB es por lo tanto una serie de 8 líneas verticales que corresponden a los 8 niveles de amplitud transmitidos. Las líneas que existen alrededor de la constelación 8-VSB indican que la portadora de RF esta en constante cambio tanto en amplitud como en fase.

TESIS CON
FALLA DE ORIGEN

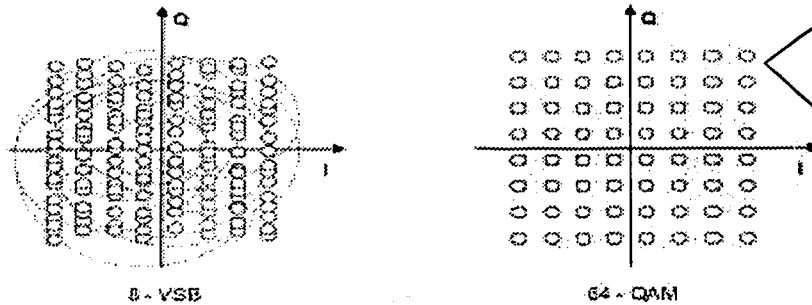


Figura 6.9 Constelación de la señal 8-VSB vs 64 QAM.

TESIS CON FALLA DE ORIGEN

6.4.3 Conversor ascendente de RF y el resto de los procesos de 8-VSB.

Después del filtro Nyquist VSB, la señal de frecuencia intermedia (IF) es convertida a una frecuencia portadora mayor para ser transmitida, esto se hace tradicionalmente con circuitos oscilador-mezclador-filtro, para que quede dentro del canal de banda de frecuencia en VHF (muy alta frecuencia) o UHF (ultra alta frecuencia) asignado. Posteriormente la señal pasa por un amplificador de potencia de RF tradicional. La última etapa es la antena la cual transmite toda la potencia de la señal. La antena es diseñada de acuerdo a las necesidades de transmisión como son el patrón de radiación y el alcance requerido, así como la banda de frecuencia en la cual transmitirá.

6.5 Descripción de un receptor VSB

En la figura 6.10 se presenta el diagrama a bloques de un receptor de 8-VSB típico, en un receptor la señal es demodulada aplicando los mismos principios que mencionamos anteriormente pero en reversa, es decir en un proceso contrario. El transmisor VSB es la última etapa en la gran cadena de procesos que sufre la señal de video, audio y datos a transmitir. Al contrario de lo que sucede con la recepción, en este caso el receptor VSB es la primera etapa por la cual debe de atravesar la señal que contiene la información que deseamos recuperar.

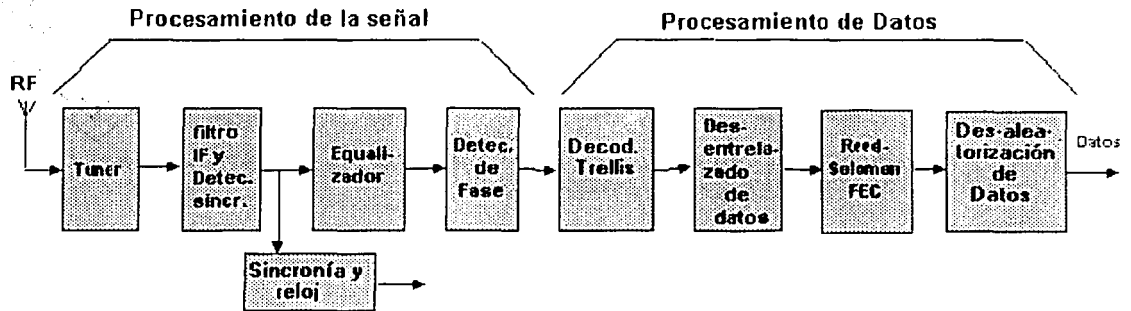


Figura 6.10 Diagrama a bloques de un receptor VSB

❖ **Tuner (sintonizador).**

Este dispositivo se encarga de realizar el proceso contrario al que realiza el convertor ascendente, es decir, la señal recibida en una alta frecuencia vuelve a pasar a frecuencia intermedia IF, esto se puede lograr utilizando la misma técnica que se utilizó en el convertor de fase (circuitos oscilador-mezclador-filtro) utilizando el principio de heterodinización.

❖ **Filtro IF y detector de sincronía.**

Esta etapa realiza prácticamente lo contrario a la etapa del modulador VSB, es decir que en esta etapa se filtra IF, con esto se logra tener a la señal en banda base, además se retiran las señales de piloto y las de sincronía que son utilizadas por los demás circuitos para la decodificación adecuada de los datos y también se recupera el reloj.

❖ **Ecualizador.**

Esta etapa solo maneja el flujo de bits de información, ya sin información de sincronía y reloj (esta es manejada por otro circuito al cual fue entregada esta información después de su extracción de la señal), el ecualizador se encarga de remover las distorsiones de canal (inclinaciones, fantasmas,...) con la ayuda de la señal de sincronía de campo y piloto.

❖ **Detector de fase.**

Se encarga de remover el ruido de fase que se suma a la señal durante la conversión de frecuencia y detecta la fase de la señal.

TESIS CON
 FALLA DE ORIGEN

❖ **Decodificador Trellis.**

Realiza la operación inversa del Codificador Trellis, lo cual es sencillo debido a que se conoce el código aplicado por el transmisor. En esta etapa se interpretan los 8 niveles y se obtiene el código de acuerdo a estos.

❖ **Desentrelazado de datos.**

Aquí se realiza la operación contraria a la efectuada en el entrelazado de datos. Aplica un código convolucional, el cual es conocido, ya que se aplicó para la transmisión, con lo cual se logra regresar a la colocación original a los bytes de datos.

❖ **Reed-Solomon FEC.**

En esta etapa se interpretan los bloques de bits de FEC sumados durante la transmisión y se realiza la corrección de errores en donde se requiera (puede corregir un máximo de 10 bytes por paquete) y es muy efectivo para combatir principalmente el ruido impulsivo.

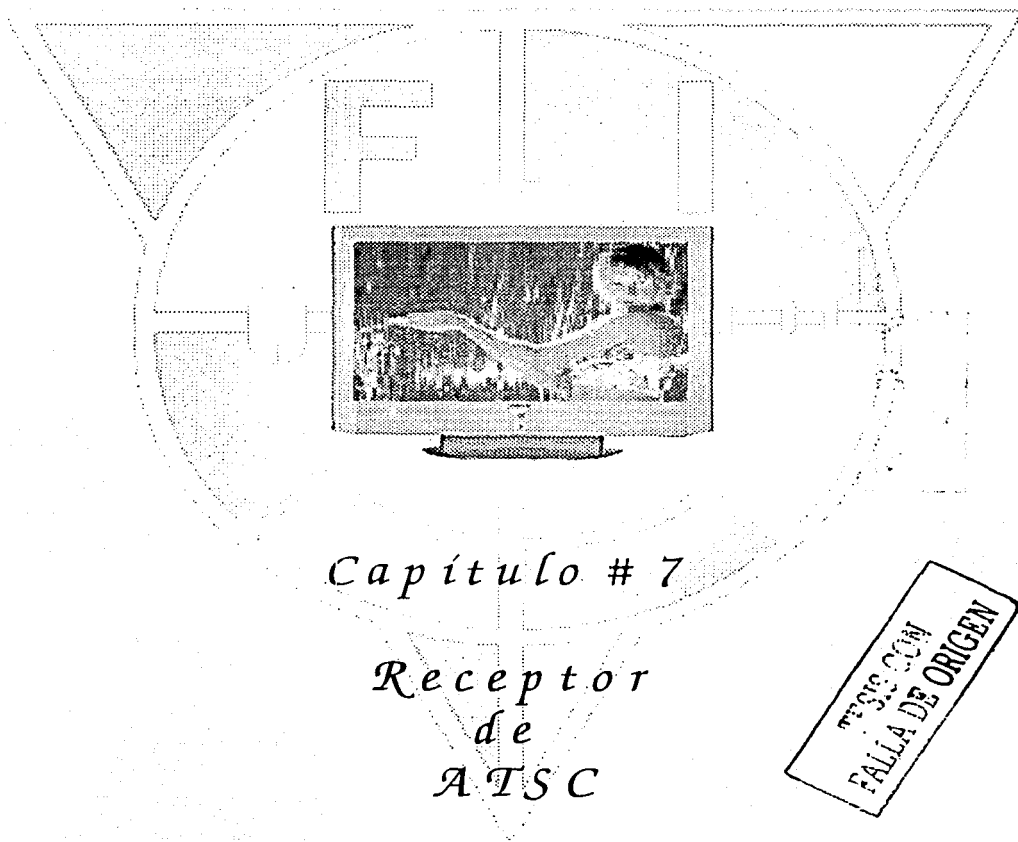
❖ **Desaleatorizador de datos.**

Regresa los bytes a su forma original, es decir, realiza la operación contraria al aleatorizador de datos del transmisor. En esta etapa los bytes recuperan su valor original, al ser colocados los datos en su forma original, utilizando el código pseudoaleatorio conocido que fue usado en la transmisión.

Los procesos anteriores de decodificación serán explicados de una manera más amplia en el siguiente capítulo.

REFERENCIAS

- [1] Comité de Servicios de Televisión Avanzada (ATSC), organización internacional no lucrativa que desarrolla estándares voluntarios para la TV Digital. Sitio web www.atsc.org
- [2] Índice de recursos MPEG en Internet, se enfoca en los estándares MPEG-1 y MPEG-2. Sitio web www.mpeg.org
- [3] Comité del Sistema de Televisión Nacional (NTSC). Organización internacional que regula el estándar de televisión americano. Sitio web www.ntsc-tv.com
- [4] Libro de SMITH "Modern Communication Circuits", editorial Mc Graw Hill, Tercera edición, Estados Unidos 1986, cap 12.

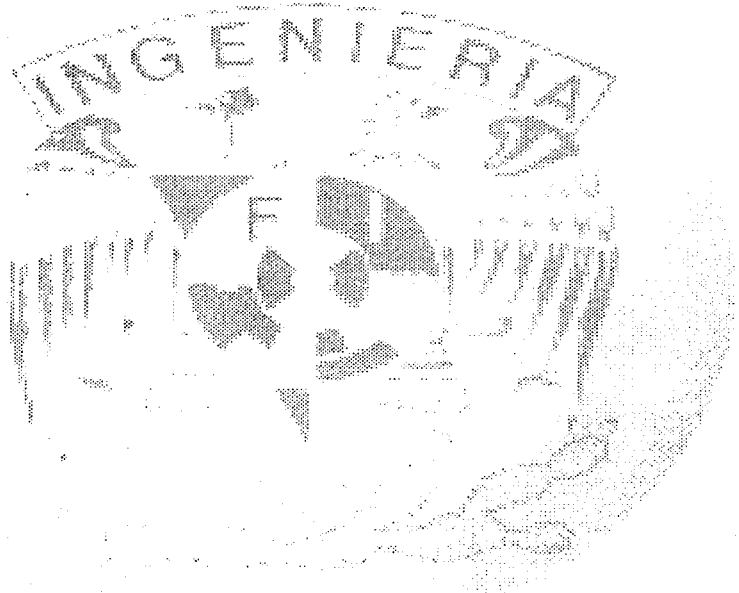


Capítulo # 7

Receptor
de
ATSC

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



191

7

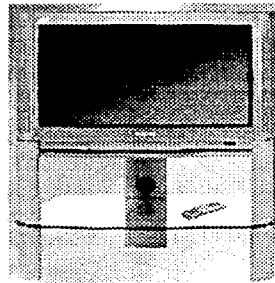
ANÁLISIS CON
FALLA DE ORIGEN

7.1. Diagrama general de un receptor de TV digital.

A un receptor de DTV (Televisión Digital) se le nombra de muchas maneras, pero principalmente se le conoce como decodificador de DTV o Set-Top-Box (ver figura 7.1). Este es un dispositivo que tiene la función de decodificar la señal de DTV que se encuentra en estándar ATSC (Advanced Television Systems Committee) y convertirla a una señal de televisión analógica cuyo funcionamiento esta basado en el estándar NTSC (National Television Standards Committee) y así esta pueda ser reproducida por un televisor analógico. También existen dispositivos para la PC (computadora personal) que permiten ver la señal de DTV. Se considera que en un futuro, estos dispositivos dejarán de existir debido a que se crearán televisores totalmente digitales, capaces de decodificar la señal digital sin necesidad de convertirla a analógica. Sin embargo por el momento una televisión de este tipo resulta incosteable, es decir, el usuario no podría pagar el precio tan alto al que tendrían que venderse estos televisores. Es por eso que se recurre al uso de los decodificadores, mientras se logra desarrollar un televisor digital de menor costo. Entonces el uso de un decodificador es la opción más factible y cómoda para el usuario, ya que puede continuar utilizando su televisor analógico para recibir la señal de la DTV.



Decodificador Sagem de Quiero TV



*Philips 32DW6834C DigitalTV
Televisor Digital preparado para
recibir la señal
digital del operador inglés
ONDigital*



Decodificador Nokia de Via Digital

Figura 7.1 Decodificadores para TV digital.

En la figura 7.2 se muestra un diagrama a bloques de un receptor general para DTV en el estándar de ATSC con el propósito de explicar la función que desempeña cada uno de los dispositivos o bloques en este capítulo. El objetivo del receptor de DTV es recuperar la información de vídeo y audio para ser proyectada en la pantalla de un televisor analógico.

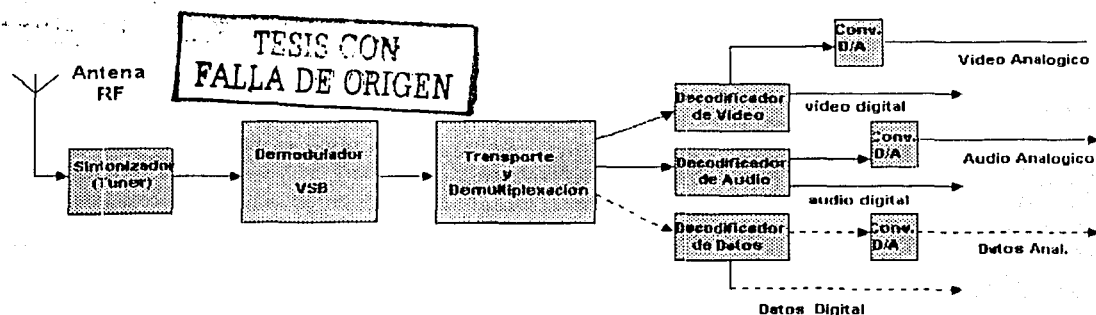


Figura 7.2 Diagrama a bloques de un receptor ATSC.

7.2. Antena.

La antena es un dispositivo muy importante en la recepción de la señal para la TV analógica y lo sigue siendo para la recepción de la DTV. Ahora se requiere que este dispositivo tenga incluso una recepción con la mayor eficacia posible, debido a que se busca una imagen totalmente nítida y de muy alta calidad en la pantalla del televisor. De ahí que no se puede permitir que la imagen en el televisor contenga el efecto que se presenta muy comúnmente en la señal analógica como son los fantasmas (en la imagen aparecen sombras de los objetos que existen en ella debido a la multitrayectoria de la señal) y el granizo (es un efecto en el cual aparecen puntos de brillo recorriendo toda la pantalla, parecido a la lluvia). En VSB el efecto que más se critica y al cual es más susceptible la señal es el de fantasmas provocado por la multitrayectoria. Las antenas juegan un papel muy importante para que esto no suceda, ya que este efecto puede ser eliminado con una orientación y polarización adecuada de la antena receptora.

El problema es que en cada canal la antena tendría que ser orientada para evitar el efecto de fantasmas. Por esta razón se han propuesto varios tipos de antenas para recepción. En los inicios de la DTV en Estados Unidos se propuso una antena de uso exterior la cual contaba con un motor que la hacía girar para variar su orientación y mantener buena recepción en distintos canales; pero este sistema de antena era muy caro, difícil de instalar e incomodo para el usuario. Esta fue la razón principal para que en sus inicios ATSC fuera superado por el DVB (Digital Video Broadcasting) y muy criticado por usar modulación VSB (Banda Lateral Residual). Pero en los años recientes se han realizado mejoras y agregado etapas en la modulación VSB que protegen a la señal y la hacen de tan alta calidad, que al realizarse pruebas en los Estados Unidos, se tuvo una imagen muy aceptable, utilizando como antena receptora una antena "V" o de "conejo" como se le conoce en México.

calidad de filtrado y menor susceptibilidad al ruido, y desde luego este dispositivo ha sido implementado para el "tuner" de la DTV.

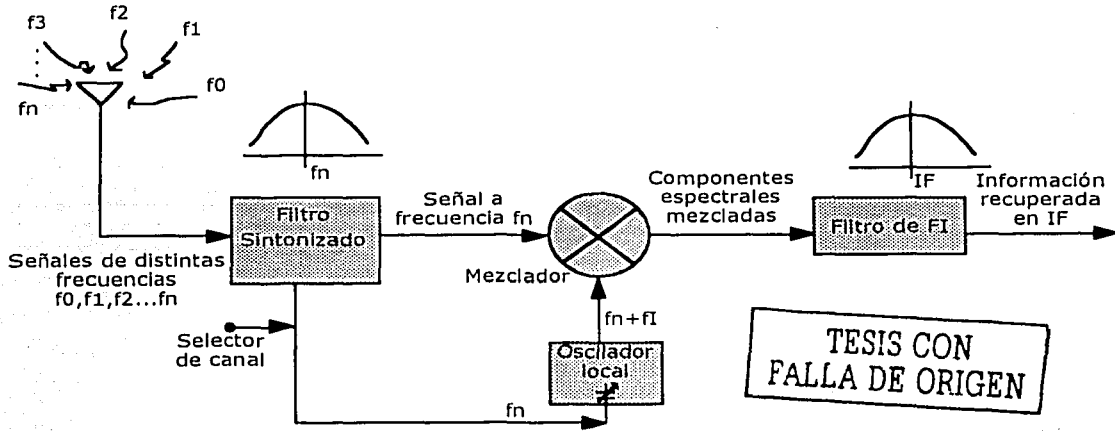
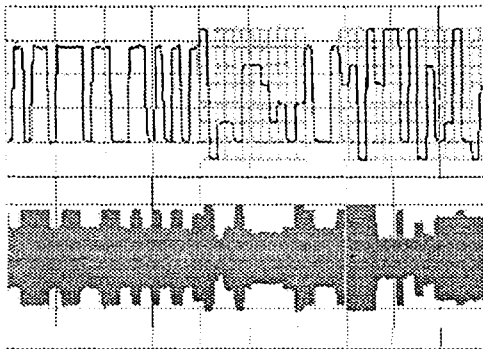


Figura 7.5 Sintonizador de canales (Tuner).

Después que la señal pasa por la etapa del sintonizador de canales cambiará de una alta frecuencia (figura 7.4 a) a una frecuencia intermedia (IF) con lo cual la señal tendrá la apariencia que se muestra en la figura 7.6. En la figura 7.6 a) se muestra en la parte superior la señal en niveles de voltaje que fue modulada por la portadora de IF y en la parte inferior se observa la señal modulada en IF.



a) Forma de onda en el osciloscopio



b) Espectro en frecuencia

Figura 7.6 Señal de 8-VSB en RF de frecuencia intermedia

7.4 Demodulador de 8-VSB.

El demodulador de 8-VSB tiene como objetivo transformar la señal, que es recuperada en banda IF por el sintonizador de canales, a un tren de bits sin alterar el orden ni el contenido de este. Esto se logra guiando a la señal a través de distintos procesos encargados de decodificarla para obtener a la salida el tren de bits que contiene la información de audio, video y datos. Dichos procesos se muestran en un diagrama a bloques (figura 7.7), que en conjunto forman lo que se denomina el demodulador de 8-VSB.

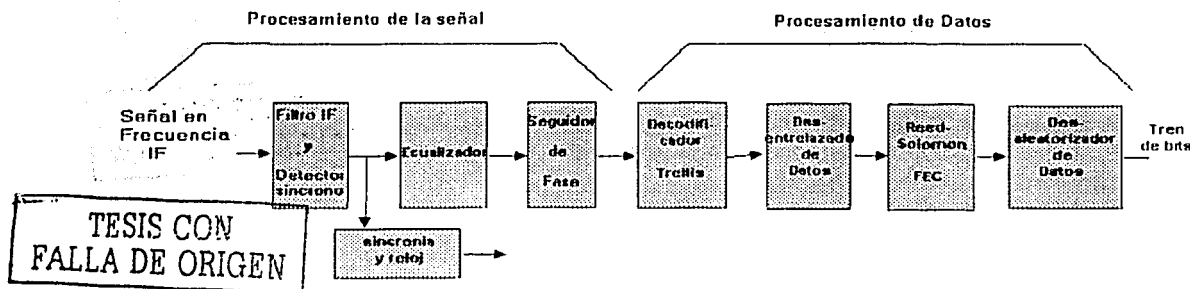


Figura 7.7 Demodulador de 8-VSB.

7.4.1 Filtro IF y detector síncrono.

En esta etapa el objetivo es demodular a la señal de frecuencia IF (figura 7.6 a) para obtener la información (video, audio y datos) en contenido y frecuencia originales. La frecuencia que tendrá la señal después de esta etapa será menor que la de IF y se le conoce como banda base (BB). Es decir, después de este proceso obtendremos la señal de 8 niveles en los que se encuentra la información (ver figura 7.8). La técnica utilizada para demodular en este caso es la detección síncrona debido a que elimina la distorsión causada por el desvanecimiento de la portadora. Esto lo logra utilizando como oscilador local (OL) un circuito FPLL (Fraccional Phase-Locked Loop), lo cual reduce el ruido de fase. En la figura 7.9 se muestra el diagrama a bloques del demodulador por detección síncrona.

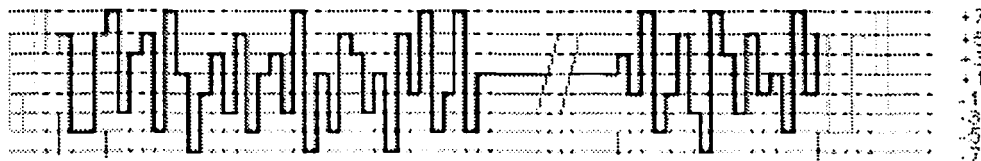


Figura 7.8 Señal de 8 niveles de 8-VSB.

Un "Fraccional Phase-Locked Loop (FPLL)" es básicamente un PLL (Phase-Locked Loop), con la diferencia de que cuenta con dos divisores en lugar de uno, como en el caso de un PLL normal, produciendo un control mas fino en frecuencia. A continuación se describirá el funcionamiento de dichos dispositivos.

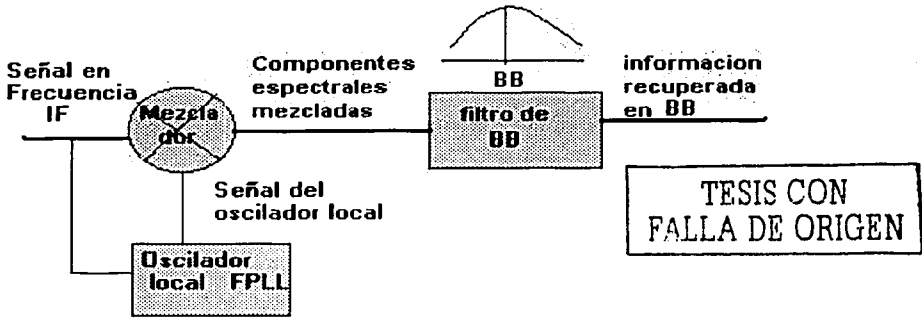


Figura 7.9 Demodulador AM por detección sincrónica.

❖ Phase-Locked Loop PLL.

Este dispositivo es básicamente un sistema de control de frecuencia de lazo cerrado, el cual funciona con un detector de fase sensible, obteniendo la diferencia de fases entre la entrada y la salida del oscilador controlado (CO). Un PLL básico consiste en un detector de fase/frecuencia, un controlador de corriente, un filtro de lazo, un oscilador controlado por voltaje (VCO) y un divisor (ver figura 7.10). Con una constante en el divisor igual a N, el lazo obliga a la frecuencia del VCO a ser exactamente N veces la frecuencia de referencia (entrada), esto es $F_{vco} = N \cdot F_{ref}$. El detector de fase/frecuencia y el controlador de corriente liberan pulsos de corriente positivos o negativos dependiendo de cómo es la fase de la señal de referencia, es decir, si se encuentra adelantada o atrasada con respecto a la fase de la señal dividida de VCO (ver figura 7.11). Estos pulsos de corriente son integrados por el filtro de lazo para generar un voltaje sintonizado que desplaza la frecuencia de VCO arriba o abajo hasta que las fases son sincronizadas.

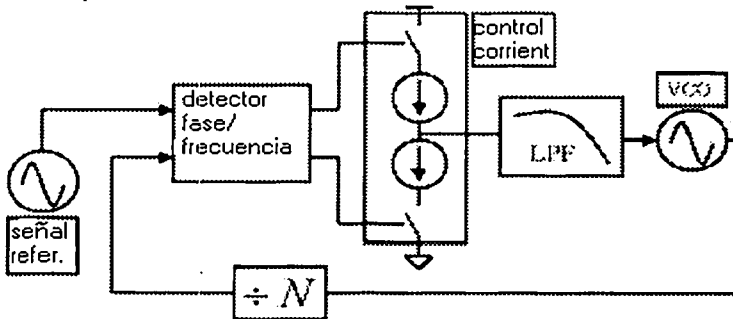


Figura 7.10 Esquema de bloques básico de un PLL.

El detector de fase/frecuencia es un dispositivo que compara dos frecuencias de entrada F_{ent} , generando una señal de salida que es la medida de la diferencia entre sus fases (por ejemplo si existe una diferencia en las frecuencias, la salida es una señal periódica que representa la diferencia entre las frecuencias), esto se muestra en la figura 7.11. Si la frecuencia de entrada no es igual a la frecuencia F_{vco} del VCO, la señal fase-error provoca que la F_{vco} sea desviada en la dirección de la F_{ent} . Si las condiciones son correctas el VCO puede llegar rápidamente a seguir a la F_{ent} y mantener una relación fija con esta.

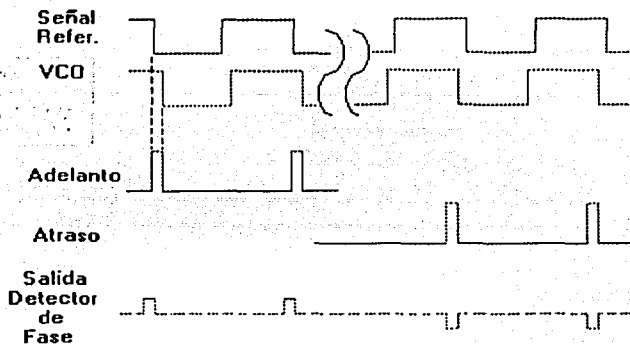


Figura 7.11 Detector de fase.

❖ Fractional Phase-Locked Loop FPLL.

Este dispositivo es una técnica popular para lograr frecuencias de resolución más fina que la frecuencia de referencia. Aprovecha que el divisor es variado periódicamente entre dos valores enteros.

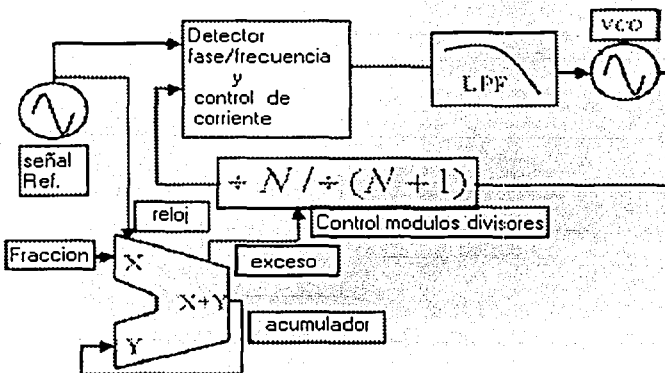


Figura 7.12 Esquema a bloques de un FPLL.

TESIS CON FALLA DE ORIGEN

La frecuencia promedio en VCO de este tipo de PLL (ver figura 7.12) es determinada mediante la siguiente ecuación:

$$F_{vco} \text{ promedio} = [T_n N F_{ref} + T_{n+1} (N+1) F_{ref}] / (T_n + T_{n+1}), \text{ la cual puede ser reducida a}$$
$$F_{vco} \text{ promedio} = [N + (T_{n+1}) / (T_n + T_{n+1})] * F_{ref} = (N + \text{fracción}) * F_{ref}.$$

Por lo tanto, la parte fraccional de la división (divisor) es determinada por el ciclo: la fracción de tiempo que el divisor esta dividiendo por $N+1$. Este tipo de PLL puede ser modelado usando un acumulador que suma las fracciones cada ciclo, cuando el acumulador rebasa su capacidad los excesos y el acumulador son reseteados a $N+1$.

7.4.2 Sincronía y reloj

En esta etapa se recupera la sincronía de segmento, con ayuda de la cual se genera el reloj del sistema, el receptor necesita dividir el tiempo en intervalos de símbolo para saber exactamente cuando debe volver a decodificar un nuevo símbolo. Además, para garantizar que la señal tiene tiempo de alcanzar un nivel adecuado reduciendo la posibilidad de equivocarse, es importante decodificar en el centro del intervalo de un símbolo. Como la frecuencia de los relojes (osciladores) del emisor y receptor no es idéntica, la desviación se va acumulando con el tiempo, al cabo de un cierto número de bits, el receptor puede estar mirando la señal fuera del intervalo correspondiente al símbolo. Para evitar esto, el receptor debe extraer su reloj de recepción de la señal recibida para permanecer sincronizado con el emisor. Para no transmitir la señal de reloj por una línea adicional se insertan transiciones en la señal que codifica los datos, que en nuestro caso son los pulsos de sincronía de segmento, insertados durante la modulación 8-VSB. Esto se puede lograr con un PLL (Phase Locked Loop) que es capaz de generar un reloj a una frecuencia determinada, sincronizado con las transiciones de una señal de entrada. Para poder usar un PLL es obligatorio que aparezcan transiciones en la señal cada varios símbolos. La señal de datos que transporta información de reloj codificada en sí misma se denomina self-clocking signal (Señal con reloj propio).

La recuperación de la sincronía de segmento es usada por el receptor para regenerar el reloj del sistema y el muestreo de la señal recibida. Debido a que la señal de sincronía es de alta frecuencia de repetición, nivel alto y de duración extendida, hace que su recuperación por el receptor sea fácil. Se trata de un pulso positivo-negativo-positivo de +5 a -5 el cual se encuentra al principio de cada segmento de bits. Los segmentos tienen una longitud de 207 bytes y la sincronía de segmento tiene una duración de 1 byte (ver figura 7.13).

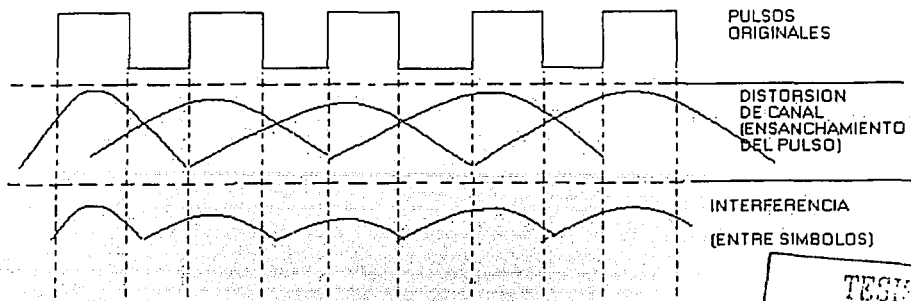


Figura 7.14 Interferencia Inter. Símbolo (ISI).

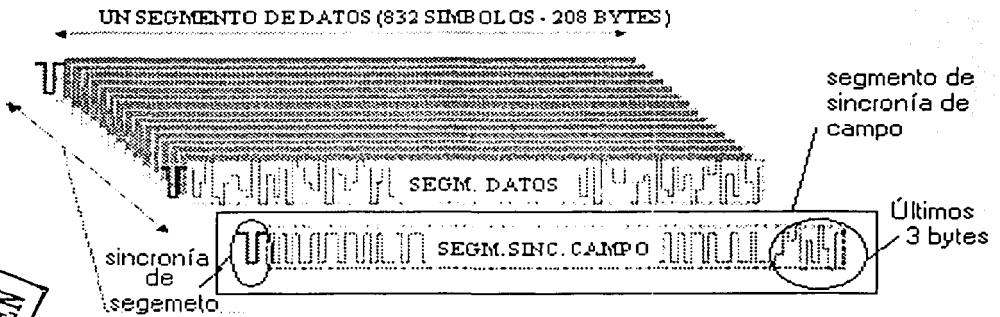
TESIS CON
FALLA DE ORIGEN

Los ecualizadores adaptables son, sin embargo, más que un filtro. Cuando se establece por primera vez un enlace, el ecualizador debe ajustar y perfeccionar los datos recibidos para compensar el efecto que tuvo el canal durante la transmisión. Debe recordarse que los fenómenos que causan la multitrayectoria constantemente están cambiando, y que el ecualizador debe adaptarse para compensar estos cambios. La adaptación puede lograrse porque el modulador inserta secuencias de instrucciones predeterminadas en la cadena de datos que son conocidas y guardadas por el demodulador, que en esta aplicación es el segmento de sincronía de campo. Un algoritmo se usa para adaptar el ecualizador comparando las secuencias de instrucciones transmitidas con las almacenadas localmente, enviando periódicamente las secuencias de instrucciones. El ecualizador puede estar adaptado para compensar el canal continuamente.

La sincronía de campo es un segmento de datos que se repite una vez por campo (cada 24.2 mseg). Un campo contiene 330 segmentos de datos consecutivos, su patrón de pulso es de positivo-negativo y es usado por el receptor para eliminar el efecto de fantasmas o duplicación de imagen de la señal causado por una pobre recepción. Esto se logra realizando una comparación entre el segmento de sincronía de campo recibido y el segmento de sincronía de campo conocido. El resultado de esta comparación son vectores de error que son usados para ajustar la etapa de ecualización en el receptor el cual tiene como objetivo cancelar el efecto de fantasmas. La sincronía de campo tiene una longitud de 208 bytes y al final de cada segmento de sincronía de campo, los últimos 12 símbolos (3 bytes) del último segmento de datos, son repetidos en orden para reiniciar el Decodificador Trellis en el receptor. Ver figura 7.15.



a) Pulsos de sincronía de campo



UN CAMPO DE DATOS (313 SEGMENTOS)

b) Campo de datos.

Figura 7.15 Segmento de sincronía de campo

7.4.4 Seguidor de fase (Phase Tracker).

En la señal de un oscilador la potencia principalmente se acumula en la frecuencia central, sin embargo algo de esta potencia es distribuida sobre las bandas laterales de la frecuencia central del oscilador. Estas componentes de frecuencia creadas no son deseadas y se les conoce como ruido del oscilador. Este ultimo puede tener diferentes impactos sobre el desarrollo del sistema, uno de los cuales es el ruido de fase. En un receptor el ruido de fase del oscilador local se mezcla con la señal de información y puede provocar que se genere una nueva señal con un frecuencia no deseada. Además, esta frecuencia podría estar dentro de la banda de paso del sistema de IF. Esto generaría dos señales con la misma información pero defasadas provocando el efecto ya comentado de duplicación de la imagen (fantasmas).

TESIS CON FALLA DE ORIGEN

El seguidor de fase elimina el ruido de fase generado por el transmisor, el sintonizador y el mezclador analógicos, dispositivos que tienden a generar y sumar ruido de fase a la señal de información durante la modulación y demodulación. Utiliza la sincronía de campo, dicho dispositivo es un sintonizador rápido de lazo (alrededor de 50 KHz).

7.4.5 Decodificador Trellis.

En esta etapa la entrada es una señal de 8 niveles (figura 7.16), la cual ha atravesado los bloques en los que se procesa la señal (procesos analógicos y digitales). Esto quiere decir que esta señal de 8 niveles se encuentra libre de ruido y efectos de multitrayectoria (caso ideal), además de que a partir de este bloque inicia el procesamiento de los datos (procesos totalmente digitales). El Decodificador Trellis interpreta estos niveles de voltaje de acuerdo a la tabla 7.1 y obtiene un tren de bits, el cual se encuentra codificado de acuerdo al algoritmo utilizado en el transmisor por el Codificador Trellis.

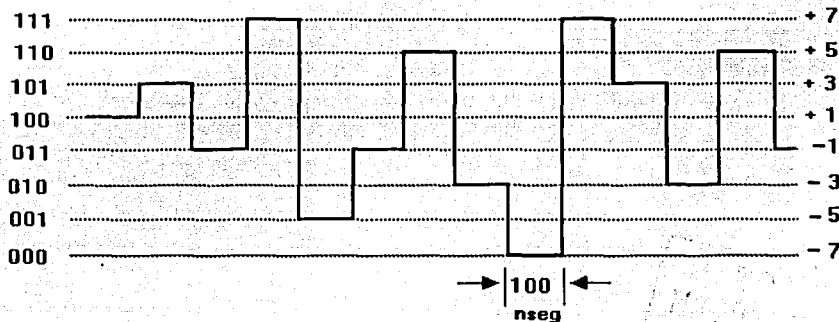


Figura 7.16 Señal de 8 niveles.

La Codificación Trellis es una forma de corregir errores utilizando un código convolucional en el transmisor. El Codificador Trellis toma cada palabra de dos bits que entra y la compara con la historia pasada (palabras previas), un código de 3 bits es generado matemáticamente para describir la transición de la palabra previa de 2 bits con la palabra presente. Este código de 3 bits sustituye a las palabras originales de 2 bits y es enviada a través del espacio en forma de símbolos de 8 niveles (Figura 7.16).

Nivel de voltaje	Código binario
-7	000
-5	001
-3	010
-1	011
+1	100
+3	101
+5	110
+7	111

Tabla 7.1 Asignación de niveles al código binario.

El Decodificador Trellis en el receptor interpreta los niveles de voltaje que recibe para obtener el código de 3 bits que requiere para reconstruir el paquete de datos compuesto por palabras de 2 bits. El Decodificador Trellis sigue una "pista, huella o rastro" de la forma como la señal se mueve de palabra a palabra. El potencial de este método consiste en la habilidad para rastrear la historia de una señal a través del tiempo y descartar la información potencialmente fallida (errores) basándose en la señal pasada y en un comportamiento futuro. Cuando alguno de los códigos de 3 bits son corrompidos, el Decodificador Trellis puede considerar muchas alternativas (pistas, huellas o rastros) para encontrar el candidato mas adecuado para realizar la corrección. La decodificación se lleva a cabo utilizando el Algoritmo de Viterbi (por lo cual es necesario utilizar un Decodificador de Viterbi). Al utilizar el código convolucional de Trellis logramos realizar el proceso de demodulación y decodificación en un solo paso, sin implicar esto un incremento en el ancho de banda del sistema.

El Algoritmo de Viterbi: es una de las funciones de procesamiento de señales digitales mas utilizadas, es usado en el diseño de decodificadores para códigos convolucionales. De entre las técnicas de decodificación, el Algoritmo de Viterbi es uno de los mejores métodos para decodificación de códigos convolucionales. El Algoritmo de Viterbi puede ser gráficamente representado por un diagrama de estado llamado "*Diagrama de Trellis*", el cual es graficado como función del tiempo. Un ejemplo simple de un Diagrama de Trellis se muestra en la figura 7.17; como se ve en esta figura cada nodo corresponde a un estado distinto en un tiempo dado, y cada rama (línea) representa una transición a otro estado en la siguiente etapa.

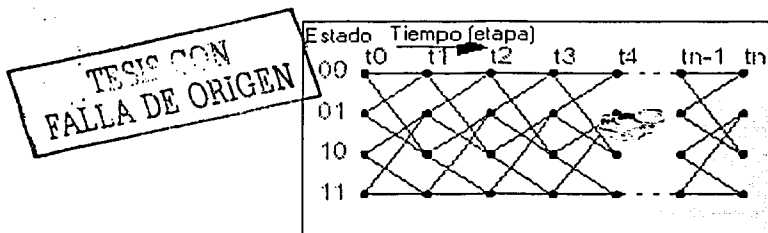


Figura 7.17 Diagrama de Trellis.

El proceso consiste en desechar algunos de todos los caminos posibles. Lo que se consigue aplicando este método es reducir el número de cálculos. Según el Algoritmo de Viterbi, para reducir el número de cálculos, cada vez que dos trayectos (también llamados ramas) se junten en un estado en el Diagrama de Trellis, el de mayor métrica (distancia Hamming) acumulada se desecha en la búsqueda del trayecto óptimo. Esto se debe a que el estado actual resume la historia de todos los estados anteriores en cuanto a su influencia en los estado posteriores. Esto se hace en los $2k-1$ estados, se pasa al intervalo t_i+1 y se repite el proceso. Realmente, esta eliminación de posibles caminos no comienza hasta el tercer nivel de representación. Esto se debe a que hasta ese instante, no han podido converger dos ramas en un estado ya que por cada k bits de palabra de datos ($k = 1$ en el ejemplo) surgen $2k$ estados. Es decir, del primer estado (el 00 por definición) tendremos dos estados en el siguiente nivel de representación, y por cada uno de estos, otros dos en el posterior nivel.

Por tanto será en el tercer nivel cuando empecemos a tomar decisiones porque ahora sí convergerán dos ramas en cada estado.

Para ilustrar mejor lo anterior utilizaremos el siguiente ejemplo, suponemos una secuencia recibida con una serie de errores producidos por el ruido en el canal (en "negritas" se encuentran los bits erróneos en la recepción), dado que también conocemos el Trellis que sigue la palabra enviada, podremos observar gráficamente en el proceso de Decodificación de Viterbi cómo se van obteniendo los caminos que conducen al resultado óptimo. De esta manera:

Datos:	1 1 0 0 1 0 1 1
Estado presente:	10 11 01 00 10 01 10 11
Codificado:	11 10 10 11 11 01 00 10
Recibido:	10 10 11 11 01 01 00 10

Veamos las distintas representaciones gráficas que desarrollan el proceso según dicta nuestra máquina de estados.

El primer y segundo nivel son fijos para cualquier entrada. Del estado 00 existen dos posibles ramas, hacia el estado 00 y hacia el 10.

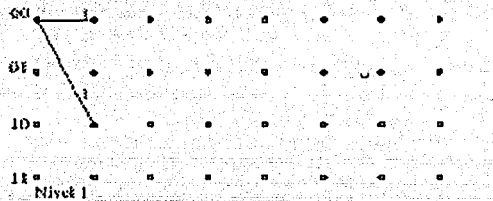


Figura 7.18 Primer nivel del Diagrama de Trellis.

Todavía no hay que realizar ninguna decisión ya que, como hemos dicho, estos caminos son los mismos independientemente de la entrada.

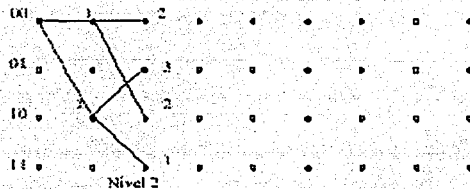


Figura 7.19 Segundo nivel del Diagrama de Trellis.

A partir del tercer nivel, en cada estado convergerán dos ramas, por lo que se analizan los errores en la métrica de cada camino posible en relación con la señal recibida.

TESIS CON
FALLA DE ORIGEN

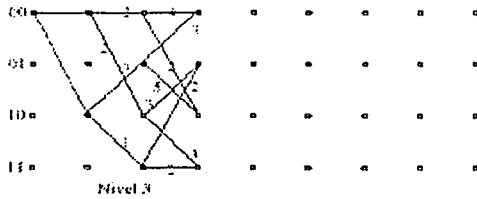


Figura 7.20 Tercer nivel del Diagrama de Trellis.

Según el número de errores acumulados en relación con la señal recibida, se elige un camino u otro. Nos quedamos con el de menor error acumulado. Así pues, se parte de menos ramas para el siguiente nivel del proceso:

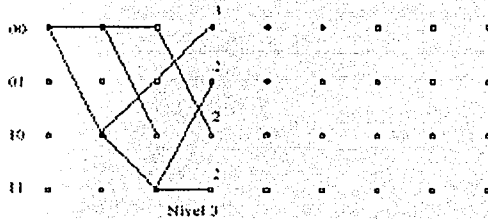


Figura 7.21 Tercer nivel elección de camino.

De cada nodo parten de nuevo otras dos ramas, colocando únicamente los errores correspondientes al último nivel, para una mejor visualización:

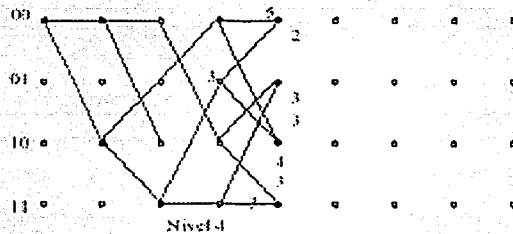
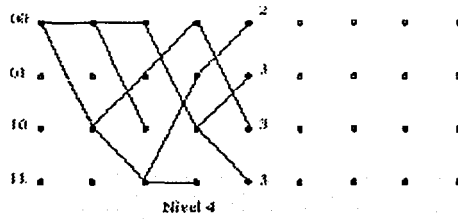


Figura 7.22 Cuarto nivel del Diagrama de Trellis.

Procede del mismo modo que para el tercer nivel y nos quedamos con las ramas de menor error acumulado:



TESIS CON FALLA DE ORIGEN

Figura 7.23 Cuarto nivel elección de camino

Con esta metodología aplicada al resto de niveles, las cuatro ramas seleccionadas que llegarían a los cuatro últimos estados dada su menor acumulación de error son:

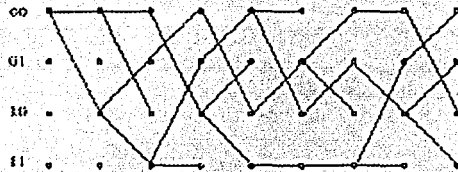


Figura 7.24 Nivel final del Diagrama de Trellis.

Por lo tanto, el trayecto óptimo es aquel que finaliza en el estado 11, ya que tiene tres errores acumulados frente a los cuatro de los otros estados. Así pues, el trayecto recorrido óptimo es:

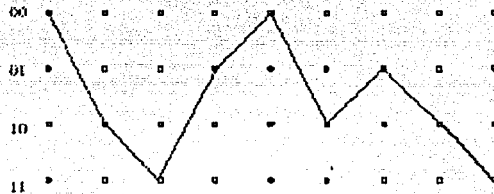


Figura 7.25 Camino óptimo elegido.

Observamos que coincide exactamente con la secuencia de datos enviada y codificada. Podemos concluir que mediante el Algoritmo de Viterbi hemos reconstruido perfectamente nuestra secuencia de datos a partir de una secuencia codificada con errores introducidos por el ruido del canal.

Arquitectura del Decodificador de Viterbi: El Decodificador de Viterbi se divide en tres partes funcionales (ver figura 7.26). La primera parte llamada por sus siglas en inglés Add-Compare-Select ACS (suma-compara-selecciona) es la unidad que es usada para calcular la métrica de un camino (cantidad de errores en el camino). La segunda parte es la Survivor Memory Unit (SMU) para el manejo de la memoria. Esta unidad se utiliza para almacenar las secuencias sobrevivientes (caminos con menor número de errores) y es la

última parte del Decodificador de Viterbi, la tercera parte es una memoria RAM común que auxilia al SMU

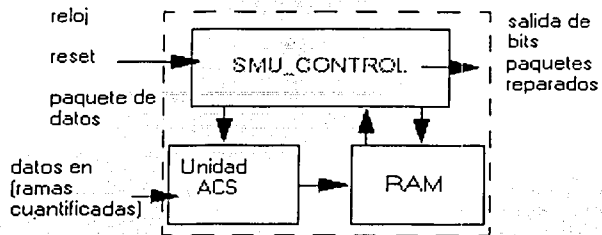


Figura 7.26 Bloque funcional de un Decodificador Viterbi.

El diagrama de bloques funcional del ACS se muestra en la figura 7.27, su función es sumar la métrica de la rama con su correspondiente métrica de camino para seleccionar el camino más pequeño (con menor numero de errores). El ACS también genera el indicador del estado previo (valor de decisión) y lo coloca en la memoria.

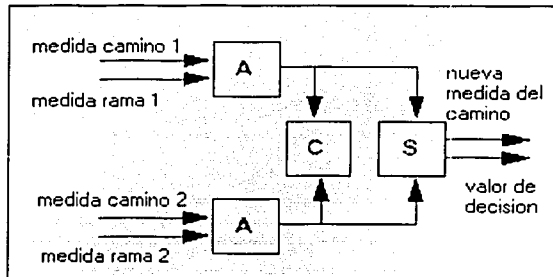


Figura 7.27 Bloque Add-Compare-Select ACS.

TESIS CON
FALLA DE ORIGEN

7.4.6 Des-entrelazado de datos.

En esta etapa se lleva acabo el proceso contrario al entrelazado realizado en la etapa de transmisión, es decir, en esta etapa se reciben los paquetes TS (Transport Stream Packet) que tienen una longitud de 208 bytes, los cuales se encuentran entrelazados (Figura 7.28). La forma en la cual se realiza este entrelazado en el transmisor es repartiendo estos 208 bytes en una trama de igual tamaño pero de una forma pseudo-aleatoria con un algoritmo.

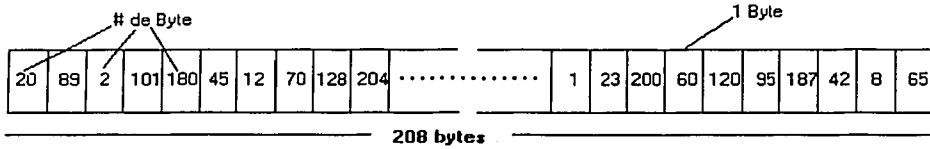


Figura 7.28 Trama de bytes entrelazados.

Como el ruido afecta a la trama de registros, aparecerán errores, los cuales afectarán cierta porción de la trama. Al des-entrelazar estos bytes por medio de este método el error será propagado a lo largo de la trama, es decir que se colocaran los bytes en su lugar y se tendrán bytes erróneos a lo largo del TS en lugar de perder todo un paquete de información. Además aumenta la eficiencia de la decodificación Reed-Solomon, ya que al llegar errores mas separados es mas probable que pueda repararlos. En la figura 7.29 a) se observa como el ruido afecta a un gran número de bytes que se encuentran juntos; si los datos no estuvieran entrelazados perderíamos una gran cantidad de información de un solo paquete, y sería imposible reparar ese error y por lo tanto imposible recuperar ese paquete. En cambio en la figura 7.29 b) se observa que después del des-entrelazado de datos el error se propaga en todos los paquetes pero se tienen 1 o 2 bytes erróneos por paquete, con esto se logra que no se pierda un paquete entero y el error sea más fácil de corregir.

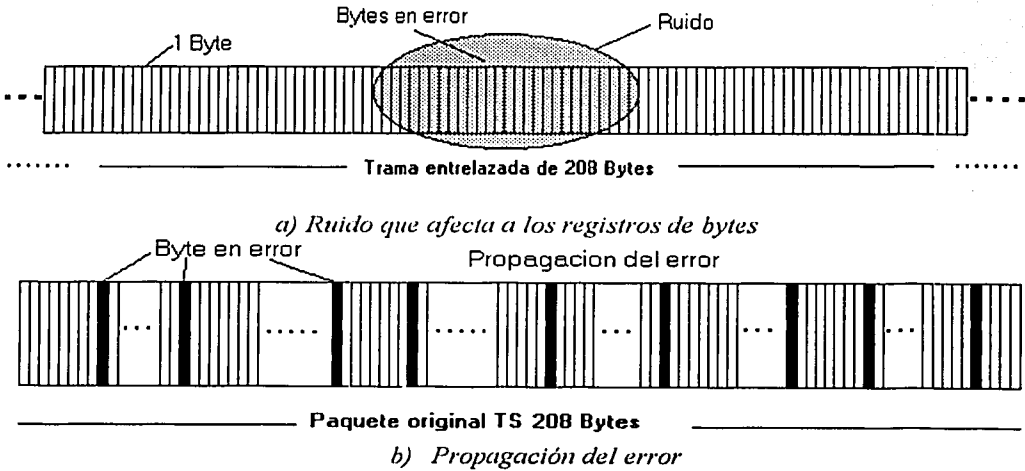


Figura 7.29 Des-entrelazado de Datos.

TESIS CON FALLA DE ORIGEN

7.4.7 Reed-Solomon FEC.

En esta etapa que se encuentra después del des-entrelazado de datos, se tiene el paquete original TS con sus 4 bytes de encabezado, 184 de datos y 20 de Reed-Solomon (Ver figura 7.30). Aunque los datos aún contienen algunos errores los cuales serán corregidos en esta etapa, los 20 bytes de Reed-Solomon se generan en el transmisor se toman los 188 bytes del TS para manipularlos matemáticamente como un bloque para generar identificadores (ID) del contenido del bloque. Estos ID ocupan 20 bytes los cuales son colocados al final del paquete original de 188 bytes, a estos 20 bytes se les conoce como bytes de paridad Reed-Solomon.

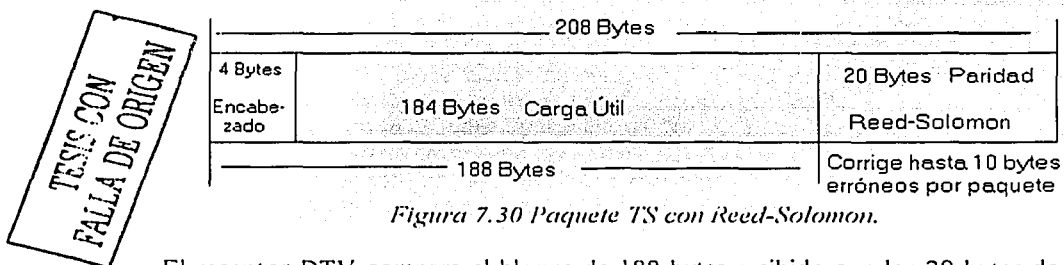
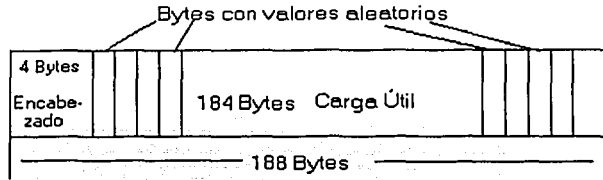


Figura 7.30 Paquete TS con Reed-Solomon.

El receptor DTV compara el bloque de 188 bytes recibido con los 20 bytes de paridad para determinar la validez de los datos recuperados. Si son detectados errores, el receptor determina que ID no corresponde al contenido del paquete y busca un paquete similar (que tenga pocos cambios en la posición de los bits). Por ejemplo diremos que este procedimiento es similar a las placas de los automóviles, un testigo describe el carro en el que huyeron los asaltantes de un banco, el dice que es un auto rojo Toyota 1999, con placas ABC123, en la base de datos de la policía no hay ningún registro de un auto Toyota 1999 rojo con estas placas o similares a ABC123, hay sin embargo el registro de un auto Nissan 1999 rojo con placas ABC123. La conclusión lógica es que el testigo confundió el tipo de auto, debido a que son autos similares, por lo cual la policía concluye que el auto visto fue el Nissan. El decodificador Reed-Solomon en el receptor DTV realiza una operación similar para comparar los bytes de paridad y determinar el paquete que fue transmitido de acuerdo con la opción mas lógica para corregirlo.

7.4.8 Des-aleatorización de datos.

En esta etapa se recibe el paquete de datos TS pero ahora con una longitud de 188 bytes debido a que los 20 bytes de paridad fueron ya utilizados por el decodificador Reed-Solomon (Ver figura 7.31). Pero este paquete de datos tiene sus bytes desordenados de una manera pseudo-aleatoria, debido al proceso aplicado en el transmisor (para lograr esto el transmisor aplica al valor de cada byte un cambio de acuerdo con un patrón de generación de números pseudo-aleatorios). Esto se hace con el fin de evitar, que los datos que contengan patrones de repetición en sus secuencias de bits, al ser modulados en RF, la energía se acumule solo en algunas frecuencias y en otras no exista energía a lo largo de canal de 6 MHz.



TESIS CON FALLA DE ORIGEN

Figura 7.31 Paquete TS con bytes de valores aleatorios.

El proceso de des-aleatorización consiste simplemente en aplicar el algoritmo de generación de números aleatorios inverso al que se aplicó en el transmisor con lo cual se recuperará al TS con sus datos, En el mejor de los casos, en el orden correcto y sin errores, estos datos son los paquetes de audio, video y datos que aun se encuentran multiplexados dentro de este paquete TS.

7.5 Transporte y demultiplexaje.

A continuación se describirá el bloque denominado transporte y demultiplexaje, el cual tiene como objetivo convertir el flujo de transporte (transport stream TS) en información digital pura de la compresión de video MPEG-2; el tren de bits elemental de audio; también el reloj de referencia del programa (PCR) y las marcas de tiempo (PTS), que son información sobre medida de tiempo; así como la información específica del programa (PSI) y las tablas de servicio de información (SI) relacionadas con la operación del servicio de difusión para la reconstrucción de los programas multiplexados en el TS. En la siguiente figura se muestra de forma más clara dicho proceso.

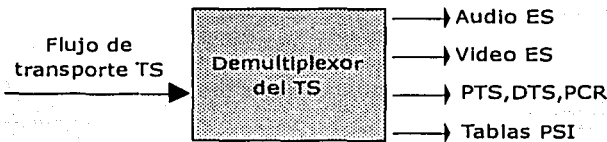


Fig. 7.32 Función básica del bloque de transporte y demultiplexaje.

El demultiplexor cumple una serie de pasos que se muestran en la figura 7.33

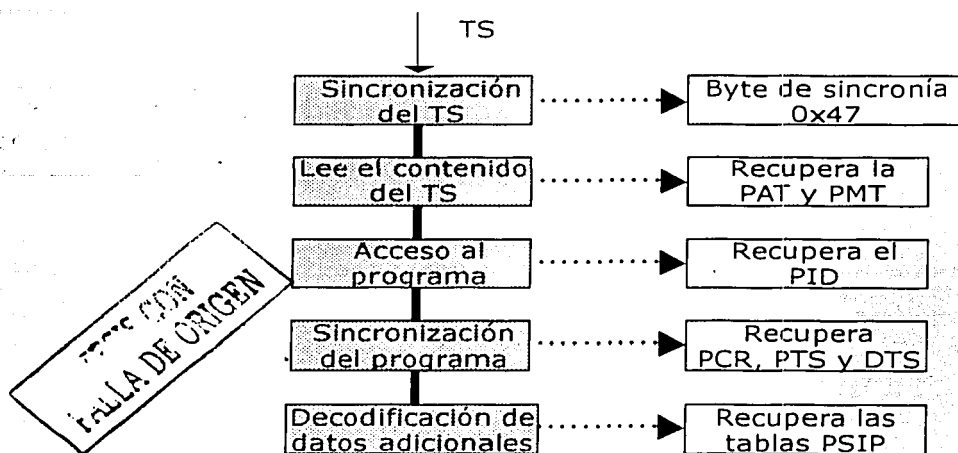


Fig. 7.33. Funciones del decodificador

La primera función del decodificador es la sincronización del flujo de transporte por medio de la detección del byte de sincronía marcado por el número 0x47. Se debe sincronizar dicho flujo en un máximo de 5 paquetes de información, para que el usuario no se de cuenta de dicha acción y para que sea un sistema eficiente, de acuerdo con las características de los sistemas que se encuentran diseñados en la actualidad.

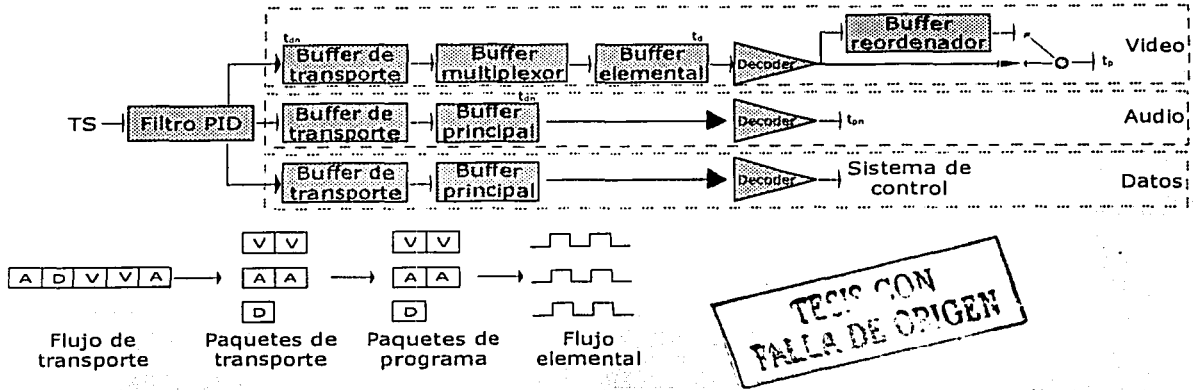
Posteriormente se procede a leer el contenido de cada paquete del flujo de transporte para poder recuperar las tablas PAT (Tabla de Asociación de Programa) y la PMT (Tabla de Mapa del Programa), que como se había mencionado con anterioridad son las que contienen los números de PID (Identificador de Paquete) de cada programa en el caso de la PMT, mientras que la PAT contiene los PID's de las tablas PMT que se ocupan en el proceso. Estas tablas tienen el "mapa" por así llamarlo, para la reconstrucción de los datos en el decodificador.

La tercera función es el acceso al programa, el cual se realiza con la clasificación de los paquetes de transporte por medio de su identificador o PID. (Cabe mencionar que esta etapa se realiza parcialmente antes de llegar a este punto del proceso, para poder identificar los paquetes que contienen las tablas).

Seguido de esta acción se tiene la sincronización del programa, que se realiza con base en las marcas de tiempo PCR (Referencia del Programa), PTS (Marca de Tiempo de Presentación) y DTS (Marca de Tiempo de decodificación), las cuales indicaran el tiempo y el orden en los cuales deben ser decodificados los datos de video y audio.

Para finalizar, se recuperan las tablas PSI que contienen información extra que manda la difusora según su estándar como son la hora, la programación, texto sobre los programas, por mencionar algunas.

El esqueleto del demultiplexor en su forma básica consta de los bloques mostrados en la figura 7.34, el cual se describirá en los siguientes párrafos.



TESIS CON FALLA DE ORIGEN

Fig. 7.34. Diagrama general de transporte y demultiplexaje.

Analizando el filtro PID, este debe realizar las siguientes funciones:

- Recuperar cada paquete de transporte del flujo
- Deberá encontrar el valor PID de cada paquete
- Con base en dicho valor mandará el paquete al buffer correspondiente.

Ya que se logra tener los valores constantes en tiempo del byte de sincronía, se dividirá el flujo de transporte en paquetes de transporte para que estos puedan ser analizados de manera independiente. Primero se analiza el identificador PID de cada paquete, para poder determinar el destino del paquete en algún buffer de transporte.

Los buffers de transporte reciben en la entrada los paquetes de transporte y proporcionan a la siguiente etapa paquetes de programa, como se muestra en la figura 7.34. Se realiza un análisis de la información agregada a los paquetes de programa que se convierten en paquetes de transporte en la etapa de multiplexaje y transporte del codificador. Dichos buffers tienen la capacidad de almacenar dos paquetes de transporte y se rigen por el modelo de tráfico leaky bucket^[2] para la transferencia de datos al siguiente buffer. Dicho modelo especifica que se tendrá una velocidad de transferencia de datos según el tipo de flujo que se tenga, por ejemplo para video en MPEG-2 se tendrá una velocidad fija según el nivel y el perfil en cuestión. Al regular las velocidades de los flujos de datos se controla el desbordamiento de paquetes a la siguiente etapa, siendo esta una característica primordial de este buffer. Cabe mencionar que solo se hace uso de la información de las cabeceras de los paquetes de transporte como son el PID, el campo de

adaptación, el indicador de inicio de carga, el byte de sincronía, el PCR, y los bytes de relleno, entre otros.

En el caso del audio y de los datos, del buffer de transporte se pasa al buffer principal, en el cual entran los paquetes de programa y salen los trenes elementales de información, tal como se muestra en la figura 7.34. Tiene como finalidad almacenar el flujo de entrada y mover dichos datos con base en las marcas de tiempo. Dichas marcas se obtienen de la cabecera del paquete de programa, además de obtener información como la longitud del paquete, bytes de relleno, campo opcional, el CRC y en algunos casos datos privados, entre otros.

No existen marcas de tiempo para retrasar los datos (DTS) en audio y los datos del sistema de control, solo existen marcas de tiempo de presentación al espectador (PTS), con esto se simplifica el trabajo realizado por los buffers.

Para el video se usan las marcas antes mencionadas, ya que para codificar una imagen B se tiene que codificar una imagen I y una P con anterioridad. En particular los datos de imágenes I y P son transferidos inmediatamente del buffer principal al decodificador en el tiempo denotado por DTS que se encuentra en la cabecera del paquete PES (Paquete Elemental). Esto quiere decir que saldrá del buffer principal cuando el PCR iguale a DTS para ser decodificado. Después de la decodificación se utiliza un buffer para reordenar y mostrar el retraso de la imagen I o P cuando el PCR iguale a PTS. El buffer principal es dividido en un buffer multiplexor y en un buffer de flujo elemental. El primero tiene como función al igual que el buffer de transporte direccionar y archivar la velocidad de transferencia de datos para que sea compatible con la siguiente etapa, basándose en el método leak que emplea el modelo leaky bucket para determinar la máxima velocidad de transferencia según el nivel y perfil usados, así como estudiar la cabecera del paquete PES.

Para el caso del buffer de flujo elemental, realiza la función de verificación de video de retardo (VBV) ^[3]. Dicha función transfiere los datos sobre las imágenes en una velocidad de bit calculada usando un código de retraso dentro del flujo de video. Este código de retraso consta de una cantidad de tiempo que debe de permanecer la imagen en el buffer antes de ser decodificada. Este método toma información de la tabla que contiene el mapa del programa, la cual especifica los valores de PID para la configuración del flujo de programa.

7.6 Decodificador de Video

En el capítulo 3 se describió el proceso de codificación de video usando MPEG-2. Este proceso entrega una trama de bits conocida como trama elemental de video. La decodificación es esencialmente el proceso inverso al de la codificación. Se llevan a cabo los mismos procedimientos, pero en orden inverso. Un decodificador de video tiene como entrada la trama elemental de video, la cual se va a procesar hasta llegar a la señal de video original. A continuación se explican los bloques del decodificador de video mostrado en la figura 3.25 (Capítulo 3 – Compresión de Video) de forma más detallada.

7.6.1 Buffer

Un buffer o "memoria temporal" es un área de datos compartida por dispositivos de hardware o procesos de programas que operan a distintas velocidades o con diferentes conjuntos de prioridades. El buffer permite que cada dispositivo o proceso opere sin verse interferido por otro. Para que un buffer sea efectivo, su tamaño y los algoritmos para agregarle y retirarle datos deben ser considerados. Un buffer es "un punto intermedio de almacenamiento" que existe para apoyar la coordinación de actividades separadas. Las memorias buffer trabajan con memoria de acceso serial, es decir que los datos se encuentran normalmente almacenados en el orden en que serán utilizados.

La memoria intermedia (buffer) de entrada recibe los datos del canal de transmisión ya en forma de una trama de video individual. En el buffer son almacenados los datos de la trama para permitir variaciones temporales en la tasa de transmisión. Si la tasa de transmisión experimenta un incremento muy grande, el buffer tenderá a llenarse. Para prevenir un sobreflujo el proceso de cuantización será más exigente. Igualmente cuando en el buffer haya un subflujo o "underflow", es decir que se esté vaciando, la cuantización será menos exigente para mantener una tasa de transmisión promedio. El buffer de entrada del decodificador contiene exactamente lo que contenía el buffer de salida usado en la codificación.

Para el decodificador de MPEG-2 Phillips SAA7201, el buffer de entrada tiene una capacidad de 2.6 Mbits. En el decodificador SF986 de EEC la memoria buffer de entrada es de 2 Mbits.

7.6.2 Demultiplexaje

Como se muestra en la figura 7.35 la TS contiene paquetes PES, que son los que contienen la información que necesita el decodificador de video. Dentro de los encabezados de la Trama Elemental Empaquetada (PES - Packetized Elementary Stream), que es un paquete con tramas elementales, hay un apuntador a la primera imagen que hay dentro de este paquete, que es con la cual se va a comenzar a efectuar el proceso de descompresión de video. También se encuentra el DTS (Decoding Time Stamp) que es la parte del encabezado que indica cuando se decodificarán los datos de video, y el PTS (Presentation Time Stamp) que señala el tiempo en que la imagen se mostrará al usuario.

Toda la información que resulta del proceso de compresión se encuentra en la estructura de datos ES de video. Los datos que lleva consigo la trama elemental no sólo son los bits de información de las imágenes. A partir de la trama de bits elemental o secuencia de video, se demultiplexa la información que esta transporta, para obtener los vectores de compensación de movimiento, los datos codificados espacialmente, las tablas de cuantización y los datos sintácticos. En la figura 7.36 se muestra el formato de la trama elemental de video.

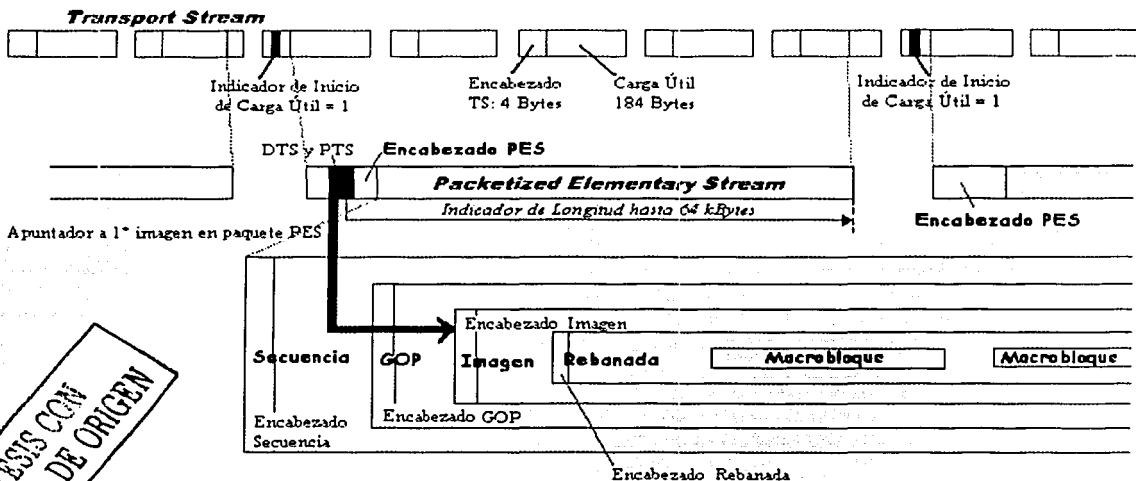


Figura 7.35 Estructura de la Trama Elemental de Video

Dentro de los encabezados de la ES de video se encuentran parámetros como el tamaño horizontal y vertical de la imagen, si se usa un barrido progresivo o entrelazado, el perfil, nivel, velocidad de transferencia de bits, y cuales matrices de cuantificación se usan para codificar imágenes espaciales y temporales. Cabe mencionar que los valores propios de la imagen de video, con los cuales se reconstruye la imagen, y los vectores de movimiento se encuentran únicamente en el campo de datos de imagen.

Sin los datos del encabezado de la secuencia, un decodificador no puede entender la trama de bits, y por lo tanto estos llegan a ser los puntos de entrada en que los decodificadores pueden comenzar su correcta operación.

Trama Elemental de Video MPEG-2

Encabezado Secuencia	Extensión Secuencia	Extensión y Datos Usuario	Encabezado GOP	Extensión y Datos Usuario	Encabezado Imagen	Extensión Codificación Imagen	Extensión y Datos Usuario	Datos Imagen	Fin Secuencia
----------------------	---------------------	---------------------------	----------------	---------------------------	-------------------	-------------------------------	---------------------------	--------------	---------------

Figura 7.36 Secuencia de Video MPEG-2

Para decodificar el video es necesario efectuar el proceso inverso que se hizo en el codificador, para llegar de la secuencia de video hasta los bloques de coeficientes. El proceso de demultiplexaje o separación de los componentes de la trama elemental de video se muestra en la figura 7.37. Como se dijo antes, una secuencia de video está formada por grupos de imágenes. Entonces, una secuencia de video se va a separar en los grupos de imágenes que la forman.

Las imágenes se obtienen a partir de los grupos de imágenes. Antes de transmitir las imágenes se reordenan. Los datos de las imágenes P se envían antes de los de las imágenes B. Las últimas imágenes B de un GOP no pueden ser transmitidas hasta que se haya transmitido la imagen I del siguiente GOP, debido a que estos datos serán necesarios para efectuar la decodificación bidireccional. En el decodificador las imágenes llegarán con un orden distinto al original, pues para realizar la predicción de imágenes es necesario que la imagen futura ya se encuentre en el decodificador, pues este no hace una predicción literal, sino que se basa en las imágenes futuras para crear una imagen. Para poner en orden a las imágenes de nuevo, se incluye una referencia temporal en cada una de ellas, que el decodificador tomará en cuenta. El encabezado de las imágenes contiene el tipo de imagen que se trata (I, P o B).

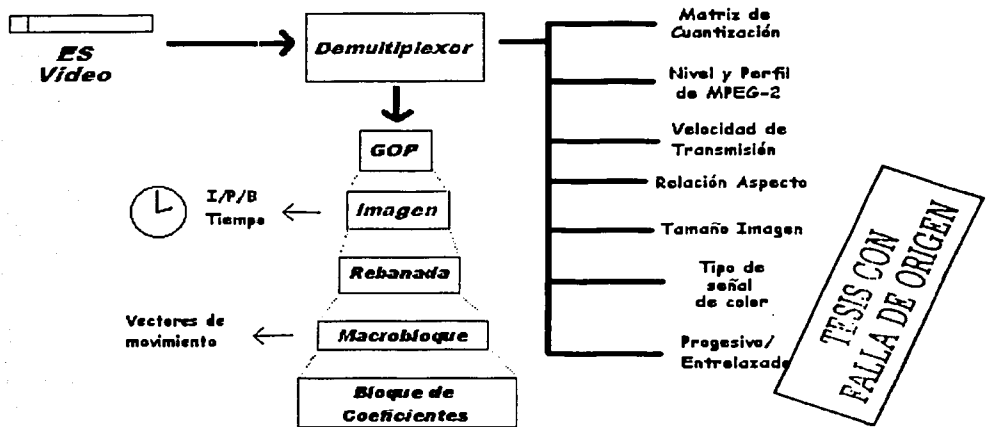


Figura 7.37 Demultiplexaje de la secuencia de video

Las rebanadas o slices, que representan una fila horizontal que está ordenada de izquierda a derecha, se obtienen a partir de las imágenes. Como las rebanadas son la unidad fundamental de sincronización para la codificación de la longitud variable y diferencial, si hay un bit erróneo en la trama elemental, puede ocurrir una falla en los símbolos del código de longitud variable, o que los coeficientes y vectores codificados diferencialmente sean incorrectos. La estructura de rebanadas permite una recuperación al proporcionar un punto de resincronización en la trama de bits.

Una rebanada está formada por macrobloques, por lo que a partir de las rebanadas se pueden obtener los macrobloques. Los macrobloques son la unidad fundamental de la imagen y contienen en su encabezado un vector de movimiento de dos dimensiones. A

TESIS CON FALLA DE ORIGEN

partir de los macrobloques se obtienen los bloques individuales de coeficientes de la transformada discreta del coseno.

Una vez que se sabe cual es la primera imagen que contiene la trama de video, se decodifica. La decodificación de las imágenes I se hace completamente por un proceso de decodificación espacial. Las imágenes B y P se decodifican primeramente espacialmente, y posteriormente se le añaden los datos de las imágenes con las que están relacionadas, los vectores de movimiento y la compensación de errores.

7.6.3 Decodificación Espacial

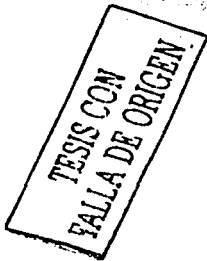
❖ Decodificación VLC y RLC

Los datos que se reciben en el decodificador se encuentran en forma de una trama de bits, la cual está codificada con los códigos de entropía VLC y RLC. Con el VLC se le asignaron a los coeficientes que ocurren con más frecuencia palabras de bits más cortas y con el RLC se nos dice cuantos ceros hay en las cadenas de estos, que ocurren al aplicar la DCT. En la figura 7.38 se esboza la decodificación de la trama de acuerdo con los códigos VLC y RLC.

Al decodificar los bits se tiene que llegar a la trama de bits original. Usando una tabla del código de Huffman o de Longitud Variable (VLC), se decodifican los coeficientes con lo que se obtiene los coeficientes de DC y los coeficientes codificados con RLC.

Después se hace la decodificación RLC, con lo que se vuelven a tener las cadenas de ceros que serán colocadas en sus lugares correspondientes en los bloques de coeficientes de la transformada de 8×8 . Un código RLC está formado por un par de números, el primero indica el número de ceros consecutivos, y el segundo el valor que se encuentra después de la cadena de ceros. El código RLC (5,8) representa la secuencia {0, 0, 0, 0, 0, 8}.

Para colocarlos en su posición original se hace el barrido. En la figura 7.40 se muestra como quedan los coeficientes anteriores en su posición original, después de realizar el barrido en Zig-Zag.



173	0	0	0	0	0	0	0
6	-1	-2	0	0	0	0	0
0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
-1	0	0	0	0	0	0	0
0	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 7.40 Coeficientes de la transformada (cuantizados) después de realizar la exploración en Zig-Zag

Una vez que se ha reconstruido el bloque de coeficientes de la transformada se tiene que decuantizar.

❖ Decuantización

En el codificador los coeficientes de la transformada fueron divididos por la matriz de cuantización, dicha división se efectúa de acuerdo con la localización del coeficiente, es decir que el coeficiente DCT (1,1) se va dividir por el valor localizado en la posición (1,1) de la matriz de cuantización. La matriz de cuantización se nombra con la letra Q. Los valores de los coeficientes también pueden ser divididos por algún factor de escala que permita mayor compresión, este factor de escala se aplica a todo el bloque de 8x8 coeficientes excepto al de DC (que es el que ocupa la posición (0,0)), pues este coeficiente debe tener la mayor precisión posible.

Existen diferentes matrices de cuantización para los bloques, según el tipo de imagen con la que se está trabajando. La matriz de cuantización de las imágenes I se muestra en la tabla 7.2 y la matriz de cuantización de las imágenes P y B se muestra en la tabla 7.3.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Tabla 7.2 Matriz de cuantización para bloques pertenecientes a imágenes I

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16



Tabla 7.3 Matriz de cuantización para bloques dentro de imágenes P y B.

Dentro del decodificador es necesario decuantizar los coeficientes de la transformada, para ello se multiplicarán por el factor de escala, si es que se aplicó alguno, y posteriormente cada coeficiente será multiplicado por el valor de la matriz de cuantización que le corresponde de acuerdo con su localización.

En vista de lo anterior y tomando los coeficientes cuantizados del ejemplo que estamos utilizando (ver figura 7.40), primero se multiplica el bloque por su factor de escala (para este ejemplo vale 2) y después se multiplica cada uno de los coeficientes del bloque por el valor que le corresponde en la matriz de cuantización para imágenes I que se muestra en la tabla 7.2. Es importante notar que al coeficiente de DC no se le multiplica por el factor de escala. El proceso anterior da como resultado los coeficientes de la DCT que se muestran en la tabla siguiente (tabla 7.4).

1384	0	0	0	0	0	0	0
192	-32	-88	0	0	0	0	0
0	0	0	0	0	0	0	0
88	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
-52	0	0	0	0	0	0	0
0	-54	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Tabla 7.4 Coeficientes de la transformada en el decodificador

Los coeficientes de la transformada originales se muestran en la tabla 7.5; se puede apreciar que existe una pequeña diferencia entre los coeficientes obtenidos en el decodificador y los originales, debido al truncamiento que se realizó en el codificador al momento de aplicar la cuantización. Tomemos el coeficiente colocado en la posición (1,0) de la tabla 7.5 que es 216, si este coeficiente lo dividimos por el valor que le corresponde de la matriz de cuantización $Q(1,0)$ que es 16 (ver tabla 7.3), nos da como resultado 13.5. Sin embargo este resultado se trunca arrojando un 13. Después se divide el 13 por el factor de escala 2, quedándonos un valor cuantizado de 6.5, que también es truncado. Como se ve dicho coeficiente cuantizado es igual a 6, que es el valor que se recibió en el decodificador (véase la figura 7.40).

1384	0	0	0	0	0	0	0	0
216	-36	-99	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
99	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
-58	0	0	0	0	0	0	0	0
0	-60	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1

Tabla 7.5 Coeficientes de la transformada originales

El proceso de decuantización de los coeficientes de la transformada DCT se puede resumir como se muestra en la figura 7.41.

Coeficientes DCT
Cuantizados
 $QF(v,u)$



$$QF(v,u) \cdot \text{Factor de Escala} \cdot Q(v,u)$$



Coeficientes DCT
 $F(u,v)$

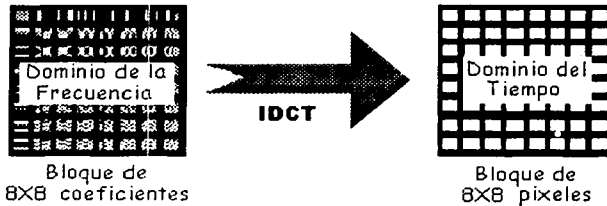
TESIS CON
FALLA DE ORIGEN

Figura 7.41 Proceso de decuantización de los coeficientes de la transformada

❖ Transformada DCT inversa

Una vez que se tienen los coeficientes de la transformada DCT, se efectúa la transformada inversa DCT para cambiar la imagen del dominio de la frecuencia al dominio del tiempo, como se muestra en la figura 7.42. La DCT se encuentra dentro de la familia de la DFT

(Discrete Fourier Transform), y permite una descripción de la señal en el dominio de la frecuencia.^[4,5,6] En el capítulo 4 de este trabajo se explicaron los procesos necesarios para realizar la DCT dentro de la compresión de video, como la segmentación de la imagen en bloques.



TESIS CON
FALLA DE ORIGEN

Figura 7.42 Bloque de la IDCT

La transformada discreta del coseno bidimensional para un bloque de 8x8 pixeles (representado por $f(x,y)$) se define como se muestra a continuación:

$$F(u,v) = \frac{C(u)C(v)}{4} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

donde $C(u) = C(v) = \frac{1}{\sqrt{2}}$ para $u = v = 0$,
 $C(u) = C(v) = 1$ en cualquier otro caso.

En seguida se muestra un algoritmo en Matlab para realizar la DCT; la entrada es la matriz f de 8x8 valores, la matriz de salida se llama sal .

```
for u=1:8
    for v=1:8
        sal(u,v)=0;
        for x=1:8
            for y=1:8
                sal(u,v)=sal(u,v)+(f(x,y)*cos((2*(x-1)+1)*(u-1)*pi/16)*
                    cos((2*(y-1)+1)*(v-1)*pi/16));
            end
        end
    end
    if (u==1)
        cu=1/sqrt(2);
    else
```

```

        cu=1;
    end

    if (v==1)
        cv=1/sqrt(2);
    else
        cv=1;
    end

    sal(u,v)= sal(u,v)+0.25*cu*cv;
end
end
end

```

La IDCT o Transformada Discreta del Coseno Inversa se aplica en el decodificador para recuperar el valor de los datos de la señal de video en el dominio del tiempo que hasta este momento se encuentra en el dominio de la frecuencia. Si la matriz de coeficientes transformados de 8×8 se encuentra nombrada por $F(u,v)$, su IDCT se define como se muestra en la expresión siguiente:

$$f(x,y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(u,v) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

en donde $C(u) = C(v) = \frac{1}{\sqrt{2}}$ para $u = v = 0$,
 $C(u) = C(v) = 1$ en cualquier otro caso.

A continuación se muestra un algoritmo en Matlab para realizar la IDCT; la entrada es la matriz de coeficientes transformados **sal**, y su salida es la matriz **ent**, en la que se recupera la matriz "original" de la señal digital de video (hay que recordar que hay error de cuantización pues se hicieron truncamientos).

```

for x=1:8
    for y=1:8
        ent(x,y)=0;
        for u=1:8
            for v=1:8

                if (u==1)
                    cu=1/sqrt(2);
                else
                    cu=1;
                end

                if (v==1)
                    cv=1/sqrt(2);
                else

```

TESIS CON
FALLA DE ORIGEN

```

        cv=1;
    end
    ent(x,y)=ent(x,y)+(sal(u,v)*cv*cv*cos((2*(x-
1)+1)*(u-1)*pi/16)*
        cos((2*(y-1)+1)*(v-1)*pi/16));
    end
end
ent(x,y)=ent(x,y)*0.25;
end
end

```

La IDCT nos entrega bloques de 8x8 píxeles con los valores de las señales de video digital (Y, Cr o Cb) en el dominio del tiempo. Para obtener la salida digital de video, los bloques deben almacenarse en memoria RAM para reconstruir cada imagen.

7.6.4 Reconstrucción de la señal de video 4:2:0

Por medio de la transformada inversa, los bloques de 8x8 píxeles son recuperados. Para obtener la salida, los bloques son almacenados en memoria DRAM (Dynamic Random Acces Memory) o SDRAM (Synchronous Dinamic Random Acces Memory), en donde se lee una línea a la vez, estas líneas se agrupan hasta formar la imagen final.

Los circuitos integrados DRAM o RAM (Random Acces Memory) Dinámica contienen arreglos rectangulares de celdas de memoria, en las cuales se pueden leer y escribir datos, además tienen circuitos que permite refrescar y mantener la integridad de los datos almacenados. Los arreglos de memoria están organizados en renglones y columnas de celdas de memoria, llamados respectivamente líneas de palabra y líneas de bits. Cada celda de memoria tiene una sola dirección definida por la intersección de un renglón y una columna. La SDRAM o DRAM Síncrona difiere de los tipos anteriores de RAM en que no funciona de forma asíncrona al reloj del sistema, como hacen las memorias convencionales. La SDRAM está sujeta al reloj del sistema y está diseñada para que pueda leer o escribir desde la memoria en forma de ráfaga, es decir, sin interrupciones.

La necesidad de memoria en el decodificador es de unas 3 imágenes (dos imágenes de referencia más la imagen en vía de reconstrucción), siendo para 4:2:0, de aproximadamente 16 Mbits. Este es el requerimiento de memoria SDRAM total para efectuar la decodificación, y sirve para colocar los datos del canal, la información decodificada y el espacio para el procesamiento del decodificador.

Se han implementado técnicas para reducir la cantidad de memoria, las cuales han permitido llegar de 16 Mbits a 8 Mbits. Existe una tecnología llamada MPEG Memoria Reducida (REMPEG: Reduced Memory MPEG), que reduce la memoria hasta unos 4 Mbits.

Como se dijo antes la señal de video está compuesta por las componentes de luminancia y de diferencia de color, de modo que se almacenan en memoria los bloques de cada una de estas componentes, para efectuar el proceso inverso que se hizo en la codificación. Las muestras de color en 4:2:0 se posicionan entre las muestras de luminancia en los ejes horizontal y vertical, tomando en cuenta el submuestreo que se realizó, para que se encuentren igualmente espaciadas cuando se utiliza una fuente entrelazada. La señal de video digital a la salida queda como se muestra en la figura 7.43.

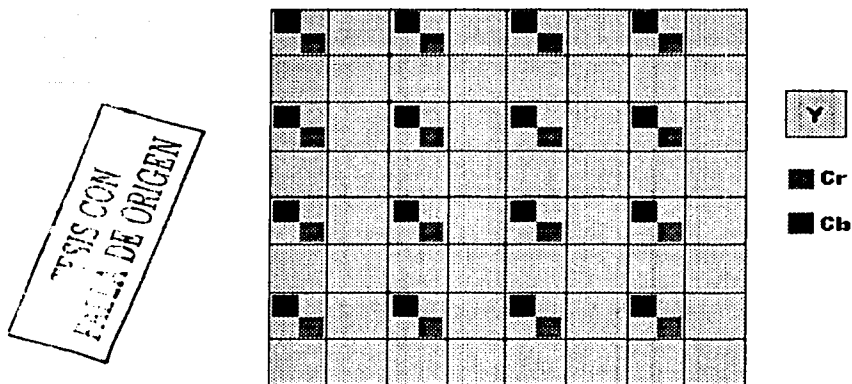


Figura 7.43 Salida de video con formato 4:2:0

Las imágenes a la salida del decodificador usando el Main Level/Main Profile, ya se encuentran en el dominio del tiempo, es decir que se encuentran en forma de píxeles, y tienen una dimensión de 720 píxeles por 480 líneas (es la parte de la imagen visible en la pantalla). Para visualizar la imagen en un televisor NTSC, se tiene que hacer una conversión D/A de las señales de video.^[7]

7.6.5 Compensación de movimiento y predicción de imágenes

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

Para realizar la decodificación se toma en cuenta el DTS que se encuentra en el encabezado del PES. El DTS indica el instante en que una imagen debe ser decodificada. Cuando una secuencia IPBB es recibida, las imágenes I y P deben ser decodificadas antes que la primera imagen B. La decodificación empieza con la primera imagen I y la imagen P siguiente, que una vez decodificadas serán almacenadas en la memoria del decodificador para que se utilicen como referencia para las imágenes P o B que dependen de ellas.

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. Como se dijo en el capítulo#3 las imágenes P están relacionadas con las imágenes I y P anteriores, mientras que las imágenes B dependen de imágenes I y P anteriores y posteriores a esta. Los datos de las imágenes P consisten en vectores que describen de que parte de la imagen anterior puede ser tomado cada macrobloque y coeficientes transformados que describan la corrección que debe añadirse a los macrobloques. También hay que tener presente que las imágenes fueron reordenadas para que las imágenes "futuras", de las cuales dependen las imágenes B, se transmitan antes que estas y pueda realizarse la decodificación. Los datos de las imágenes B son vectores que describen de que porciones de las imágenes pasadas y futuras se tomarán los datos. También contiene coeficientes transformados para corregir los macrobloques.

La compensación de movimiento se envía hacia el decodificador en forma de vectores de movimiento, los cuales señalan la parte de la imagen pasada, o futura, que va a ser desplazada hacia su nueva posición en la imagen nueva. Los vectores de movimiento se encuentran en el encabezado de los macrobloques. De acuerdo con el concepto de la codificación bidireccional, tomando como base los macrobloques, una imagen bidireccionalmente codificada puede obtener compensación de movimiento de acuerdo con las imágenes pasadas o futuras, por medio del vector de movimiento.

En el receptor se decodifican los vectores de movimiento así como los datos de diferencia para corrección de los macrobloques de las imágenes previstas. Esta información se añade a los macrobloques de las imágenes que se toman como base, y dan como resultado los cambios necesarios para obtener los macrobloques de la imagen que se está reconstruyendo (ver la figura 7.44).

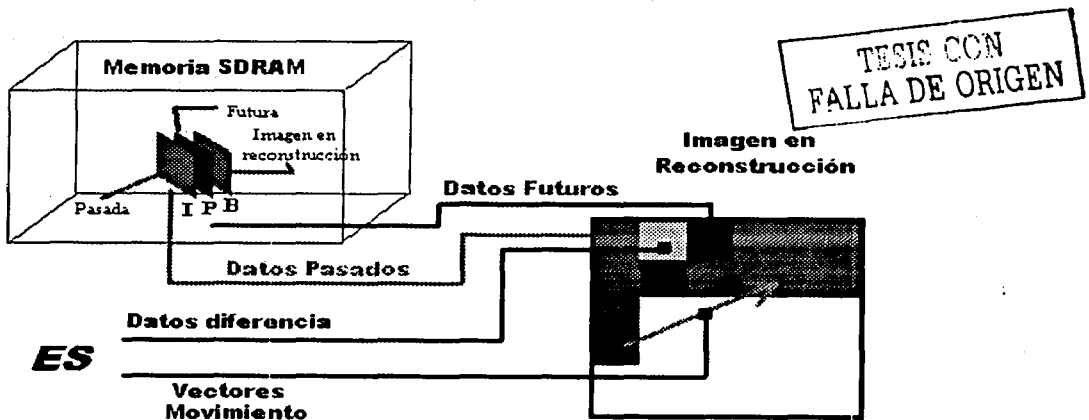


Figura 7.44 Predicción de imágenes y compensación de movimiento

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

7.6.6 Reordenación de imágenes.

La ultima etapa de la decodificación es poner las imágenes en el orden inicial de visualización. Para hacer esto se recurre al PTS que se encuentra en el encabezado del PES, el cual indica en que momento debe de ser presentada una imagen a la salida del decodificador. Esto se debe hacer pues en el proceso de codificación hubo un proceso de reordenamiento. Las imágenes B son decodificadas y presentadas simultáneamente de tal forma que solo contienen PTS. Cuando se recibe una secuencia IPBB, las imágenes I y P deben ser decodificados antes que la primera imagen B. Un decodificador solo puede decodificar una imagen a la vez; por lo tanto la imagen I es decodificada y almacenada. Mientras la imagen P está siendo decodificada, la imagen I decodificada se entrega a la salida para que pueda ser seguida por las imágenes B. La figura 7.45 muestra que cuando una trama elemental de video que contiene una imagen I es recibida, tendrá en el encabezado las marcas de tiempo DTS y PTS, las cuales se encuentran separadas por el periodo de una imagen. Como se está usando codificación bidireccional, una imagen P debe seguir a la imagen I, y está también contiene PTS y DTS, pero la separación entre las dos marcas de tiempo es de un periodo de tres imágenes para permitir la intercalación de las imágenes B. Entonces cuando se recibe una secuencia IPBB, la imagen I tiene un retardo del tiempo de una imagen, la imagen P es retrasada por un periodo de tres imágenes, a las dos imágenes B no se les aplica retardo, y la secuencia de presentación llega a ser IBBP.

	I	P1	B1	B2	P2	B3
PTS	N+1	N+4	N+2	N+3	N+7	N+5
DTS	N	N+1	—	—	N+4	—

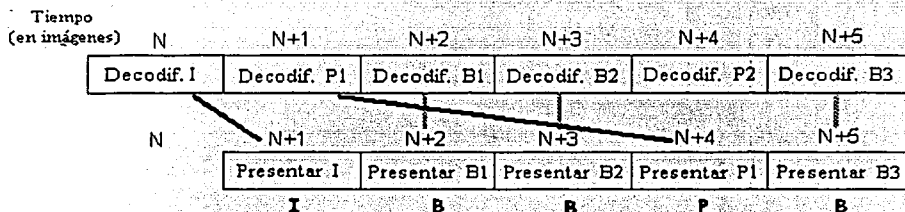


Figura 7.45 Reordenación de las imágenes

Una vez que se hace la reordenación se tienen a la salida del decodificador las imágenes de video digital del programa de televisión que quiere ver el usuario.

TESIS CON FALLA DE ORIGEN

7.7 Decodificador de Audio

Un resumen del proceso de decodificación de AC-3 se muestra en el pseudocódigo siguiente:

Alineación del Frame AC-3/Verificación CRC

Desempaquetar los datos BSI

For bloque_de_audio = 1 to 6

{

Desempaquetar datos fijos

For canal = 1 to Número_de_canales_codificados

{

Desempaquetar exponentes

For banda = 1 to Número_de_bandas

{

Calcular la asignación de bits

Desempaquetar mantisas

Escalar mantisas/Desacoplar

Desnormalizar mantisas con exponentes

}

Calcular la transformada inversa parcial

Mezclado para obtener canales de salida apropiados

}

For canal = 1 to Número_de_canales_de_salida

{

Añadir ventana y traslape con un buffer de retraso

Almacenar las muestras PCM en un buffer de salida

Copiar los valores del buffer de mezclado al buffer de retraso

}

}

A fin de mantener la memoria y los requerimientos de latencia del codificador tan pequeños como sea posible, cada paquete de AC-3 (Audio Coding –3) es decodificado en una serie de “ciclos” anidados.

El primer paso establece el alineamiento del paquete. Esto implica encontrar la palabra de sincronía, y entonces confirmar que las palabras CRC (Cyclic Redundance Check: Procedimiento de detección de errores por chequeo de redundancia cíclica)^[8] no indiquen errores. Una vez que el paquete de sincronía es encontrado, los datos BSI son desempaquetados para determinar la información importante del paquete tal como el número de canales codificados.

El paso siguiente es desempaquetar cada uno de los seis bloques de audio. A fin de minimizar los requerimientos de memoria de los buffers de salida PCM (Pulse Codec Modulation), los bloques de audio son desempaquetados uno a la vez. Al final del periodo de cada bloque los resultados PCM son copiados a los buffers de salida, los cuales se

encuentran generalmente con dos buffers a la salida para interrumpir directamente el acceso a los convertidores D/A del sistema.

7.7.1 Tiempo de sincronización del decodificador.

La figura 7.46 muestra las etapas de regulación de tiempo usadas para la sincronización del decodificador AC-3. Mediante el diseño, todos los codificadores AC-3 deben asegurar que los primeros dos bloques de audio estén completamente contenidos dentro de los primeros 5/8 del paquete. También la primera palabra CRC protege solo los primeros 5/8 del paquete. Esto ayuda a reducir la latencia del decodificador, pues puede comenzar a desempaquetar antes de que el paquete haya sido completamente recibido.

Una vez que los primeros 2/3 del paquete hayan sido recibidos, el decodificador comienza a verificar la primera palabra CRC. Si no hay errores, el decodificador trabajará con el Bloque 0 y copiará las muestras PCM reconstruidas al buffer de salida. Al final del periodo de dicho bloque (5.33 ms para 48 kHz de frecuencia de muestreo del sistema), el manejador de la interrupción a la salida comienza pasar las muestras del Bloque 0 hacia el DAC (Convertidor Digital/Anologo), y puede empezar la decodificación del Bloque 1.

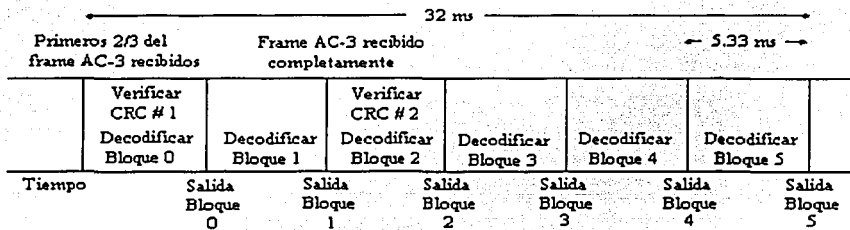


Figura 7.46 Regulación del Tiempo de Sincronización de la Trama AC-3

Cuando haya terminado el periodo de decodificación del Bloque 1, el paquete completo habrá sido recibido por el decodificador y la segunda palabra CRC puede ser verificada. Como puede verse, este método proporciona una latencia total en el decodificador igual al periodo de un bloque mas el tiempo para recibir 2/3 del paquete de entrada (27 ms asumiendo una entrada continua con una frecuencia de muestreo del sistema de 48 kHz).

7.7.2 Procesamiento de datos en el decodificador.

El procesamiento del bloque de audio puede dividirse en dos etapas distintas, denominadas procesamiento a la entrada y a la salida. El procesamiento a la entrada incluye todo el desempaquetamiento de la trama de bits y la manipulación del canal codificado. El procesamiento a la salida se refiere primeramente a las etapas de la ventana y adición de traslape de la transformada MDCT inversa.

Esta diferencia se hace debido a que el número de canales de salida generados por un decodificador AC-3 no coincide necesariamente con el número de canales codificados en la trama de bits. Al usar una técnica llamada "downmixing" o mezclado, un decodificador puede aceptar una trama de bits con cualquier número de canales codificados y producir un número arbitrario de canales de salida. El procesamiento a la entrada trabaja en una base por canal codificado, mientras el procesamiento a la salida funciona en una base por canal de salida.

❖ **Procesamiento a la entrada del decodificador**

El procesamiento a la entrada comienza cuando el decodificador desempaqueta el bloque de datos de audio fijo, el cual es una colección de parámetros y banderas localizadas al inicio del bloque de audio. Estos datos fijos incluyen elementos tales como banderas de encendido, información de acoplamiento, exponentes y los parámetros de los bits de asignación. El término "datos fijos" se refiere al hecho de que el tamaño de las palabras para estos elementos de tramas de bits es conocido a priori, y no requiere que se lleve a cabo una asignación de bits.

Los exponentes componen el campo más grande en la región de datos fijos, ya que incluyen todos los exponentes por cada canal codificado. Dependiendo del modo de codificación, pueden haber tantos como uno por mantisa, y pueden existir hasta 253 mantisas por canal. En lugar de desempaquetar todos estos exponentes a una memoria local, es preferible guardar los apuntadores de los campos de exponentes, y desempaquetarlos hasta que sean necesarios, un canal a la vez.

Una vez que los datos fijos se encuentran desempaquetados, el decodificador comienza a procesar cada canal codificado (véase la figura 7.47). Primero, los exponentes para un canal dado son desempaquetados a partir del paquete de entrada. Un cálculo de asignación de bits es entonces efectuado, en el se toman los exponentes y los parámetros de asignación de bits y se calcula el tamaño de las palabras para cada mantisa empaquetada. Entonces se desempaquetan las mantisas a partir del paquete de entrada. Las mantisas son ajustadas para proporcionar el rango dinámico de control apropiado (y para deshacer la operación de acoplamiento, si es necesario), y entonces se denormalizan los exponentes. Finalmente, se calcula una transformada inversa parcial, y los resultados son mezclados en los buffers apropiados para el procesamiento de salida que se realiza después.

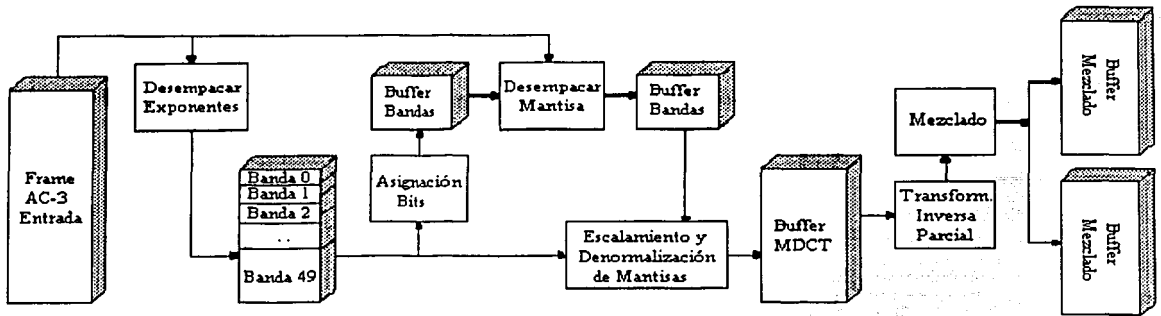


Figura 7.47 Procesamiento a la entrada del decodificador

TESIS CON
 FALLA DE ORIGEN

En el primero de estos pasos, los exponentes para cada canal individual son desempaquetados dentro de un buffer con un tamaño de 256, llamado buffer MDCT. Estos exponentes son entonces agrupados en tantas bandas como 50 para propósitos de asignación de bits. El número de exponentes en cada banda incrementa mientras las frecuencias de audio son más grandes, siguiendo de manera burda una división logarítmica que modela las bandas críticas psicoacústicas.

Para cada una de estas bandas de asignación de bits, los exponentes y los parámetros de asignación de bits son combinados con el propósito de generar un tamaño de palabra de la mantisa, para cada mantisa en esa banda. El tamaño de estas palabras se almacena en el buffer de la banda con un tamaño de 24 (la banda de asignación de bits más amplia está construida por 24 compartimentos de frecuencia). Una vez que se ha calculado el tamaño de las palabras, las mantisas correspondientes son desempaquetadas a partir del paquete de entrada y almacenadas dentro del buffer de la banda. Finalmente estas mantisas son adaptadas y denormalizadas por el correspondiente exponente y escritas dentro del buffer MDCT. Después que todas las bandas han sido procesadas, y todas las mantisas han sido desempaquetadas, cualquier lugar restante en el buffer MDCT es llenado con ceros.

En este punto, una transformada inversa parcial se lleva a cabo y se coloca dentro del buffer MDCT. La transformada inversa parcial se realiza mediante una etapa de premultiplicación compleja de 128 puntos, una FFT (Fast Fourier Transform) inversa compleja de 128 puntos, y una etapa de postmultiplicación compleja de 128 puntos. La salida de este proceso es entonces mezclada en los buffers de mezclado apropiados.

Note que el mezclado podría hacerse completamente en el dominio de la frecuencia, antes de la transformada inversa parcial, si no fuera por el uso de switcheo en los bloques. En el caso de transformadas de bloques switcheados, se usa un proceso diferente de transformada inversa parcial, y los resultados de las dos transformadas diferentes no pueden combinarse directamente hasta justo antes de la etapa de ventana.

Como se mostrará posteriormente, la etapa de procesamiento a la entrada utiliza relativamente poca cantidad de memoria, ya que puede procesar la trama de entrada, un canal a la vez, e incluso una banda a la vez. Por otra parte, este es la etapa más intensiva del

proceso de decodificación por el número de MIPS (millones de instrucciones por segundo) que requiere, ya que la asignación de bits y las rutinas de desempaqueado tienen por sí mismas la mitad de complejidad del total de la decodificación.

❖ Procesamiento a la salida del decodificador.

Una vez que el procesamiento a la entrada se ha completado y los buffers de mezclado se han llenado, el codificador puede llevar a cabo el procesamiento a la salida (véase figura 7.48). Para cada canal de salida, un buffer de mezclado y su correspondiente buffer de retraso con una longitud de 128 para medio bloque son puestos en una ventana y combinados para producir 256 muestras PCM de salida. Estas muestras son entonces redondeadas según la longitud de la palabra del DAC y copiadas en el buffer de salida. Una vez que se ha hecho esto, la mitad del buffer de mezclado es copiada a su buffer de retraso correspondiente, proporcionando el 50% de información de traslape necesaria para la reconstrucción apropiada del siguiente bloque de audio.

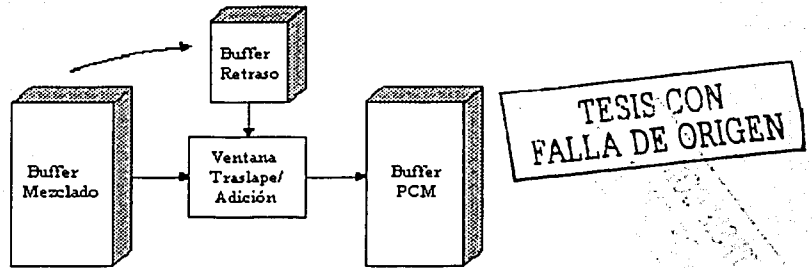


Figura 7.48 Procesamiento a la salida

A diferencia de la etapa de procesamiento a la entrada, el procesamiento a la salida en el decodificador utiliza muy pocas MIPS y una cantidad relativamente grande de memoria. Para cada canal de salida, el decodificador requiere un buffer de mezclado con una longitud de 256, un buffer de retraso con una longitud de 128, y un buffer de salida con una longitud de 512 (asumiendo una salida con dos buffers). Mas adelante estos buffers requieren espacios lo suficientemente grandes para mantener la relación señal a ruido de la banda ancha dentro de los estándares de audio profesionales. Para una salida de precisión de 16 bits PCM, el buffer de retraso debería tener al menos un tamaño de 18 bits, y los buffers de mezclado y MDCT deberían tener un tamaño de 20 bits. Para una salida PCM más exacta, estos números deben incrementarse adecuadamente.

7.7.3 Resumen de la memoria del decodificador.

Los requerimientos de memoria total y MIPS del decodificador AC-3 se resumen en las Tablas 7.6 y 7.7. La estructura de la memoria se divide en RAM Y ROM (Read Only Memory). En estas tablas, el parámetro N es el número de canales de salida. Estas tablas también asumen que la trama de entrada pasa por un solo buffer, mientras que los buffers de la salida son dos, y que la trama de bits de AC-3 es un programa de seis canales

codificados a 384 kbps con una frecuencia de muestreo de 48 khz (la configuración estándar de HDTV).

Un decodificador de dos canales requiere significativamente menos RAM que un decodificador de seis canales. En cuanto a MIPS, podemos decir que la complejidad computacional es igual para un decodificador de dos canales y de seis canales. Esto indica que un diseño de un circuito integrado para un decodificador de dos canales puede ampliarse hacia un dispositivo de seis canales, sin un cambio significativo del procesador.

El tamaño del MIPS y del código de programa ROM se determinaron usando la serie de instrucciones del procesador DSP de Zoran ZR38001. Este circuito integrado, es un procesador de propósito general, que proporciona un solo ciclo de desplazamiento y cuatro ciclos de FFT de mariposa^[6], además de las características usuales de un DSP. Como ejemplo de comparación, el DSP56002 de Motorola requiere 45 MIPS para realizar la decodificación AC-3 de seis canales.

Estructura de Datos RAM	Palabras	# de Bits
Frame AC-3 Entrada	768	16
Datos fijos bloque de audio	384	16
Buffer MDCT	256	≥ 20
Buffers Mezclado	N+256	≥ 20
Buffers Retardo	N+128	≥ 18
Buffers PCM	N+512	≥ 16

Tabla 7.6 Requerimientos de memoria RAM

Estructura de Datos ROM	Palabras	# de Bits
Tablas de Transformadas	768	≥ 20
Otras tablas	1250	16
Código de Programa	3500	32

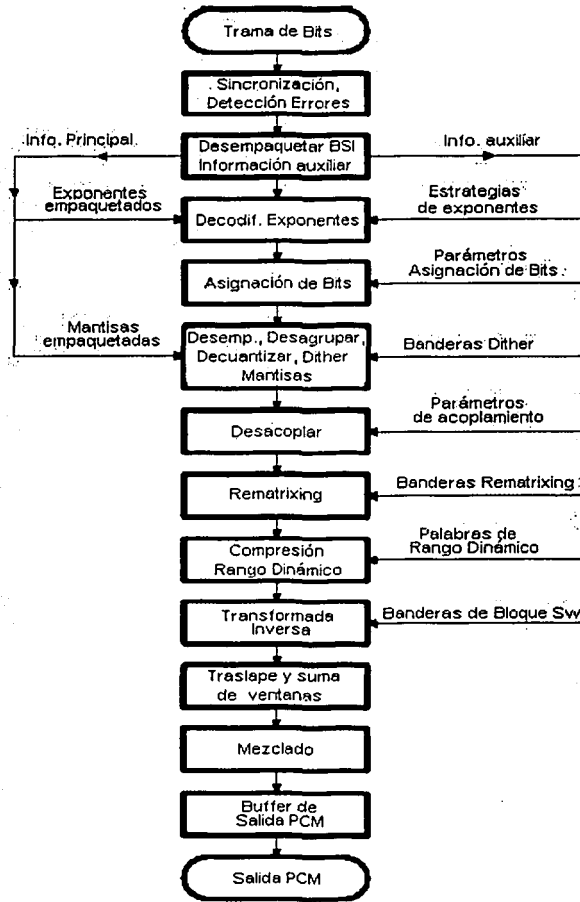
Tabla 7.7 Requerimientos de memoria ROM

7.7.4 Decodificando la Trama de bits AC-3

El proceso de decodificación de la trama de audio AC-3 se muestra en la figura 7.49. A continuación se explica cada uno de los bloques indicados en dicho diagrama.

❖ Trama de bits de entrada

La trama de bits AC-3 vendrá típicamente de un sistema de transmisión o de almacenamiento. Los datos AC-3 codificados pueden entrar en el decodificador como una trama de datos continua con una tasa de transmisión nominal



TESIS CON FALLA DE ORIGEN

Figura 7.49 Diagrama de bloques del proceso de decodificación de la trama de bits de AC-3.

Alineación de bytes o palabras: La mayoría de las aplicaciones transportarán la trama de bits elemental AC-3 con alineación de bytes o palabras (16 bits). El frame *sync* siempre tiene un número de palabras que lo integran, hablando de su longitud. El decodificador puede recibir datos como una trama de bits serial continua sin alineación alguna, o con alineación de bytes o palabras de 16 bits. La alineación de datos de entrada permite que exista simplificación en el decodificador. La alineación ayuda a reducir la probabilidad de una detección errónea de la palabra *sync*.

❖ Sincronización y detección de errores

El formato de la trama de bits AC-3 permite una sincronización rápida. La palabra de sincronía es siempre 0x0B77, es decir "0000 1011 0111 0111". Esta palabra de 16 bits tiene una baja probabilidad de una detección falsa. Sin una alineación de la trama de entrada la probabilidad de detección equivocada de la palabra de sincronía es de 0.0015 % por posición de bits de la trama de entrada. Para una tasa de transmisión de 384 kbps, la probabilidad de una detección errónea de la palabra de sincronía es de 19% por frame. La alineación de bytes de la entrada disminuye esta probabilidad hasta 2.5%, y la alineación de palabras la disminuye hasta 1.2%.

Los datos de la trama AC-3 pueden arrojar aquellos errores que existan dentro de el frame de datos. El decodificador puede ser informado sobre este hecho mediante el sistema de transporte que ha liberado los datos. La integridad de los datos puede ser verificada usando los CRCs. También, algunas verificaciones de consistencia simple en los datos recibidos pueden indicar que hay errores presentes. La estrategia a seguir por el decodificador cuando hay errores presentes la define el usuario. Las respuestas posibles incluyen silencio, repetición de bloques o repetición de frames.

Cada frame AC-3 contiene dos palabras CRC de 16 bits. La segunda palabra de 16 bits en el frame AC-3 es *crc1*, siguiendo la palabra *sync*. La última palabra de 16 bits en el frame de audio es *crc2*, precediendo la palabra *sync* del frame siguiente. *crc1* se aplica a los primeros 5/8 del frame, sin incluir la palabra de sincronía; mientras *crc2* cubre los últimos 3/8 del frame así como el frame completo (sin incluir la palabra *sync*). La decodificación de la(s) palabra(s) CRC permiten que los errores sean detectados.

El generador polinomial que se usa para generar cada una de las palabras CRC de 16 bits es:

$$x^{16} + x^{15} + x^2 + 1$$

Las restricciones de tamaño del bloque dentro de cada frame (forzada por los codificadores), garantizan que los bloques 0 y 1 estén completamente cubiertas por *crc1*. Entonces, los decodificadores pueden empezar a procesar inmediatamente el bloque 0 cuando los 5/8 del frame de datos se alcancen. Esto permite que los buffers de entrada sean más pequeños en algunas aplicaciones. Los decodificadores que almacenan un frame completo pueden seleccionar utilizar solo *crc2*, por lo cual los decodificadores no comenzarían a procesar el bloque 0 de un frame hasta que todo este sea recibido.

Cuando un patrón de *sync* es detectado, el decodificador puede estimar si está en sincronía y las dos palabras CRC pueden ser verificadas. Debido a que *crc1* viene primero y cubre los primeros 5/8 del frame, el resultado de una verificación de *crc1* puede estar disponible después de que hayan sido recibidos los 5/8 del frame. O se puede recibir todo el frame y verificar *crc2*. Si las dos CRC no indican errores, el decodificador puede presumir que seguramente está en sincronía, y puede proceder la decodificación y reproducción de audio. La posibilidad de una sincronía falsa en este caso será la concatenación de las probabilidades de una detección de palabra *sync* errónea y una falla en la detección de error CRC. La prueba CRC es confiable hasta 0.0015%. Esta probabilidad, concatenada con la probabilidad de una detección errónea de sincronía en una trama de bits de entrada alineada por bytes, arroja una probabilidad de sincronización equivocada de 0.000035% (o cerca de una en tres millones de intentos de sincronización).

Existe siempre la posibilidad de que un frame AC-3 pueda tener información *sync* válida y CRCs válidos, y por otra parte no se pueda decodificar. Esta condición puede aparecer si un frame es adulterado de tal forma que la palabra CRC es todavía válida, o en el caso de error en el codificador (bug). Una protección contra este problema es el efectuar pruebas de verificación de errores dentro del decodificador AC-3. A pesar de su eficiencia de codificación, hay algunas redundancias inherentes en la trama de bits AC-3. Si la trama de bits AC-3 contiene errores, pueden aparecer un número de construcciones sintácticas ilegales. Al realizar la búsqueda de estas construcciones ilegales se detectarán un gran número de condiciones significativas de error.

❖ **Desempaquetado de BSI, Información auxiliar**

Inherente al proceso de decodificación, se encuentra el desempaquetado (demultiplexaje) de los diversos tipos de información incluidos en la trama de bits. Algunos de estos pueden ser copiados desde el buffer de entrada a los registros dedicados, otros pueden ser copiados a una localidad de memoria de trabajo específica, e incluso pueden ser localizados en el buffer de entrada con apuntadores que los manden a otra localidad para su uso cuando se requiera la información.

❖ **Decodificación de exponentes**

Los exponentes son liberados en la trama de bits de forma codificado. Para desempaquetar y decodificar los exponentes se requieren dos tipos de información auxiliar.

Primero, el número de exponentes debe conocerse: Para los canales con ancho de banda completo este puede determinarse para los canales no acoplados y los acoplados. Para el canal de acoplamiento, el número de exponentes puede determinarse. Para el canal LFE (cuando esté en uso), hay siempre 7 exponentes.

Segundo, se debe conocer la estrategia de exponentes que usa cada canal: Los exponentes diferenciales se combinan en grupos en el bloque de audio. La agrupación se hace por uno de los tres métodos, D15, D25 o D45, los cuales se llaman estrategias de los exponentes. La estrategia de exponentes para cada canal acoplado o independiente se

incluye en un grupo campos de 2 bits designados como `chexpstr[ch]`, como se observa en la Tabla 7.8.

<code>chexpstr[ch]</code> , <code>chexpstr</code>	Estrategia de Exponentes	Exponentes por Grupo
00	reutilización exponentes previos	0
01	D15	3
10	D25	6
11	D45	12

Tabla 7.8 Codificación de Estrategia de Exponentes

Cuando el canal LFE está habilitado, el campo `lfeexpstr` está presente. Se decodifica como se muestra en la Tabla 7.9.

<code>lfeexpstr</code>	Estrategia de Exponentes	Exponentes por Grupo
0	reutilización exponentes previos	0
1	D15	3

Tabla 7.9 Codificación de Estrategia de Exponentes de Canal LFE

La decodificación de un conjunto de exponentes agrupados codificados creará conjuntos de exponentes absolutos de 5 bits. Los exponentes se decodifican como sigue:

- Cada uno de los valores mapeados de agrupamiento de 7 bits es decodificado usando el procedimiento inverso a la codificación.
- Cada valor mapeado es convertido en un exponente diferencial restándole el offset de mapeo.
- El grupo de exponentes diferenciales se convierte en exponentes absolutos añadiéndole a cada exponente diferencial el exponente absoluto del bin de frecuencia previo.
- Para los modos D25 y D45, cada exponente absoluto es copiado a los miembros que faltan del par o del cuarteto de exponentes.

❖ Asignación de bits

El cálculo de la asignación de bits revela cuántos bits son usados por cada mantisa. Las entradas del cálculo de la asignación de bits son los exponentes decodificados, y la información auxiliar de la asignación de bits. Las salidas del cálculo de la asignación son un grupo de apuntadores de asignación de bits (`baps`), uno por cada mantisa codificada. El `bap` indica el cuantizador usado por la mantisa, y cuántos bits en la trama de bits fueron usados por cada mantisa.

La asignación de bits debe ser calculada en el decodificador siempre que la estrategia de exponentes para uno o más canales no indique reutilización. De acuerdo con esto, la asignación de bits puede ser actualizada a una razón que va desde una vez por bloque de

TESIS CON
TALLA DE ORIGEN

audio hasta una vez cada 6 bloques de audio. Un grupo completo de información de asignación de bits nueva se transmite siempre en bloque de audio 0. Debido a que la rutina paramétrica de asignación de bits debe generar resultados idénticos en todas las implementaciones de codificadores y decodificadores, cada paso se define en términos de operaciones con enteros de punto fijo y búsquedas en tablas. Se emplea aritmética complemento a dos signada. Todas las adiciones se realizan con un acumulador de 14 o más bits. Todos los resultados intermedios y valores almacenados tienen 8 bits.

❖ Procesamiento de mantisas

Las mantisas cuantizadas toscamente componen la mayor parte de datos de AC-3. Cada mantisa es cuantizada al nivel de precisión indicado por su *bap* correspondiente. Para empaquetar los datos de mantisas más eficientemente, algunas mantisas son agrupadas en un solo valor transmitido. Por ejemplo, dos valores cuantizados de 11 niveles son transportados en un código de 7 bits (3.5 bits/valor) en la trama de bits.

Los datos de la mantisa son desempaquetados quitando grupos de bits de la forma indicada por los *baps*. Las mantisas agrupadas deben ser desagrupadas. Los valores individuales de la mantisa codificada son convertidos en valores decuantizados. Las mantisas que tienen cero bits pueden ser reproducidos como cero, o como un valor de dither aleatorio (bajo el control de la bandera dither).

Durante el proceso de decodificación, la trama de datos de mantisas se separa en mantisas individuales de longitud variable, que se intercalan con grupos que representan la codificación combinada tanto de los pares como de los tríos de mantisas. En la trama de bits, las mantisas en cada conjunto de exponentes están colocadas en orden ascendente de frecuencia. Sin embargo, cuando hay mantisas agrupadas se desempaca la primera mantisa contenida en el grupo. Nada es desempacado de la trama de bits para las subsecuentes mantisas en el grupo.

Para los valores del apuntador del arreglo de asignación de bits de $6 \leq \text{bap} \leq 15$ ($32 \leq$ niveles del cuantizador $\leq 65,536$), se usa cuantización fraccional asimétrica de complemento a dos. Cada mantisa, junto con su exponente, son la representaciones en punto flotante del coeficiente de la transformada. El punto decimal se encuentra a la izquierda del bit más significativo¹⁹, por lo tanto la palabra de la mantisa está representada en el rango de $(1.0 - 2^{-(\text{qntztab}[\text{bap}]-1)})$ a -1.0 . El número de la mantisa *k* se extrae de la trama de bits, de longitud $\text{qntztab}[\text{bap}[k]]$. Se convierte de nuevo a la representación de punto fijo mediante un desplazamiento de la mantisa hacia la derecha según su exponente.

Para los *baps* entre 1 y 5 ($3 \leq$ niveles de cuantizador ≤ 15), las mantisas están representadas por valores codificados. Los valores codificados se convierten a palabras binarias fraccionales de complemento a 2 mediante una consulta de una tabla. El número de bits indicado por el *bap* de la mantisa se extrae de la trama de bits y se justifica a la derecha. Este valor codificado se usa como un índice para buscar el valor de la mantisa en la tabla. El valor de la mantisa resultante se desplaza a la derecha por su exponente correspondiente para generar el valor del coeficiente de la transformada.

El decodificador utiliza valores de ruido aleatorio (dither) en lugar de valores cuantizados cuando el número de bits asignados a la mantisa es cero ($bap = 0$). El uso del ruido aleatorio está condicionado al valor de dithflag. Cuando el valor de dithflag es 1, se usa un valor de ruido aleatorio. El dither es aplicado después de que los canales individuales son extraídos del canal de acoplamiento.

Cualquier secuencia aleatoria razonable puede ser utilizada para generar valores dither. La longitud de palabra de los valores dither no es crítica. Con ocho bits es suficiente. La escala óptima de la palabra de dither es el tomar una distribución uniforme con valores entre -1 y $+1$, y escalarlos por 0.707 , resultando en una distribución uniforme entre $+0.707$ y -0.707 .

Una vez que el valor dither es asignado a la mantisa, la mantisa se desplaza a la derecha de acuerdo con su exponente para generar el coeficiente de la transformada correspondiente.

❖ Desacoplamiento

Cuando el acoplamiento está en uso, los canales que están acoplados deben ser desacoplados. El desacoplamiento envuelve la reconstrucción de la sección de alta frecuencia (exponentes y mantisas) para un canal acoplado, a partir del canal de acoplamiento y las coordenadas de acoplamiento para un canal individual. Dentro de cada banda acoplada, los coeficientes de acoplamiento de canal (exponente y mantisa) se multiplican por las coordenadas de acoplamiento para ese canal y esa sub-banda de frecuencia.

Las coordenadas de acoplamiento existen para cada banda de acoplamiento en cada canal que es acoplado. Las coordenadas de acoplamiento se envían en formato de punto flotante. El exponente se envía como un valor de 4 bits, que indica el número de desplazamientos hacia la derecha que deben ser aplicados al valor de la mantisa fraccional. Las mantisas se transmiten como valores de 4 bits, que deben ser propiamente escalados antes de su uso. Las mantisas son valores no signados. Las mantisas de canales individuales se reconstruyen a partir de los canales acoplados.

❖ Rematrixing^[9]

En el modo de codificación de audio 2/0 puede ser empleado el “rematrixing”, si así lo indican sus banderas. Cuando la bandera indica que una banda está bajo este tipo de codificación, los coeficientes codificados en la trama de bits son los valores sumados y restados, en lugar de los valores izquierdo y derecho.

El “rematrixing” en AC-3 es un técnica para combinar canales en los que las sumas y diferencias de los canales altamente correlacionados se codifican en lugar de los canales originales. En lugar de codificar y empaquetar los canales izquierdo y derecho en un codificador de dos canales, se construye

$$\begin{aligned}\text{izquierdo}' &= 0.5 * (\text{izquierdo} + \text{derecho}); \\ \text{derecho}' &= 0.5 * (\text{izquierdo} - \text{derecho});\end{aligned}$$

Las operaciones de cuantización y paquetización se realizan en izquierdo' y derecho'. Si la señal estéreo original fuera idéntica en ambos canales, esta técnica nos daría como resultado que la señal izquierdo' es idéntica a los canales izquierdo y derecho originales, y que la señal derecho' es igual a cero. Como resultado, se puede codificar el canal derecho' con muy pocos bits, y se incrementa la exactitud en el canal izquierdo', que es más importante. Esta técnica es especialmente importante para preservar la compatibilidad con Dolby Surround.

❖ Compresión de Rango Dinámico^[9]

Para cada bloque de audio puede incluirse un valor de control de rango dinámico en la trama de bits. El decodificador, por defecto, usará este valor para alterar la magnitud de los coeficientes (exponente y mantisa).

Un problema común en la liberación de la programación de audio es que los diferentes miembros de la audiencia desean disfrutar de diferentes rangos dinámicos. La programación de alta calidad (como la de las películas) se encuentran mezcladas típicamente en un rango dinámico un poco ancho. Usando los diálogos como referencia, los sonidos fuertes como las explosiones son unos 20 dB más grandes o más, y los sonidos débiles como el crujido de las hojas son 50 dB menores. En muchas condiciones de audición es inaceptable permitir que el sonido llegue a ser muy fuerte, y por lo tanto los sonidos más fuertes deben ser comprimidos hacia un nivel más bajo. De igual manera, en muchas situaciones los sonidos muy bajos podrían ser inaudibles, y deben ser puestos en un nivel más alto para que sean escuchados. Como la mayor parte de la audiencia se beneficiará de un rango dinámico de programa limitado, las pistas sonoras que han sido mezcladas con un rango dinámico amplio son generalmente comprimidas: el rango dinámico se reduce mediante bajar el nivel de los sonidos altos y subir el nivel de los sonidos bajos. Aunque esto satisface las necesidades de la mayor parte de la audiencia, quita la posibilidad para que una parte de la misma escuche el programa de audio completo original. La tecnología de codificación AC-3 resuelve este conflicto al permitir que los valores del control del rango dinámico sean colocados en la trama de bits de AC-3.

Los valores del control del rango dinámico, *dynrng*, indican un cambio en la ganancia que será aplicada en el decodificador para implementar una compresión del rango dinámico. Cada uno de estos valores puede indicar un cambio de ganancia de ± 24 dB. La secuencia de estos valores son una señal de control de compresión. Un codificador AC-3 generará esta secuencia de valores. Cada valor es utilizado por el decodificador AC-3 para alterar la ganancia de uno o más de los bloques de audio. Para el escucha, es deseable que los sonidos más fuertes se disminuyan hacia el nivel del diálogo, y que los sonidos bajos se incrementen en su nivel, hacia el nivel de diálogo. Los sonidos con la misma intensidad que el diálogo hablado normal, típicamente no cambiarán su ganancia.

La compresión de rango dinámico se aplica en el decodificador de audio AC-3. El audio codificado tiene un rango dinámico completo. Se permite que el decodificador ignore los valores de compresión de rango dinámico en la trama de bits. También se permite que el decodificador utilice una fracción del valor de control.

❖ Transformada Inversa, Ventana y Traslape/Suma

Los pasos de decodificación descritos antes resultarán en un grupo de coeficientes de frecuencia para cada canal codificado. La transformada inversa convierte los bloques de coeficientes de frecuencia en bloques de muestras de tiempo.

A los bloques individuales de muestras temporales se les aplica una función ventana, y los bloques adyacentes deben ser traslapados y sumados juntos para reconstruir la salida de la señal PCM de audio final.

Transformada IMDCT (Transformada Inversa del Coseno Discreto Modificado) de 512 muestras

A continuación se presenta el algoritmo para hacer la IMDCT^[9] para un bloque de $N=512$ datos reales usando una IFFT^[4,5,6] compleja de $N/4$ puntos, cuando existe ausencia de transitorios.

- Se definen los coeficientes de la transformada MDCT $\rightarrow X[k], k=0,1,\dots,N/2-1$
 - Se realiza un paso de pre multiplicación compleja IFFT
- Se calcula el producto de multiplicación $Z[k]$ de $N/4$ puntos, donde $k=0,1,\dots,N/4-1$:

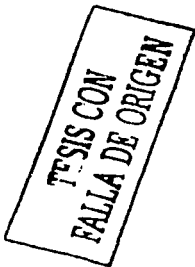
```
for k=0,1,...,(N/4)-1
{
     $Z[k] = (X[N/2 - 2 * k - 1] + j * X[2 * k]) * (x \cos 1[k] + j * x \text{sen}1[k])$ 
}
```

donde $x \cos 1[k] = -\cos(2\pi * (8k+1)/(8N))$, y
 $x \text{sen}1[k] = -\text{sen}(2\pi * (8k+1)/(8N))$.

- Se calcula la IFFT

Se calcula una IFFT compleja de $N/4$ puntos de $Z(k)$ para generar la secuencia de valores complejos $z(n)$.

```
for n=0,1,...,(N/4)-1
{
     $z[n]=0$ ;
    for k=0,1,...,(N/4)-1
    {
         $z[n] += Z[k] * (\cos(8 * \pi * k * n / N) + j * \text{sen}(8 * \pi * k * n / N))$ 
    }
}
```



➤ Paso de post multiplicación compleja IFFT

Se calcula el producto complejo de $N/4$ puntos $y(n)$, $n=0,1,\dots,N/4-1$

```
for n=0,1,...,(N/4)-1
{
   $y[n] = z[n] * (x \cos l[n] + j * x \text{sen} l[n])$ 
}
```

➤ Ventana y desintercalación

Se calculan las muestras en el dominio del tiempo ventaneadas $x[n]$. La coeficientes de la función ventana salen de una tabla⁹⁾.

```
for n=0,1,...(N/8)-1
{
   $x[2n] = -yi[(N/8)+n]*w[2n];$ 
   $x[2n+1] = yr[(N/8)-n-1]*w[2n+1];$ 
   $x[(N/4)+2n] = -yr[n]*w[(N/4)+2n];$ 
   $x[(N/4)+2n+1] = yi[(N/4)-n-1]*w[(N/4)+2n+1];$ 
   $x[(N/2)+2n] = -yr[(N/8)+n]*w[(N/2)-2n-1];$ 
   $x[(N/2)+2n+1] = yi[(N/8)-n-1]*w[(N/2)-2n-2];$ 
   $x[3(N/4)+2n] = yi[n]*w[(N/4)-2n-1];$ 
   $x[3(N/4)+2n+1] = -yr[(N/4)-n-1]*w[(N/4)-2n-2];$ 
}
```

en donde $yr[n]$ = parte real de $y[n]$
 $yi[n]$ = parte imaginaria de $y[n]$
 $w[n]$ es la secuencia de la transformada de ventana, cuyo valor se encuentra en una tabla del Estándar A/52.

➤ Traslape y suma

La primera mitad del bloque ventaneado se traslapa con la segunda mitad del bloque previo para producir las muestras PCM (El factor de escala '2' deshace el headroom^{10,11,12} producido por el codificador).

```
for(n=0;n<N/2;n++)
{
   $\text{pcm}[n] = 2 * (x[n] + \text{delay}[n]);$ 
   $\text{delay}[n] = x[(N/2)+n];$ 
}
```

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

Transformada IMDCT de 256 muestras

Las ecuaciones siguientes se deben utilizar para calcular la transformada inversa IMDCT^[9] cuando hay presencia de un transitorio y se hacen dos transformadas de 256 muestras. (N sigue siendo igual a 512).

- Se definen los coeficientes de la transformada MDCT $\rightarrow X[k]$, $k=0,1,\dots,N/2$.

```
for(k=0;k<N/4;k++)
{
    X1[k]=X[2*k];
    X2[k]=X[2*k+1];
}
```

- Pre-multiplicación compleja IFFT

Se calculan los productos complejos de N/8 puntos Z1(k) y Z2(k), $k=0,1,\dots,N/8-1$.

```
for k=0,1,...,(N/8) -1
{
    Z1[k]=(X1[(N/4) -2k-1] + j X1[2k])*(xcos2[k] + j sen2[k]);
    Z2[k]=(X2[(N/4) -2k-1] + j X2[2k])*(xcos2[k] + j sen2[k]);
}
```

donde
 $xcos2[k] = -\cos(2\pi*(8k+1)/4N)$, y
 $xsen2[k] = -\sin(2\pi*(8k+1)/4N)$.

- Se calcula la IFFT.

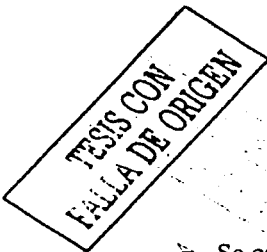
Se calculan la IFFT compleja de N/8 puntos de Z1[k] y Z2[k] para generar Z1[n] y Z2[n] para generar secuencias de valores complejos z1[n] y z2[n].

```
for n=0,1,...,(N/8) -1
{
    z1[n]=0;
    z2[n]=0;
    for k=0,1,...,(N/8)-1
    {
        z1[n] += Z1[k]*(cos(16\pi*k*n/N) + j sen(16\pi*k*n/N));
        z2[n] += Z2[k]*(cos(16\pi*k*n/N) + j sen(16\pi*k*n/N));
    }
}
```

- Se calcula la post multiplicación compleja IFFT.

Se calculan los productos complejos de N/8 puntos y1[n] y y2[n], $n=0,1,\dots,N/8-1$

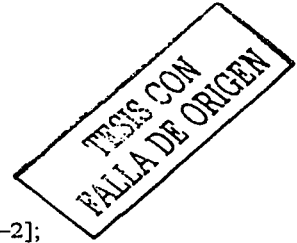
```
for n=0,1,...,(N/8) -1
{
    y1[n]= z1[n]*(xcos2[n] + j xsen2[n]);
}
```



$$y2[n] = z2[n] * (x\cos2[n] + j \text{ xsen2}[n]);$$

➤ Multiplicación por la ventana y desintercalación
Se calculan las muestras en el dominio del tiempo $x[n]$.

```
for n=0,1,...,(N/8) -1
{
  x[2n] = -yi1[n] * w[2n];
  x[2n+1] = yr1[(N/8)-n-1]*w[2n+1];
  x[(N/4)+2n] = -yr1[n]*w[(N/4)+2n];
  x[(N/4)+2n+1] = yi1[(N/8) -n-1]*w[(N/4)+2n+1];
  x[(N/2)+2n] = -yr2[n]*w[(N/2) -2n-1];
  x[(N/2)+2n+1] = yi2[(N/8)-n-1]*w[(N/2)-2n-2];
  x[3(N/4)+2n] = yi2[n]*w[(N/4) -2n-1];
  x[3(N/4)+2n+1] = -yr2[(N/8) -n-1]*w[(N/4) -2n-2];
}
```



donde

$yr1[n]$ = parte real $y1[n]$,

$yi1[n]$ = parte imaginaria $y1[n]$,

$yr2[n]$ = parte real $y2[n]$,

$yi2[n]$ = parte imaginaria $y2[n]$, y

$w[n]$ es la secuencia de la ventana de transformación que se encuentra en una tabla del estándar AC-3.

➤ Traslape y suma.

La primera mitad de los bloques que ya fueron ventaneados es trasladada con la segunda mitad del bloque previo para producir las muestras PCM.

```
for n=0,1,...(N/2)-1
{
  pcm[n] = 2*(x[n]+delay[n]);
  delay[n] = x[(N/2)+n];
}
```

❖ Mezclado

Si el número de canales requeridos en la salida del decodificador es más pequeño que el número de canales que están codificados en la trama de bits, se requiere del mezclado (downmixing)^[9]. El mezclado se puede realizar en el dominio del tiempo y en el dominio de la frecuencia, pues la transformada inversa es una operación lineal. En el dominio de la frecuencia el mezclado se realiza antes de la transformación.

El proveedor de programas puede usar los elementos sintácticos `cmixlev` y `smixlev` para afectar el balance relativo de los canales surround y central con respecto a los canales izquierdo y derecho. El mezclado del canal LFE es opcional. Un mezclado ideal del canal LFE se reproduce a un nivel acústico de +10 dB con respecto a los canales izquierdo y derecho. Se debe tener cuidado para asegurar que los altavoces no sean sobrepasados por el contenido completo de bajas frecuencias del canal LFE.

❖ **Buffer de salida PCM**

Los decodificadores típicos proporcionarán muestras PCM a la salida con la frecuencia de muestreo PCM. Debido a que los bloques de muestras resultan del proceso de decodificación, se requiere un buffer de salida .

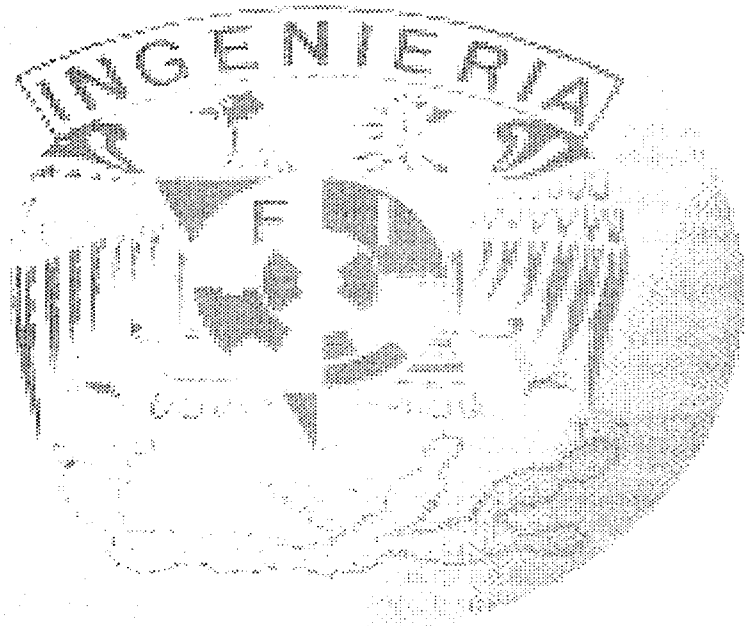
❖ **Salida PCM**

Las muestras PCM de salida pueden ser liberadas en una forma adecuada para su interconexión con un convertidor digital a analógico (DAC) ^[7,8,13].

REFERENCIAS

- [1] Productor de filtro SAW COM DEV. Sitio web www.saw-device.com
- [2] Universidad Carlos III de Madrid, España. Sitio web <http://www.it.uc3m.es>
- [3] Universidad de Granada, España. Sitio web <http://www-etsi2.ugr.es>
- [4] Libro de PORAT, B.; "A course in Digital Signal Processing"; John Wiley & Sons; E.E.U.U., 1997.
- [5] Libro de INGLE, V.; "Digital Signal Processing using Matlab V.4", International Thomson Publishing; E.E.U.U., 1997.
- [6] Libro de PŠENIČKA, B.; "Procesamiento Digital de Señales –La transformada de Fourier y Filtros Digitales"; Facultad de Ingeniería, UNAM; México, 1994.
- [7] Libro de INGLIS, A.; "Video Engineering"; Mc Graw-Hill; E.E.U.U., 1996.
- [8] Libro de TOMASI, W.; "Sistemas de comunicaciones electrónicas"; Pearson Educación; Segunda Edición; México, 2001.
- [9] Documento oficial de la norma A/52 de ATSC "Digital Audio Compression (AC-3), Revision A"; Estándar ATSC A/52A; ATSC; E.E.U.U. 2001.
- [10] Colegio cibernético para producción de televisión y radio. Sitio web <http://www.internetcampus.com/diglevels.htm>
- [11] Centro de investigación de Música y Acústica por computadora. Sitio web http://ccrma-www.stanford.edu/~jos/mdft/How_Many_Bits_are.html
- [12] Soluciones Audio Total. Sitio web <http://www.totalaudio.co.uk/support/FAQdigital.htm>
- [13] Libro de CASTRO, A.; "Teleinformática para Ingenieros en Sistemas de Información"; Editorial Reverté; Segunda Edición; España, 1999.

RECIBIDO
CON
FE DE ORIGEN



TESIS CON
FALLA DE ORIGEN

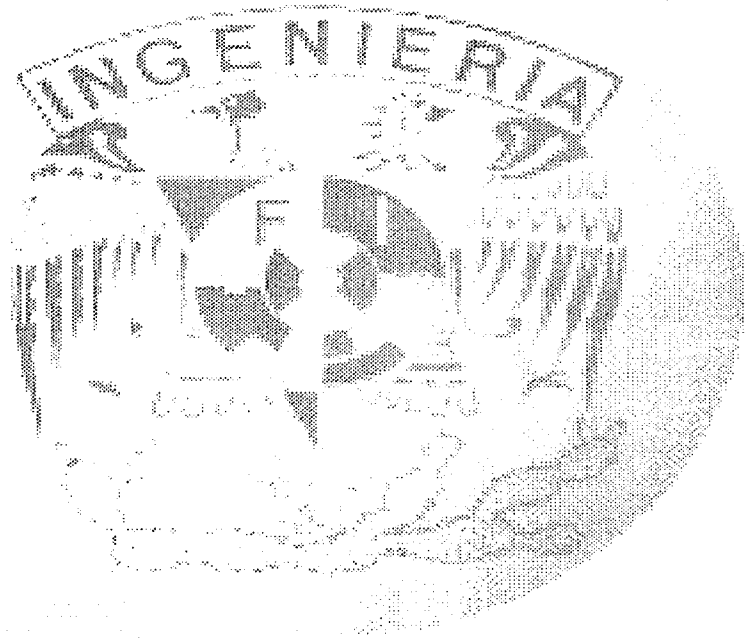


Capítulo # 8

*Diseño de
un Receptor
para DTV*

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

250

TESIS CON FALLA DE ORIGEN

8.1 Planteamiento teórico.

En el capítulo anterior se analizaron las distintas etapas necesarias para hacer la decodificación de la señal de DTV (televisión digital) de acuerdo con el estándar ATSC (Advanced Television Systems Committee). En el presente capítulo se propone el diseño de un aparato decodificador para DTV en el estándar ATSC que cumpla con los procesos necesarios para el tratamiento y decodificación de dicha señal. Nuestro diseño se basa en el diagrama de bloques general de un receptor para ATSC (figura 8.1), para proponer los distintos elementos eléctricos (circuitos integrados, dispositivos, etc) que se utilizarán para la fabricación y/o armado de este aparato receptor.

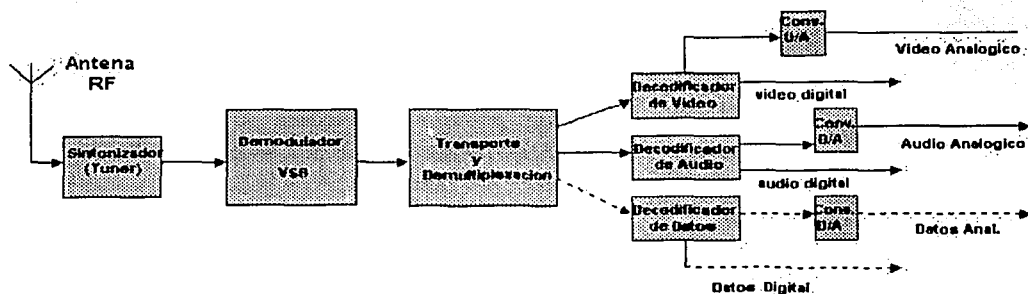


Figura 8.1 Diagrama general de un receptor de ATSC.

TESIS CON FALLA DE ORIGEN

8.2 Componentes necesarios para el diseño.

A continuación se describirán los componentes que elegimos, con los cuales se implementará el receptor de ATSC; tomando en cuenta que existen muchos mas componentes que los mencionados aquí, solo se eligieron los que a nuestra consideración son los mas útiles para nuestro propósito de diseño y cuya información se encuentra disponible para ser consultada en el sitio web de su fabricante, debido a que existen mejores componentes, pero no tuvimos a nuestro alcance ninguna información sobre estos.

8.2.1 Antena.

La antena es un muy importante dispositivo para la recepción de la señal, ya que de esta dependen la alimentación del receptor con la señal de elección del usuario (canal de TV). El sistema ATSC según estudios realizados en Washington, D.C.^[1] es capaz de captar la señal con antenas logoperiódicas de uso externo (ver figura 8.2 a); o con antenas de uso interno (ver figura 8.2 b)) como la antena "V", con muy buena calidad, pero es necesaria su reorientación cada que el usuario cambie de canal (cambio de frecuencia, polarización, etc).

Sugerimos el uso de la antena interna ya que resulta mas cómodo cambiar la orientación de la antena si se encuentra en el interior de la casa. Actualmente ATI ha desarrollado una antena capaz de cambiar su orientación y polarización automáticamente sin necesidad de que el usuario lo haga manualmente, este es un desarrollo de un nuevo tipo de antenas, llamado antena inteligente (ver figura 8.2 c)). ATI ha realizado pruebas con esta antena (llamada DTV-5000)^[2] y obtenido excelentes resultados para los decodificadores de ATSC. Según las pruebas realizadas para el estándar ATSC, este requiere de una figura de ruido de 17 dB por lo cual la antena a elegir deberá contar como mínimo con esta figura de ruido para poder captar la señal de ATSC, las antenas que sugerimos a continuación cumple con esta condición. Cabe mencionar que en este proyecto de TV UNAM se diseñarán también las antenas tanto de transmisión como de recepción para el canal asignado (canal 60).

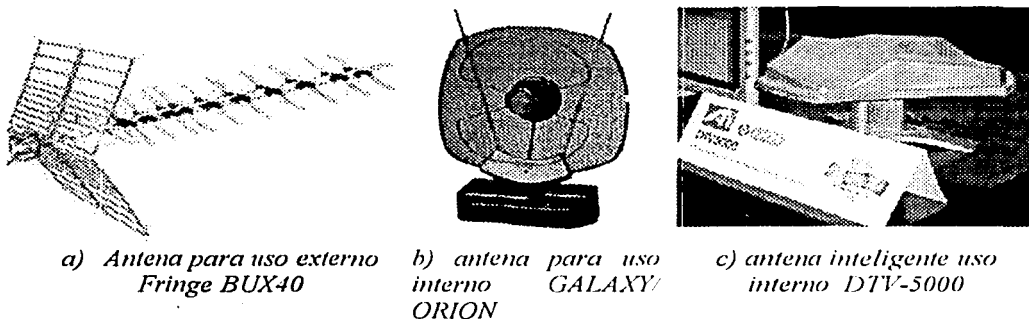


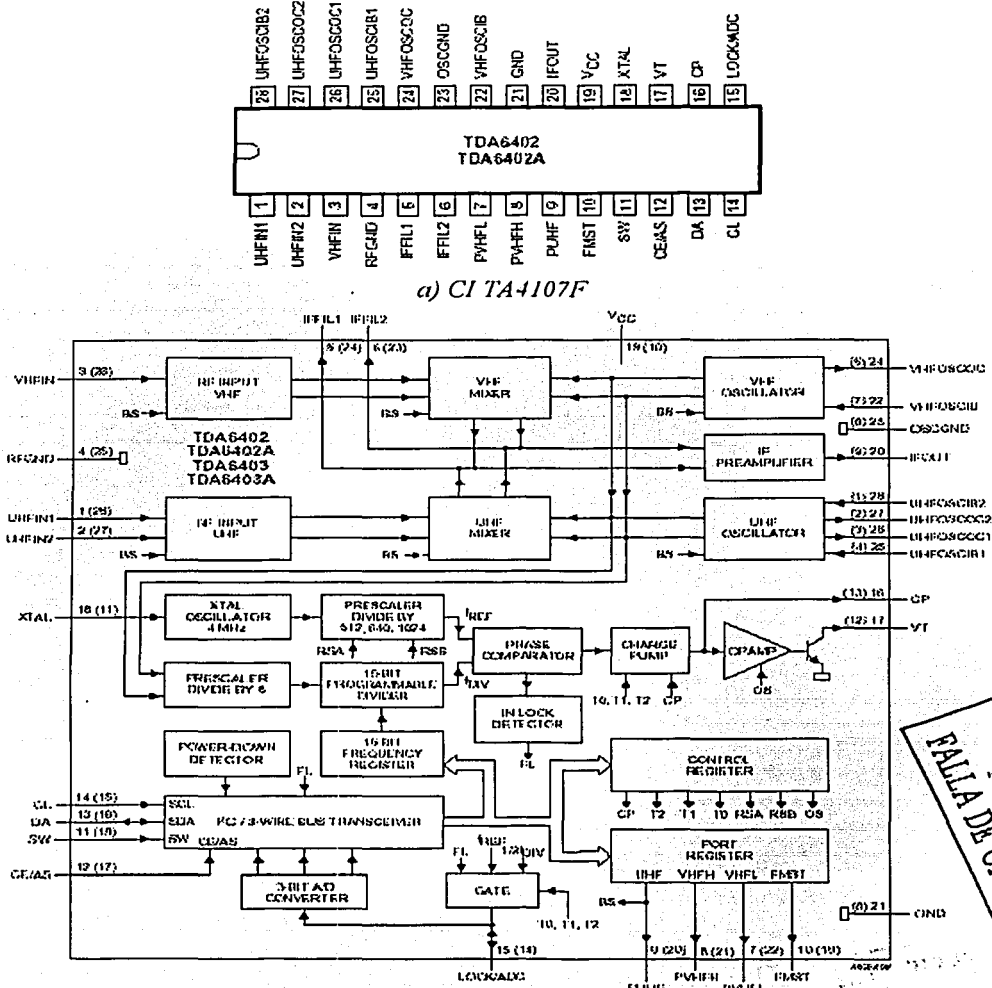
Figura 8.2 Antenas aptas para la captación de señal ATSC.

Es importante mencionar que se puede utilizar un amplificador para la antena como auxiliar en el caso de no contar con una antena que tenga la suficiente sensibilidad para asegurar la recepción de la señal de DTV.

8.2.2 Sintonizador (Tuner).

A continuación se encuentra el dispositivo encargado de discriminar las señales de distintas frecuencias que le entrega la antena y de seleccionar aquella elegida por el usuario. En el mercado existen varios tipos de circuitos integrados que pueden utilizarse con este propósito pero trabajaremos con el CI (circuito integrado) TDA6402 de Philips (Ver figura 8.3) el cual contiene 26 pines para su operación y polarización.

TESIS CON
FALLA DE ORIGEN



b) Diagrama de bloques equivalente

Figura 8.3 CI Tuner de Phillips para DTV.

De acuerdo con las hojas de especificación de operación de este CI^[3] la conexión de este para su adecuada polarización y operación se realizaría con los elementos y voltajes que se muestran en la figura 8.4 a), al armar el circuito se obtendría una tarjeta similar a la mostrada en la figura 8.4 b) la cual entregaría como salida la señal en banda IF (frecuencia

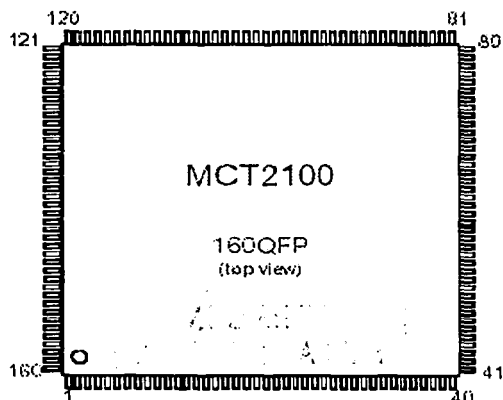
8.2.3 Demodulador VSB.

La siguiente etapa del decodificador de DTV es la demodulación 8-VSB (Banda lateral Residual de 8 niveles), la cual está compuesta por varios procesos que fueron explicados en el capítulo 7. En nuestra propuesta de diseño estos procesos no se armarán por separado, debido a que en el mercado existen CI que realizan todos estos procesos de forma conjunta. Existen varias compañías que fabrican estos CI, por ejemplo:

- NextWave con la serie de CI NXT2002, NXT2003, NXT2004 y NXT2005^[4]
- Philips con el CI TDA8961^[5]
- Motorola con el CI MCT2100^[6]

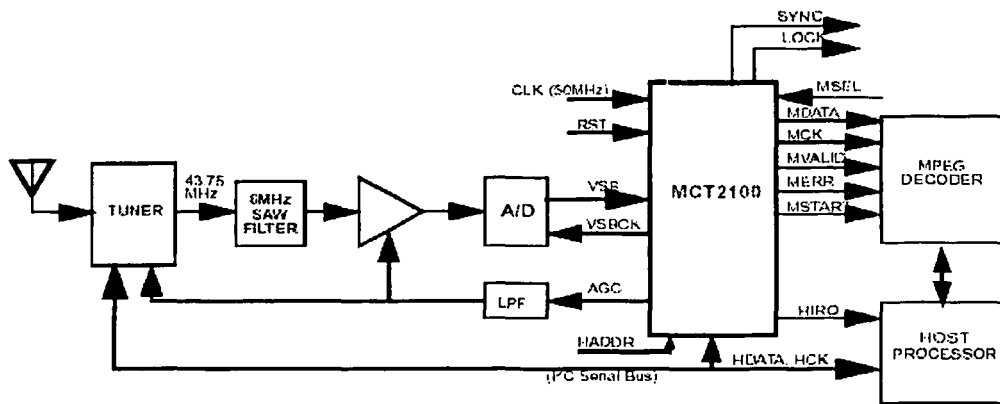
Existen mas distribuidores de estos CI como son Broadcom, Teralogic, etc.. Para nuestro diseño decidimos trabajar con el CI de Motorola (ver figura 8.5 a)). Pero hay que hacer notar que debido a que se trata de una tecnología aún en desarrollo, estos CI pueden ser cambiados o sustituidos por otros mejores o mas actualizados. El CI MCT2100 realiza los procesos de demodulación de 8-VSB como son: demodulación de IF, ecualización, seguidor de fase, Decodificación Trellis, desentrelazado, desaleatorización de datos y la corrección de errores Reed-Solomon FEC (Forward error correction). Al no tener que utilizar un CI para cada etapa del demodulador se simplifica la dificultad de ensamblado y se disminuyen costos, complejidad y tamaño del circuito, por lo que resulta muy útil para el diseño.

Este CI desarrolla la demodulación 8-VSB y la corrección de errores según las especificaciones del documento A/53 del estándar ATSC. En la figura 8.5 b) se muestra la conexión adecuada de este circuito para obtener el TS (transport stream) de bytes requerido.



TESIS CON
 FALLA DE ORIGEN

a) CI de Motorola demodulador ATSC 8-VSB/ FEC para DTV terrestre.



b) Diagrama de conexión del CI MCT2100.

Figura 8.5 CI para la etapa de demodulación 8-VSB.

Como observamos en la figura 8.5 b) el CI requiere de elementos auxiliares o externos como: Un filtro SAW (surface acoustic wave) de 6 MHz, que no es sino un filtro paso banda de muy alta calidad, como el que se muestra en la figura 8.6 a) –distribuido por Oscilent Corporation^[7] –, el cual tiene frecuencia central de 43.75 MHz, así como un ancho de banda de 6 MHz. Se requiere también de un amplificador de ganancia controlada (AGC); existen varios tipos de CI AGC, pero utilizaremos en nuestro caso el MAX3861^[8], el cual se muestra en la figura 8.6 b). Un filtro paso bajas sencillo. Un convertidor A/D de 10 bits que se utiliza para mejorar la decisión de niveles del ecualizador; proponemos usar en el diseño el de Fujitsu MB40C360^[9] (ver figura 8.6 c)). También se necesita proporcionar al CI un reloj externo de 50 MHz lo cual se puede lograr con un cristal de cuarzo como el mostrado en la figura 8.6 d) de la marca Zilog el eZ80190^[10]. También requiere de un procesador maestro (Host) del cual hablaremos más adelante, ya que este procesador es el que se encarga de administrar todos los circuitos que comprende el diseño.

TESIS CON
FALLA DE ORIGEN

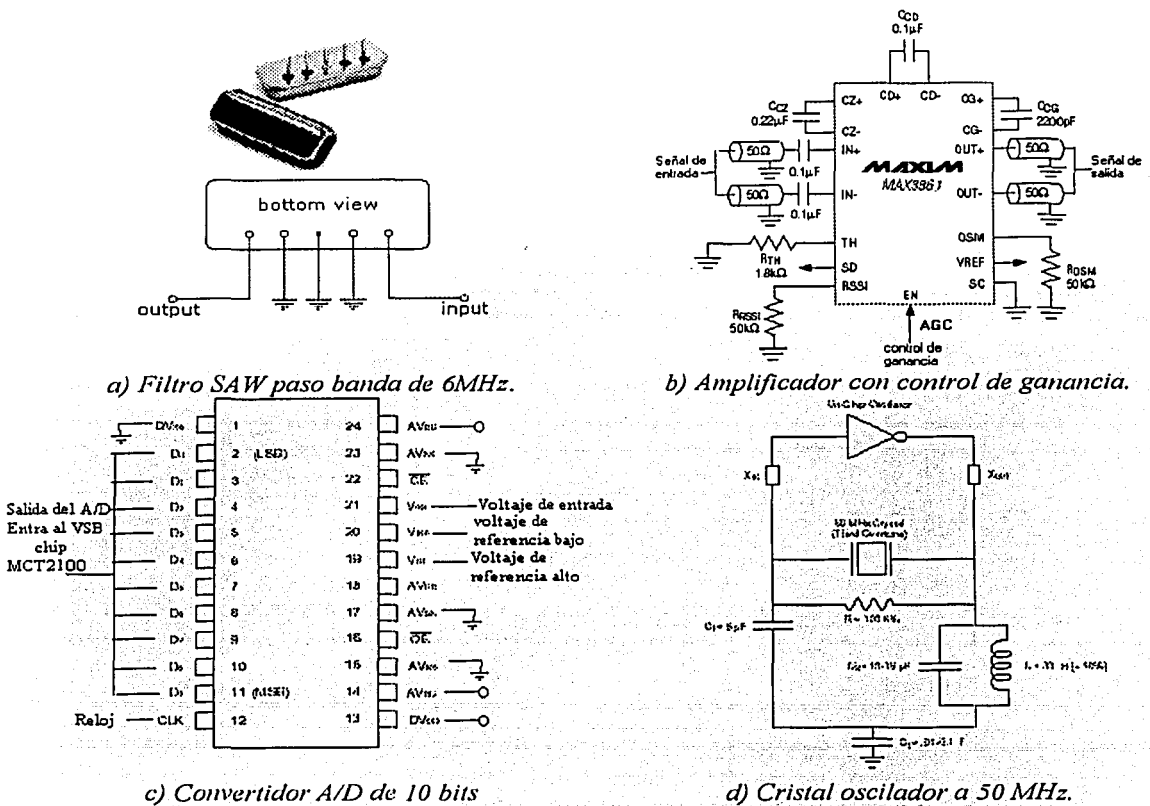
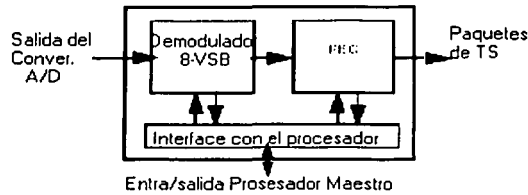


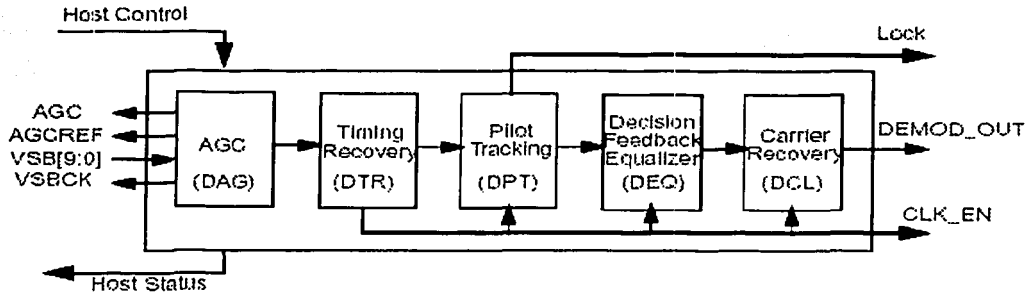
Figura 8.6 Circuitos de apoyo para el CI demodulador VSB/FEC

El diagrama de bloques de este demodulador de Motorola se muestra en la figura 8.7 a), en la cual se observa que se encuentra formado básicamente por tres bloques principales los cuales son: demodulador 8-VSB, el bloque de corrección de errores FEC y la interfase con el procesador maestro. El demodulador a su vez puede desglosarse en sub-bloques los cuales se muestran en la figura 8.7 b), al igual que el FEC cuyos sub-bloques se muestran en la figura 8.7 c). En la tabla 8.1 se hace una descripción de las entradas y salidas más importantes de este CI así como su pin correspondiente.

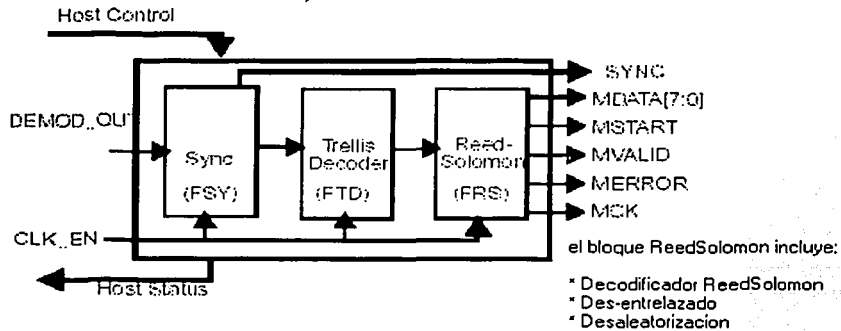
TESIS CON FALLA DE ORIGEN



a) Diagrama de bloques del CI MCT2100.



b) Demodulador 8-VSB



c) Forward Error Correction

Figura 8.7 Bloques de operación del CI MCT2100.

TESIS CON FALLA DE ORIGEN

Diseño de un Receptor para DTV

# de pin	Nombre	Descripción.
1-8	CGND0-CGND3 CVDD0-CVDD3	Son conexiones de potencia, es decir polarización con voltaje. Las GND son a tierra y las VDD son a voltaje de 1.8 volts nominal.
9	MCK	Es una señal de reloj que en modo paralelo es de 6.25MHz y en modo serie es de 50MHz. Ver figura 8.8
10	MVALID	Indica cuando se encuentra disponible un nuevo byte de datos en el bus MDATA. Ver figura 8.8
11	MSTART	Indica cuando un byte de datos corresponde a el primer byte de un paquete de datos TS. Ver figura 8.8
12	MERR	Indica cuando un byte de datos es parte de un paquete de datos con errores. Ver figura 8.8
13-20	CGND4-CGND7 CVDD4-CVDD7	Conexiones de potencia (a voltaje)
21-24	MDATA0-MDATA4	Son la salida de los paquetes de bytes en modo paralelo, el byte entero sale por los 8 pines MDATA0-MDATA7 y en modo serie el byte sale por el pin MDATA0. Ver figura 8.8
25 y 26	PGND0 y PVDD0	Pines de prueba
27-30	MDATA4-MDATA7	Bits de salida de datos (TS)
31-38	CGND8-CVDD11 CVDD8-CVDD11	Pines de Polarización
39 y 40	PGND6 y PGND7	Pines de prueba
41-52	N/C	No se conectan
53-56	CGND12-CGND13 CVDD12-CVDD13	Pines de polarización
57-64	N/C	No se conectan
65 y 66	PGND1 y PVDD1	Pines de prueba
67-72	VSB0-VSB5	Entrada de los paquetes de bits provenientes del convertidor A/D, entran en paralelo por los 10 pines de VSB.
73 y 74	CGND14 y CVDD14	Pines de polarización
75-78	VSB6-VSB9	Entrada de los paquetes de bits
79 y 80	PGND8 y PGND9	Pines de prueba
81-88	CGND15-CGND18 CVDD15-CVDD18	Pines de polarización
89	RST	Pin de reset para el CI
90	CKL	Señal de reloj requerida de entrada de 50MHz (puede provenir de un cristal de cuarzo)

**TESIS CON
FALLA DE ORIGEN**

91	MSEL	Es el pin de entrada para elegir si el CI opera en modo serie o modo paralelo.
92	LOCK	Señal piloto para compensar voltajes de Offset. Es una componente de DC.
93-100	CGND19-CGND22 CVDD19-CVDD22	Pin de polarización
101	SYNC	Salida de señal de sincronía proporcionada por el CI.
102	VSBCK	Señal de reloj de control para el convertidor A/D.
103	HCK	Señal de entrada, reloj del procesador maestro (Host)
104	HIRQ	Salida de datos auxiliares enviada al Host.
105 y 106	PGND2 y PVDD2	Pines de prueba
107-109	HADDR0-HADDR2	Entrada de datos auxiliares de procesadores periféricos.
110	HDATA	Entrada de datos provenientes del Host.
111-118	CGND23-CGND26 CVDD23-CVDD26	Pines de polarización
119 y 120	PGND10 y PGND11	Pines de prueba
121-130	N/C	No se conectan
131-134	CGND27-CGND28 CVDD27-CVDD28	Pines de polarización
135-142	N/C	No se conectan
143 y 144	PGND3 y PVDD3	Pines de prueba
145	N/C	No se conecta
146-150	TDO, TCK, TMS, TRST y TDI	Pines para el control de reloj y su manejo (vigilar, resetear etc.)
151 y 152	CGND29 y CVDD29	Pines de polarización
153 y 154	PGND4 y PVDD4	Pines de prueba
155	AGC	Pin de control para el amplificador con control de ganancia.
156	AGCREF	Voltaje de referencia para el control automático de ganancia.
157 y 158	PGND5 y PVDD5	Pines de prueba
159 y 160	PGND12 y PGND13	Pines de prueba

Tabla 8.1 Descripción de los pines del CI MCT2100.

En la figura 8.8 se muestra la forma de las señales de salida principales de este CI las cuales se encargan del manejo de los paquetes de datos, tanto en modo de operación del CI en serie como en paralelo.



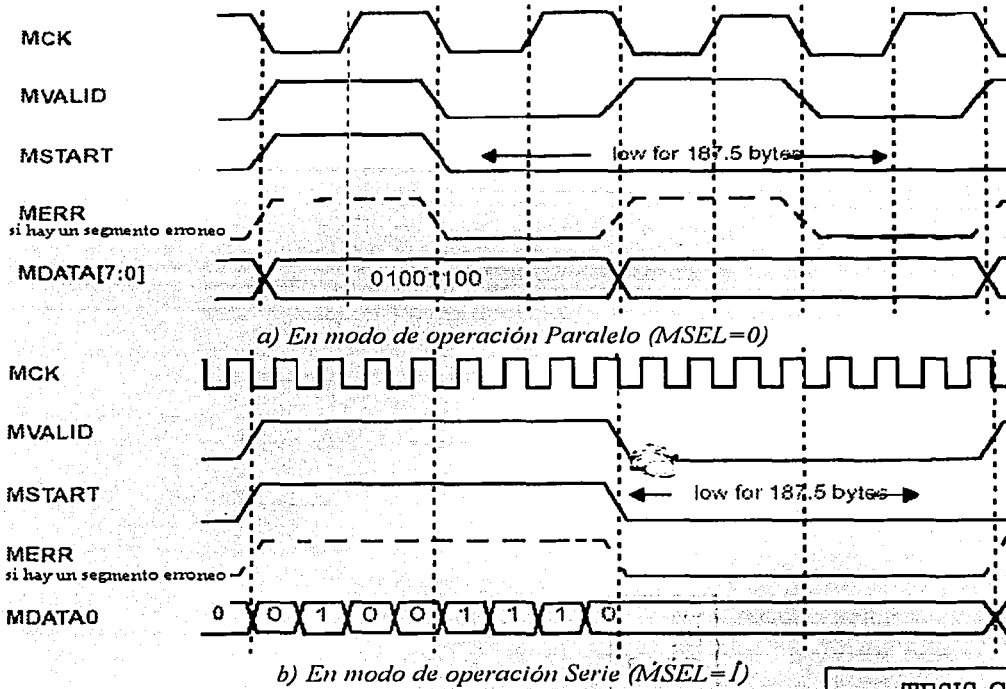


Figura 8.8 Señales de manejo de los paquetes de datos.

TESIS CON FALLA DE ORIGEN

8.2.4 Transporte, demultiplexaje y decodificación de vídeo.

La siguiente etapa de un receptor de DTV es la capa de transporte y demultiplexaje de los datos, con la que se obtienen los correspondientes PES (paquetes elementales) tanto de vídeo como de audio. Después se decodifican estos paquete de vídeo y audio para obtener una señal analógica que pueda ser interpretada y presentada por un televisor analógico. Aunque estas etapas están compuestas de una gran variedad de procesos no se arma cada proceso por separado, al igual que en el caso anterior, ya que en el mercado existen CI que contienen estos procesos en un solo CI Incluso el que elegimos para este desarrollo contiene además de la etapa de transporte y demultiplexaje, también la etapa de decodificación de vídeo, el CI es fabricado por Motorola y se le conoce como: procesador digital de vídeo ATSC MCT4000^[6]. Existen en el mercado también CI's de marcas que mencionamos con anterioridad (NextWave, Phillips, Teralogic, etc.). Los CI's pueden cambiar en cualquier momento debido a que mejoran o evolucionan e incluso se está planeando crear un CI que además de las etapas mencionadas (Demultiplexaje y Decodificación de vídeo) cuente también con la capacidad de decodificar el audio.

Las señales de entrada a este CI serán los paquetes de datos obtenido de la salida del CI de demodulación VSB, por lo cual en esta etapa se interpretaran los paquetes y se realizara la separación (demultiplexado) de los paquetes de video y audio, los paquetes de video se envían al decodificador interno de video, el cual después de procesarlas las enviara como salidas, los paquetes de audio también serán salidas del sistema. El CI requiere del auxilio de una memoria SDRAM (Synchronous Dynamic Random Acces Memory) externa para el vídeo y del procesador maestro (Host). La SDRAM de la figura 8.9 es utilizada por todos los dispositivos que requieran de una memoria auxiliar para poder guardar registros y después recuperarlos. El MCT 4000 la utiliza principalmente para guardar datos de la etapa de transporte y demultiplexaje, que son requeridos para la descompresión tanto del video y el audio, así como la recuperación y reensamblado de las tablas PCI (información específica del programa). La etapa de decodificación de video almacena en esta memoria los datos de las imágenes I e imágenes P para la reconstrucción de las imágenes B que posteriormente serán recuperadas por el post-procesador de pantalla. La capacidad de esta memoria será de 64 Mbits y la que elegimos es del fabricante Vss Burts^[11].

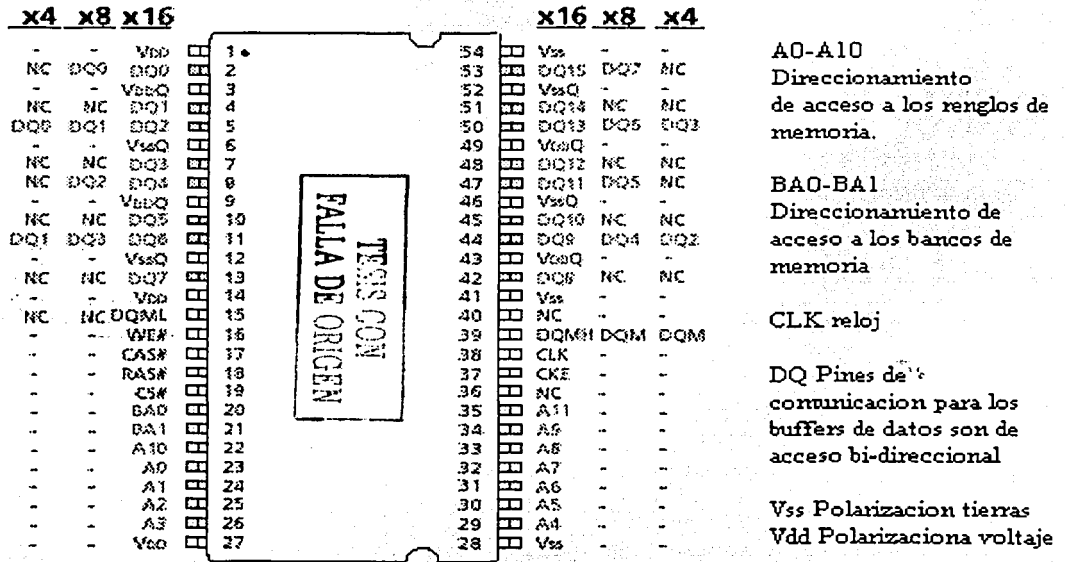
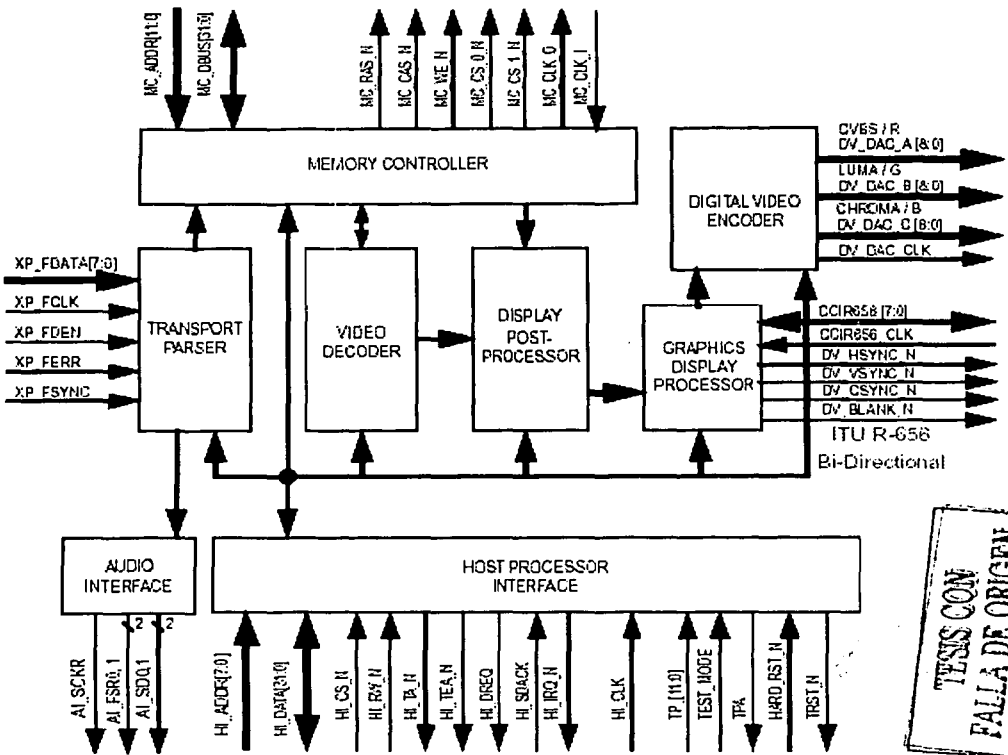


Figura 8.9 SDRAM de 64 Mbits.

El diagrama de bloques del MCT4000 se muestra en la figura 8.10. En este se representan las principales etapas y procesos que se realizan dentro de éste.



TESIS CON FALLA DE ORIGEN

Figura 8.10 Diagrama de bloques del CI MCT400.

El bloque de transporte y demultiplexaje se encarga de “separar” los paquetes de datos de audio y video para formar sus PES (paquet elementary stream) correspondientes. Se envían los PES de video al decodificador de video y los de audio a la interfase de audio. El siguiente bloque es el decodificador de video el cual decodifica y decuantiza los datos de entrada, transformando los datos al dominio de los pixeles, desarrolla la compensación de movimiento para obtener las imágenes P y B, las cuales envía a la SDRAM con auxilio del bloque de control de memoria.

El post-procesador de pantalla adapta los paquetes de video para ser desplegados sobre pantallas de 720 x 480 pixeles la cual puede ser desplegada con una relación de aspecto de 16:9 o 4:3. El control de memoria tiene como propósito vigilar, regular y controlar toda la interacción que tienen los dispositivos internos con la SDRAM externa. El procesador de imágenes en pantalla es usado para desplegar menús en pantalla como son : guías de programas, texto, letreros, etc. . Por ultimo esta la interfase de audio y la interfase con el

Capítulo # 8

host. La primera es utilizada para almacenar los PES de audio y enviarlos al decodificador de audio, a si como datos auxiliares para este, como son: el reloj, sincronía, etc. El segundo es un controlador para la interacción de los dispositivos con el Host. El codificador de vídeo digital se encarga de tomar los datos del post-procesador de vídeo y codificarlos para proporcionar a la salida paquetes de bits que representan la información de color Rojo, Verde y Azul.

Los pines de este CI se encuentran mediante coordenadas como se observa en la figura 8.11. En la tabla 8.2 se encuentra un descripción de las principales entradas y salidas de este CI, así como el número de pin al que corresponde y su nombre de identificación. Hay que tomar en cuenta que no mencionaremos todos los pines ya que algunos de ellos son para polarización, pines de prueba u otras aplicaciones; solo se hace mención de los que se ocupan de las operaciones básicas y de mayor interés.

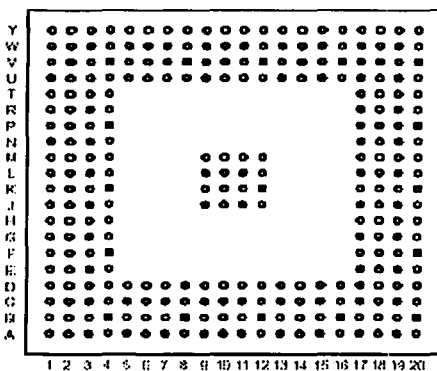


Figura 8.11 Localización de pines del CI MCT4000.

Coordenada del pin	Nombre del pin	Descripción:
P18	XP FDATA 7	Son los pines de acceso a la información proveniente del demodulador VSB por estos pines entra el paquete de datos, en el caso del diseño se conectan con la salida MDATA del CI MCT2100.
U20	XP FDATA 2	
V20	XP FDATA 0	
R18	XP FDATA 3	
R19	XP FDATA 5	
R20	XP FDATA 6	
T19	XP FDATA 1	
T20	XP FDATA 4	
W17	XP_FCLK	Este pin es la conexión del reloj proporcionado por el demodulador VSB se conecta con el pin de MCK.
Y19	XP_FDEN	Es el pin de acceso para la validación del paquete el cual indica cuando hay datos en el bus se conecta con MVALID

Diseño de un Receptor para DTV

Y17	XP_FERR	Es el pin de acceso para la indicación de que existe la transmisión de un byte de un paquete erróneo se conecta con el MERR.
W18	XP_FSYNC	Es el pin de acceso para la sincronía proporcionada por el MCT2100, se conecta con el pin SYNC.
A13-A17 B13-B16 C13-C15	MC_ADRR_11,9,6,3,0 -MC_ADRR_10,7,5,2 MC_ADRR_8,4,1	Estos pines son el control de las direcciones para la memoria SDRAM, se conectan a los pines A0-A11.
C20	MC_DBUS_31	Estos pines son utilizados para el intercambio de datos con la memoria SDRAM los pines son bi-direccionales por lo cual el CI MCT4000 puede leer y escribir en la SDRAM a través de ellos, estos se conectan con los pines de la SDRAM DQ.
D19-D20	MC_DBUS_29, 28	
E18-E20	MC_DBUS_30, 26, 25	
F18-F20	MC_DBUS_27, 24, 22	
G18-G20	MC_DBUS_23, 20, 19	
H18-H20	MC_DBUS_21, 18, 16	
J18-J20	MC_DBUS_17, 15, 14	
K18-K20	MC_DBUS_13, 12, 11	
L18-L20	MC_DBUS_8, 9, 10	
M18-M20	MC_DBUS_5, 6, 7	
A18	MC_WE_N	Pines utilizados para enviar comandos, banderas, pedir acceso, etc, a la memoria SDRAM.
A19	MC_RAS_N	
B18	MC_CAS_N	
B17	MC_CS1_N	Utilizador para consultar los bancos de memoria de la SDRAM se conectan a los pines BA0 y BA1.
C16	MC_CS0_N	
B20	MC_CLK_1	Reloj de entrada dado por SDRAM
C19	MC_CLK_0	Reloj de salida enviado a la SDRAM
A6	DV_CS_N	Envía la sincronía de señal de color (Digital)
A7	DV_HSYNC_N	Envía la sincronía horizontal (Digital)
B7	DV_VSYNC_N	Envía la sincronía vertical (Digital)
B8	DV_BSYNC_N	Envía sincronía de señal blanco/negro (Digital)
B1,B2	DV_DAC_A_5,2	Señal digital de salida de brillo rojo R (croma), la cual es enviada a un convertidor D/A.
C1-C4	DV_DAC_A_8,4,1,0	
D2,D3	DV_DAC_A_7,3	
E3	DV_DAC_A_6	
D1	DV_DAC_B_0	
E1,E2	DV_DAC_B_3,2	Señal digital de salida de brillo verde G (luminancia), la cual es enviada a un convertidor D/A.
F1-F3	DV_DAC_B_6,5,1	
G2,G3	DV_DAC_B_8,4	
H3	DV_DAC_B_7	
G1	DV_DAC_C_0	Señal digital de salida de brillo azul B (croma), la cual es enviada a un convertidor D/A.
H1,H2	DV_DAC_C_3,2	
J1-J3	DV_DAC_C_6,4,1	
K1-K3	DV_DAC_C_8,7,5	

Capítulo # 8

C8	DV_DAC_CLK	Señal de reloj enviada al convertidor D/A para su sincronía con el CI.
M3	HI_ADDR 0	Pines para que el Host controle las direcciones de acceso de MCT400, se conecta a los pines de dirección del Host.
N3	HI_ADDR 1	
P2-P3	HI_ADDR 3, 4	
R2-R3	HI_ADDR 5, 6	
T1	HI_ADDR 7	
T2	HI_DATA 0	
U1-U2	HI_DATA 1, 2	
V1-V2	HI_DATA 3, 4	
V7-V11	HI_DATA_15, 18, 21, 24, 29	
W1	HI_DATA 5	
W3-W12	HI_DATA_7, 9, 11, 13, 16, 19, 22, 25, 28, 31	Utilizados para comunicación de bus de datos entre el Host y el CI MCT4000 son pines bi-direccionales para el intercambio de información.
Y2-Y12	HI_DATA_6, 8, 10, 12, 14, 17, 20, 23, 26, 27, 30	
W13	HI_RW N	
W14	HI_TEA N	
W15	HI_SDACK	
Y13	HI_CS N	
Y14	HI_TA N	
Y15	HI_IRQ N	
Y16	HI_DREQ N	
W16	HI_CLK	
A11, B11	AI_FSR0, 1	Pines para la comunicación entre el host y el MCT4000 mediante estos pines se intercambia información como banderas y se controla el acceso y la comunicación entre estos dispositivos
A12, B12	AI_SDI, 0	
C12	AI_SCKR	
A1, A2, A20, B3, B19, C18, D4, D8, D12, D14, D17, F4, G17, H4, J9-J17, K4-K12, L9-L17, M9-M12, N17, P4, T4, T17, U4, U6, U10, U13, U15, U17, V3, V4, V6, V18, W2, W19, Y1, Y20	GND	Conexiones a tierra
C17, D6, D10, D16, D18, E17, M4, P17, T3, T18, U3, U5, U8, U11, U18, V5, V19, W20.	VDD ₃₃	Conexiones a voltaje, valor típico de 3.3 volts

**TESIS CON
FALLA DE ORIGEN**

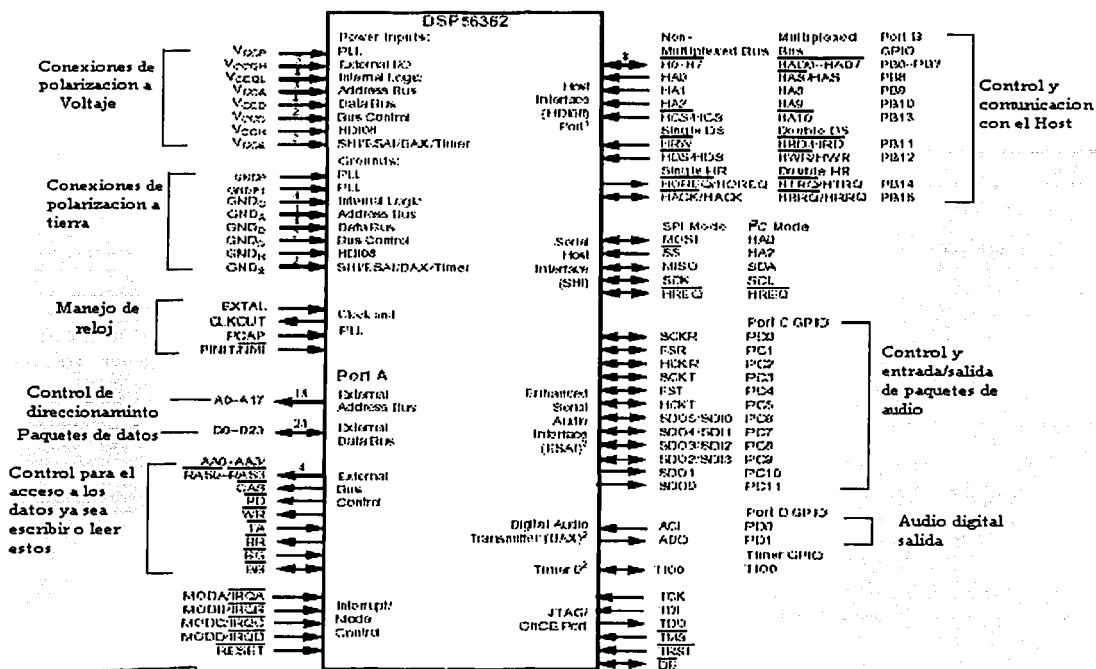


Figura 8.13 Descripción de pines del CI DSP56362.

TESIS CON FALLA DE ORIGEN

Nombre de grupo	Descripción
Conexiones de polarización Vcc y Gnd	Voltajes utilizados para el funcionamiento del CI y las etapas que lo componen. Vcc típico de 3.3 volts
Control y comunicación con el Host	Este puerto de comunicación permite obtener y enviar instrucciones y datos al procesador maestro que son de utilidad y requeridas para este proceso.
Control de entrada/salida de paquetes de audio	En nuestra aplicación este puerto se utilizará como entrada para los PES de audio. Es un puerto de acceso en serie para los datos. Los pines importantes de este puerto son SDIO-SD13, los cuales son las entradas serie de los paquete de audio, FSR para la sincronía de los datos, SCKR para la entrada del reloj externo.

Audio digital salida	En este puerto se cuenta con dos pines que son el ACI entrada de reloj para el control del flujo de los datos de audio y el ADO el pin de salida del audio decodificado y descomprimido en forma digital.
----------------------	---

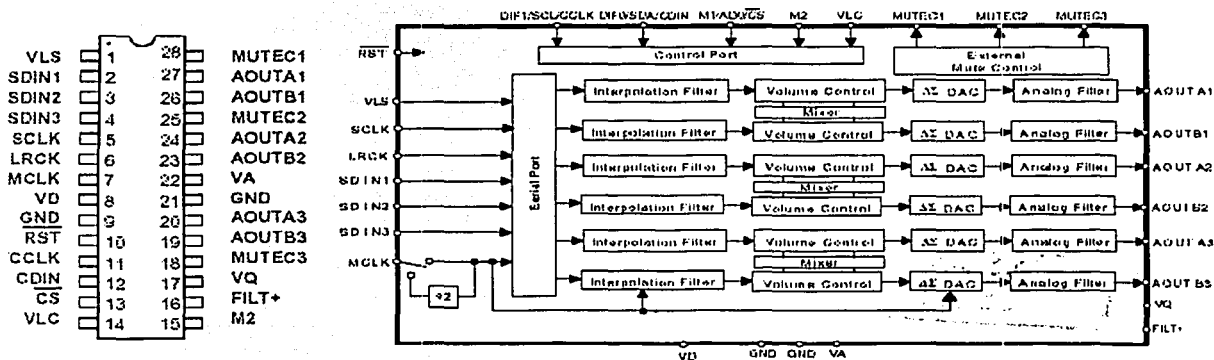
Tabla 8.3 Descripción de puertos del CI DSP56362.

TESIS CON FALLA DE ORIGEN

8.2.6 Convertidor Digital a Analógico para audio y vídeo.

❖ CDA para audio.

Debido a que los CI's de decodificación tanto de audio como de vídeo entregan las señales en forma digital es necesario utilizar un convertidor Digital-a-analógico para transformar esta señal digital a una señal analógica, la cual pueda ser interpretada por el televisor analógico, en la figura 8.14 se muestra el diagrama de bloques del CDA (convertidor digital – analógico) utilizado que es el CS4360^[12]. Este CDA se utiliza para aplicaciones específicas de audio.



a) CDA CS4360

b) Diagrama de bloques del CDA CS4360

Figura 8.14 Convertidor Digital a analógico para audio.

El circuito que debe armarse para el manejo y la polarización de este CI se muestra en la figura 8.15. Se observa que cuenta con tres entradas para datos en serie, de las cuales solo se utilizará una en este diseño. También se observa que se pueden obtener los 6 canales de salida que se requieren o soporta el estándar (los 5.1 canales). Sin embargo es necesario apuntar que debido a la aplicación que se está desarrollando, la cual es para una TV analógica, este tipo de televisores a lo mucho cuentan con dos bocinas L (izquierda) y R (derecha), aunque la mayoría solo cuenta con una bocina. Creemos que es suficiente proporcionar dos canales de audio de salida que serían L y R aunque se cuenta con la

Capítulo # 8

posibilidad de ofrecer los 5.1 canales. En la tabla 8.4 se hace la descripción de los pines de este CDA.

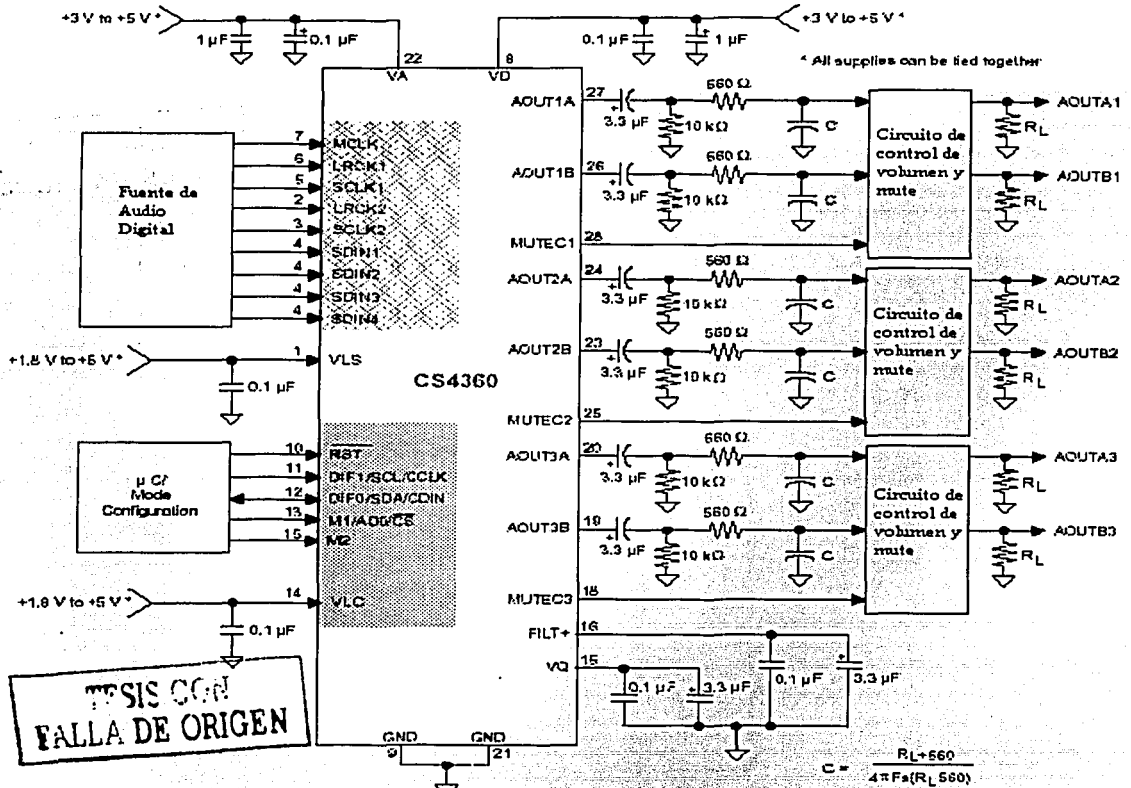


Figura 8.15 Circuito para conversión de audio digital a analógico.

# de pin	Nombre de pin	Descripción
1	VLS	Determina el nivel de voltaje requerido para los datos en serie.
2,3,4	SDIN1, 2, 3	Entrada para datos en serie.
5	SCLK	Reloj de sincronía para el bus serie de datos
6	LRCK	Reloj que indica a que bocina pertenece la secuencia de datos que se encuentra en el bus, si corresponde al canal L o R
7	MCLK	Reloj maestro
8	VD	Voltaje de polarización
9	GND	Conexión a tierra.

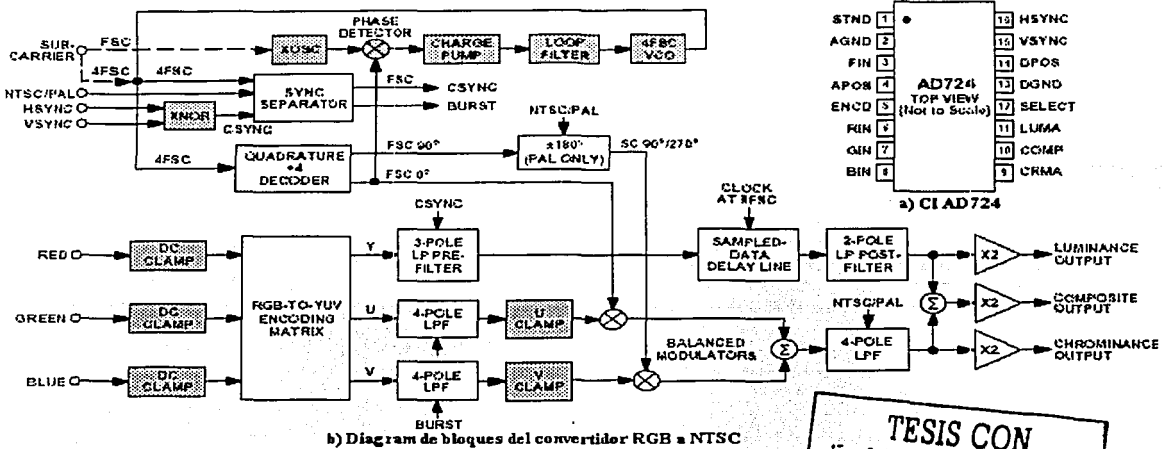
# de pin	Nombre del pin	Descripción
1-10	BPb0-BPb9	Entrada paralelo del tren de bits correspondiente a la señal color Azul (B). Conecta a pin MCT4000 DV DAC C [8:0].
11	DVss	Conexión a tierra para señal digital.
12	DVDD	Conexión a voltaje directo entre 3 y 5 volts
13-22	RPr0-RPr9	Entrada paralelo del tren de bits correspondiente a la señal de color rojo (R). Conecta a pin MCT4000, DV DAC A [8:0].
23	BLANK	Señal de sincronía de Blanquing conecta a pin MCT4000, DV BLANK N.
26	CKL	Reloj de sincronía para los trenes de datos de entrada, conecta con pin MCT4000, DV DAC CKL.
27-36	GY0-GY9	Entrada paralelo del tren de bits correspondiente a la señal de color verde (G). Conecta a pin MCT4000, DV DAC B [8:0].
37	V _{REF}	Voltaje de referencia para el CDA.
40,44	AV _{DD}	Polarización con voltaje analógico 5 volts
41	AGY	Señal analógica de salida correspondiente al color verde (G).
42,46	AVss	Conexión a tierra para señal analógica.
43	ARPr	Señal analógica de salida correspondiente al color rojo (R).
45	ABPb	Señal analógica de salida correspondiente al color azul (B).

Tabla 8.5 Descripción de pines del CDA THS8133.

Como se puede observar se obtienen las salidas de video analógicas en el formato RGB (rojo-verde-azul) y aun cuando esta salida es útil para televisores que cuenten con este tipo de acceso, es importante hacer notar que la mayoría de los televisores de nuestro país solo operan bajo el formato de NTSC (National Television System Committee) por lo cual es necesario realizar una conversión de formato RGB a NTSC.

❖ Convertidor RGB a NTSC.

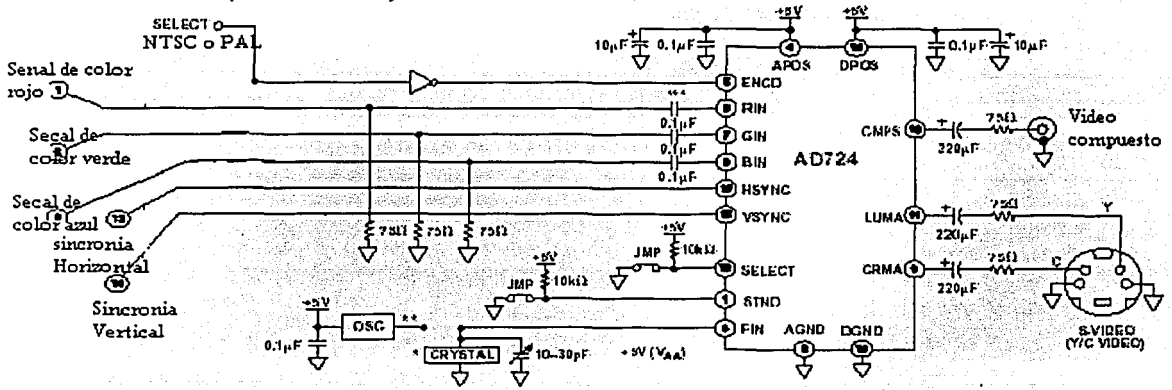
Para lograr el objetivo de convertir la señal al formato NTSC proponemos la utilización del CI AD724^[14] de Analog Devices, este CI es un codificador de video de RGB a NTSC(este circuito también puede ser utilizado para convertir a PAL, Phase Alternation Lines), y se muestra en la figura 8.17 a). El diagrama de bloques de este CI se muestra en la figura 8.17 b), este circuito ofrece 2 salidas de super-video (señal de Luminancia Y y Crominancias Cb, Cr) y la salida de video compuesto en el estándar elegido (en nuestro caso NTSC), la descripción de los pines de importancia se encuentra en la tabla 8.6.



TESIS CON FALLA DE ORIGEN

Figura 8.17 CI convertidor RGB a NTSC

El circuito que debe armarse para obtener el funcionamiento deseado del CI convertidor de RGB a NTSC es el que se muestra en la figura 8.18. Con esta etapa se completa el objetivo de nuestro diseño, que es proporcionar las señales de audio y video de tal manera que puedan ser procesadas y desplegadas en la pantalla de un televisor analógico. Las señales serán entregadas al televisor analógico por medio de los cables y conexiones de audio y video. Las salidas de audio y video del decodificador o receptor de DTV se le entregan a las entradas de audio y video de televisor analógico por medio de estos conectores.



FALL. Figura 8.18 Circuito para convertidor de RGB a NTSC.

TESIS CON FALLA DE ORIGEN

# de pin	Nombre de pin	Descripción
1	STDN	Entrada lógica en alto activa formato NTSC, entrada lógica en bajo activa PAL.
2	AGND	Conexión a tierra
3	FIN	Reloj externo, cristal de cuarzo
4	POS	Voltaje analógico de polarización 5 V
5	ENCD	CI encendido para codificar
6	RIN	Entrada de señal analógica de color rojo (R). Conecta a pin de THS8133, ARPr
7	GIN	Entrada de señal analógica de color verde (G). Conecta a pin de THS8133, AGY
8	BIN	Entrada de señal analógica de color azul (B). Conecta a pin de THS8133, ABPb
9	CRMA	Salida de señal de cromancia
10	COMP	Salida de señal analógica de vídeo compuesto en el formato seleccionado (en nuestro caso NTSC)
11	LUMA	Salida de señal luminancia
13	DGND	Conexión a tierra para señal digital
14	DPOS	Voltaje directo de polarización 5 volts
15	VSYNC	Señal de sincronía vertical. Conectada a pin de MCT4000, DV VSYNC N
16	HSYNC	Señal de sincronía horizontal. Conectada a pin de MCT4000, DV HSYNC N

Tabla 8.6 Descripción de pines del convertidor RGB a NTSC AD724.

8.2.7 Procesador Maestro (Host).

Como ya habíamos mencionado con anterioridad los circuitos integrados utilizados dependen o requieren del apoyo de un procesador maestro, para el control de la sincronía y la comunicación entre ellos. El que proponemos utilizar para este diseño es el microprocesador de Motorola para comunicaciones el MPC850^[6]. Este será el punto de referencia para el funcionamiento adecuado del sistema. En la figura 8.19 se muestra el diagrama de asignación de pines del MPC850. Esta se puede resumir al decir que la interconectividad con los demás dispositivos se realizará a través de los pines de acceso a direccionamiento, intercambio de datos, acceso a memorias, reloj y sincronía. Es procesador se encarga también de manejar la información necesaria para programar la interfase con el usuario.

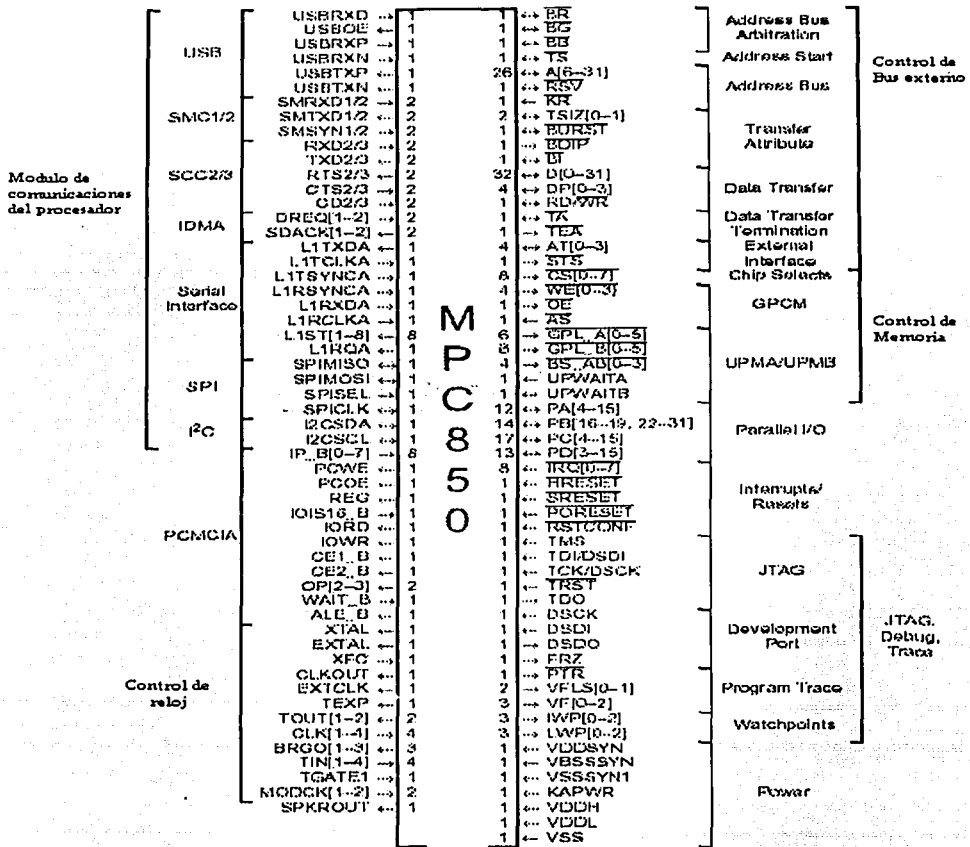


Figura 8.19 Diagrama de pines de interacción del CI MPC850.

El procesador maestro es un dispositivo necesario de acuerdo con la arquitectura que desarrollamos el diseño, y es necesario para auxiliar en los procesos de decodificación a las demás etapas. Cuando existe la selección de un nuevo canal, el procesador puede inicializar y resetear a las unidades, además es el puerto de acceso para el manejo o interfase del usuario.

TEMA CO
FALLA DE ORIG:

8.3 Fuente de alimentación.

Otro dispositivo importante en el diseño es la fuente o circuito de polarización, el cual debe proporcionar a los CI's y dispositivos del sistema un voltaje directo, constante y controlado, con la suficiente corriente. Se propone la siguiente fuente controlada de voltaje la cual será diseñada a un voltaje de 5 volts, debido a que los circuitos del sistema no requieren más voltaje y a una corriente máxima de 40 mA por CI, por lo cual en el diseño se utilizará un transformador de 127 a 12 volts con tap central y corriente de 500 mA. También se usarán diodos, regulador de voltaje y capacitores. Además se agrega al inicio de la fuente un circuito de protección; el que se eligió es uno de los mas eficientes ya que los varactores (V) tiene la capacidad de controlar los picos de corriente. El circuito de la fuente tendría las características que se muestran en la figura 8.20.

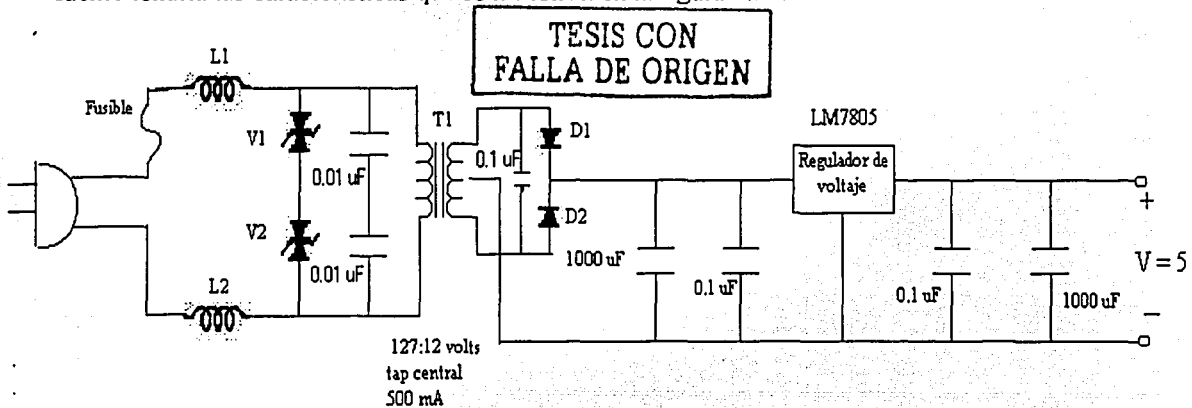


Figura 8.20 Circuito de fuente de polarización para el sistema.

La fuente anteriormente descrita, es una de las más sencillas y eficientes que se conocen, consideramos que esta tiene la capacidad suficiente para soportar la polarización del sistema de circuitos que conformará la tarjeta decodificadora o receptora de DTV en el estándar de ATSC.

8.4 Tarjeta decodificadora de DTV.

En la figura 8.21 se muestra el diagrama final de circuito completo de nuestro diseño de receptor para DTV. Es importante mencionar —como ya lo hicimos antes— que estos CI y demás componentes que conforman el diseño pueden variarse, por dispositivos que desempeñen la misma tarea y entreguen el mismo resultado, ya que existe una constante evolución de estos componentes. También se pueden sustituir por sus equivalentes en otras marcas o modelos debido a la competencia entre fabricantes de semiconductores que ofrecen nuevos y mejores CI's periódicamente.

El diseño contiene los principios y operaciones básicas esenciales para el armado y fabricación de un decodificador de DTV, por lo cual es un apoyo para comenzar el armado y fabricación de un decodificador o receptor de ATSC en México y desde luego también dentro de nuestra Universidad, sin importar el cambio o sustitución de los dispositivos mencionados.

En el siguiente apartado se muestra la lista de los componentes requeridos para armar nuestro diseño.

8.4.1 Lista de componentes

❖ Resistores

- 3 resistores de 330Ω a 1/2 Watt
- 1 resistor de $33 \text{ k}\Omega$ a 1/2 Watt
- 3 resistores de $4.7 \text{ k}\Omega$ a 1/2 Watt
- 5 resistores de $22 \text{ k}\Omega$ a 1/2 Watt
- 1 resistor de 4.7Ω a 1/2 Watt
- 3 resistores de $10 \text{ k}\Omega$ a 1/2 Watt
- 1 resistor de 680Ω a 1/2 Watt
- 2 resistores de $3.9 \text{ k}\Omega$ a 1/2 Watt
- 4 resistores de 50Ω a 2 Watt
- 2 resistores de $1.8 \text{ k}\Omega$ a 1/2 Watt
- 2 resistores de $50 \text{ k}\Omega$ a 1/2 Watt
- 1 resistor de $100 \text{ k}\Omega$ a 1/2 Watt
- 1 resistor de 560Ω a 1/2 Watt
- 6 resistores de 75Ω a 1/2 Watt
- 2 resistores de $6.8 \text{ k}\Omega$ a 1/2 Watt

❖ Capacitores

- 4 capacitores de 1 nF
- 2 capacitores de 15 pF
- 2 capacitores de 1 pF
- 2 capacitores de 1.8 pF
- 1 capacitor de 2.2 pF
- 1 capacitor de 2.5 pF
- 1 capacitor de 8.2 pF
- 1 capacitor de 47 pF
- 1 capacitor de 100 pF
- 2 capacitores de 4.7 nF
- 2 capacitores de 10 nF
- 1 capacitor de 18 pF
- 1 capacitor de 100 nF

- 1 capacitor de 1.2 nF
- 1 capacitor de 0.22 μ F a 25 v
- 17 capacitores de 0.1 μ F a 25 v
- 1 capacitor de 220 pF
- 1 capacitor de 4.7 pF
- 1 capacitor de 10 pF
- 2 capacitores de 1 μ F a 25 v
- 4 capacitores de 3.3 μ F a 25 v
- 2 capacitores de 10 μ F a 25 v
- 3 capacitores de 220 μ F a 25 v
- 2 capacitores de 1000 μ F a 225v

❖ Bobinas

- 1 bobina variable de 2 a 51 μ H
- 2 bobinas de 23 nH
- 1 bobina de 30 nH
- 2 bobinas de 80 nH
- 1 bobina de 33 H

❖ Diodos

- 1 diodo BB134
- 1 diodo BB133
- 1 diodo BA792
- 3 diodos 1N4004
- 1 Zener 1N5222 a 2.5 V 0.4 Watts
- 1 Zener 1N5225 a 3 V 0.5 Watts

❖ Cristales de Cuarzo

- 1 cristal de 4 MHz
- 1 cristal de 50 MHz
- 1 cristal de 3.5 MHz

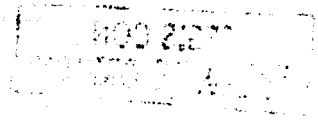
❖ Circuitos integrados

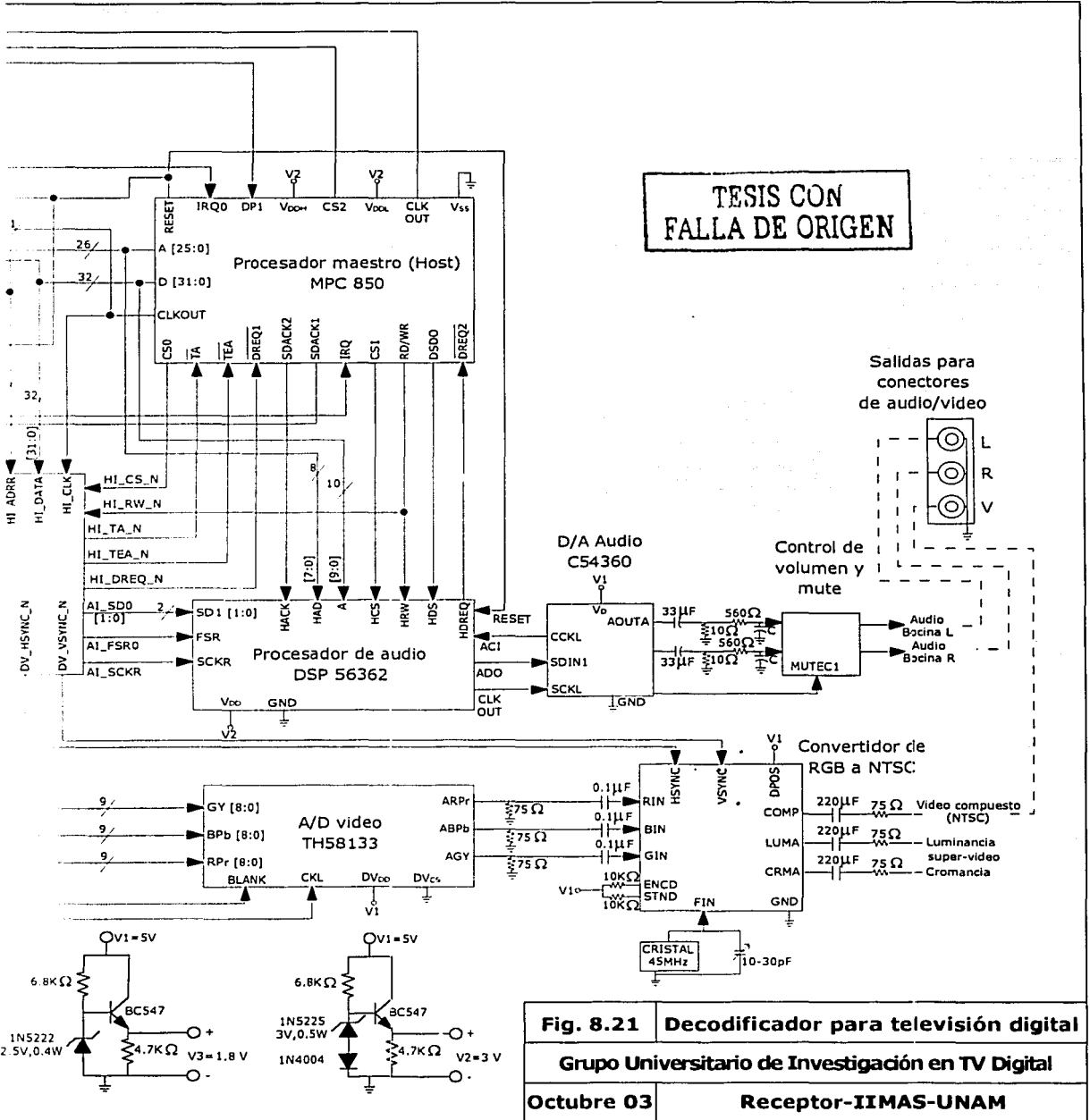
- Tuner TDA6402 Philips
- Filtro SAW de 6 MHz paso banda
- Amplificador de ganancia controlada MAX3861 Maxim
- Convertidor analógico/digital 10 bits MB40C360 Fujitsu
- Demodulador de 8-VSB&FEC MCT2100 Motorola semiconductors
- Memoria SDRAM vssBursts
- Transporte y decodificación de vídeo MCT4000 Motorola semiconductors
- Procesador de Audio DSP56362 Motorola semiconductors

- ✓ Procesador MPC850 Motorola semiconductors
- ✓ Convertidor Digital/Análogo de audio C54360 Crystal
- ✓ Convertidor Digital/Análogo de vídeo THS8133 Texas Instruments
- ✓ Convertidor de RGB a NTSC AD724 Analog Devices
- ✓ Regulador de voltaje LM7805

❖ **Otros**

- ✓ 1 Transformador de 127 a 12 volts, 500 mA con tap central
- ✓ 2 Varistores V130LA1 de 130 Vrms a 11 J
- ✓ 2 transistores BC547
- ✓ 1 fusible de 120 v a 10 A





8.4.2 Características del decodificador.

❖ Especificaciones Generales

- Estándar de transmisión: 8 VSB DTV (ATSC)
- Potencia requerida: 120 V AC, 50-60 Hz
- Consumo de potencia: 35 W
- Relación portadora ruido C/N: 14.9 dB
- Figura de ruido: 10 dB
- Ruido térmico: -106.2 dBm
- Ganancia de Antena: VHF: 4,6 , UHF 10.

❖ Señal de entrada

- Frecuencia: 54 MHz – 860MHz, canales VHF 2-13, UHF 14-69.
- Impedancia de la antena 75 ohms externa/interna (VHF/UHF).
- Ancho de Banda : 6 MHz
- Velocidad de símbolo: 10.76 Ms/s
- Velocidad de datos: 19.28Mb/s
- Formato de transporte : MPEG-2

❖ Video

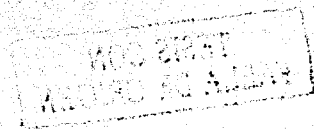
- Relación de aspecto: 4:3 y 16:9.
- Compatible con entrelazado de NTSC
- Salidas : Video compuesto (RCA) , S-video.

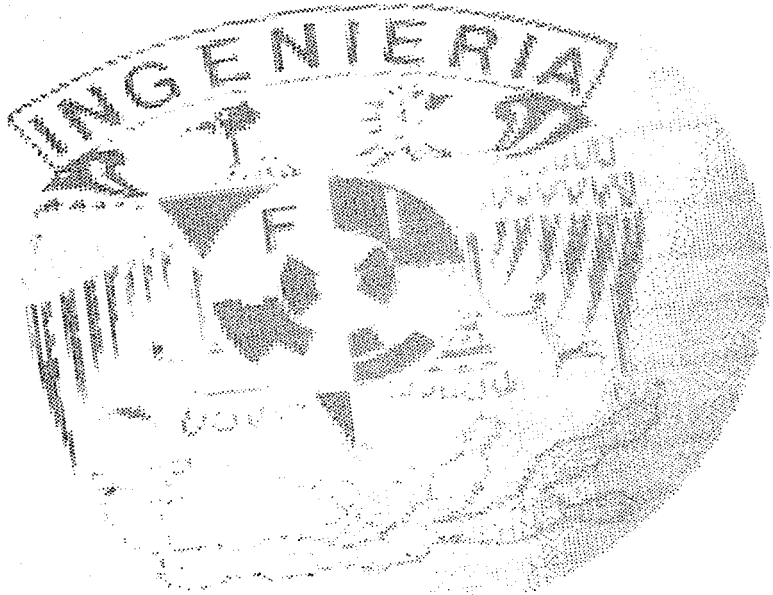
❖ Audio

- Dolby Digital AC-3
- Salida Análoga: 2 canales estéreo (RCA).

REFERENCIAS

- [1] Artículo de "Project TRB-00-1 Interim Report", 9de abril del 2001
- [2] Compañía de Tecnología ATI. Sitio en internet www.ati.com
- [3] semiconductores Philips. Sitio web www.phillipssemiconductor.com
- [4] Compañía de Tecnología ATI. Sitio web www.ati.com.
- [5] Semiconductores Philips. Sitio web <http://www.semiconductors.philips.com>
- [6] Semiconductores Motorola. Sitio web www.e-motorola.com
- [7] Corporación OSCILENT, distribuidor de osciladores de cristal, filtros etc. Sitio web <http://www.oscilent.com>
- [8] Semiconductores Dallas. Sitio web <http://www.maxim-ic.com>
- [9] Distribuidor de componentes y accesorios. Sitio web <http://www.nedis.com>
- [10] Fabricante de semiconductores, Zilog. Sitio web <http://www.zilog.com>
- [11] Compañía fabrican de Memorias y periféricos para PC. Sitio web <http://www.vssburst.com>
- [12] Soluciones para sistemas de Audio, Cirrus Logic. Sitio web <http://www.cirrus.com>
- [13] Distribuidor de semiconductores. Sitio web [http:// www.tij.co.jp](http://www.tij.co.jp)
- [14] Compañía de semiconductores, Analog Devices. Sitio web <http://www.analog.com>





TESIS CON
FALLA DE ORIGEN

204

CONCLUSIONES

La televisión es uno de los inventos que más ha revolucionado la historia humana, sin embargo en un mundo en que la tecnología digital está reemplazando a la analógica, ésta se ha quedado estancada. Las tecnologías de transmisión y recepción en la televisión no han experimentado un cambio significativo por unos 50 años. De ahí que durante los últimos se hayan desarrollado varias tecnologías de televisión digital. El cambio hacia la televisión digital es inminente, debido a las ventajas que posee, como son: codificación flexible, capacidad de almacenamiento, capacidad de procesamiento, detección y corrección de errores.

La investigación de este tipo de problemas es importante para el país ya que el cambio del televisor analógico a digital se tiene que dar, así como se dio el cambio de TV B/N al televisor a color, y debido al atraso tecnológico de nuestro país es necesario sentar las bases para cuando esta tecnología arribe al país. El mercado potencial es muy grande.

La Universidad Nacional Autónoma de México, institución que se destaca siempre por estar a la vanguardia en la investigación, a través del Grupo Universitario de Investigación en Televisión Digital, busca establecer las bases técnicas que permitan a la Universidad incursionar en el campo de la radiodifusión digital televisiva, así como en la recepción de la TV digital. Este equipo se divide en varios grupos de trabajo. El grupo de trabajo en el tema de Receptores, ha planteado diferentes actividades que permitan el conocimiento de estas tecnologías, con el propósito de diseñar un aparato receptor de TV digital, que sea competitivo con los que existen en el mercado.

Para poder desarrollar un decodificador es necesario comprender el funcionamiento de todos los procesos que se llevan a cabo en la televisión digital, para ello se tuvo que hacer una investigación sobre el tema. Lamentablemente, la bibliografía especializada en el tema disponible, es prácticamente inexistente. Se tuvo que recurrir a la información que hay sobre el tema en la internet, y no toda esta se encuentra al alcance de un equipo con recursos limitados, pues algunos estándares solo pueden ser consultados, pagando por verlos. También asistimos a presentaciones de fabricantes, como Rohde & Schwarz. Lo que se hizo en un principio, fue recopilar información, para luego estudiar los procesos de los que se vale esta tecnología para funcionar, que van desde la señal original de video, pasando por la codificación y las técnicas de trasporte, hasta llegar a la modulación.

Posteriormente se analizaron las hojas de especificaciones de los Set-Top Boxes, y de estas se obtuvieron las partes que integran a un decodificador. Durante este lapso, contactamos a las empresas fabricantes de decodificadores, para pedir asesoría, pero se negaron a darnosla, por considera que podrían verse afectados sus intereses.

Conclusiones

Después se realizó un diagrama de bloques de un decodificador, y se incluyeron en este, los elementos que no aparecían en las hojas de especificaciones, pero que durante la recopilación de datos, nos dimos cuenta que eran importantes. Una vez que se tuvo un boceto del decodificador, quisimos adquirir los componentes, las cuales se venden en grandes cantidades.

Durante nuestra estancia en el grupo de trabajo de Receptores, en el que participa el IIMAS, aprendimos en que consiste el trabajo de investigación, así como las dificultades que conlleva, sobre todo si se trata de un grupo de investigación estudiantil, con recursos limitados:

- Notamos la importancia de trazarse objetivos, así como tiempos para lograr alcanzarlos.
- Vimos que la organización es primordial, al trabajar en equipo.
- Nos dimos cuenta de que un decodificador, de televisión digital, aplica varios de los conocimientos que adquirimos durante nuestra formación: análisis y procesamiento digital de señales, antenas, programación, comunicaciones analógicas y digitales, filtrado de señales, electrónica para telecomunicaciones, etc.
- Conocimos, las dificultades de la investigación en un país como México, sobre todo si esta aborda una tecnología nueva, pues no hay muchas fuentes de información. Incluso no se puede acceder a los estándares, pues algunos de estos solo se pueden consultar, pagando por ellos.
- Asimismo vimos que es muy difícil, realizar un prototipo, si no se trabaja en alguna compañía que pueda comprar por volumen los componentes. Si no se cumple con dicha condición, los fabricantes simplemente no venden los circuitos. Las empresas manufactureras tampoco dan asesoría a los desarrolladores, si no se está trabajando con un grupo de investigación reconocido.

El grupo de trabajo eligió el estándar de ATSC debido a la influencia económica y tecnológica que Estados Unidos, país donde fue creado este estándar, ejerce en nuestro país. Además el sistema de televisión analógico actual de México opera bajo la norma NTSC originaria del mismo país.

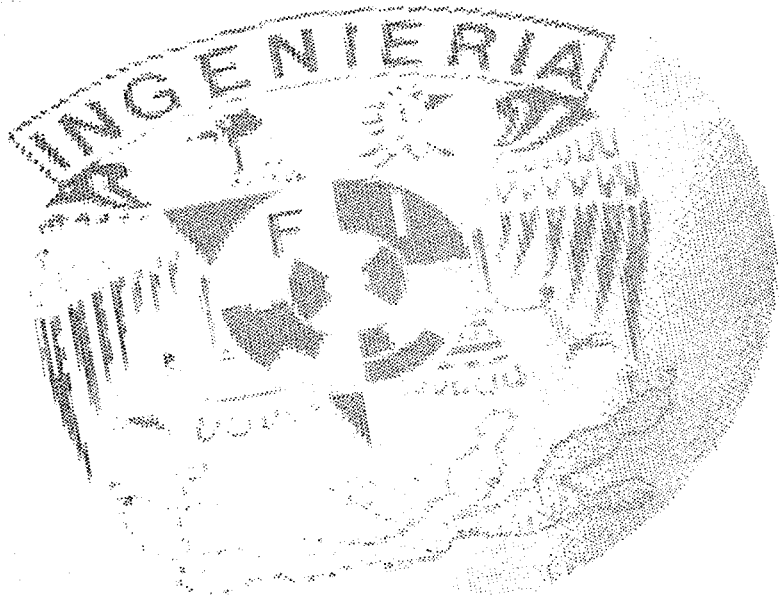
Se cumplió con el objetivo de la tesis ya que se logró obtener el diseño en papel del decodificador, se ha comenzado con la fabricación del decodificador y cabe hacer la aclaración que desde un inicio, se planteó fuera del alcance de esta tesis, pero si duda será de gran ayuda para la continuidad en el trabajo.

Estamos conscientes de que al implementar el diseño pueden ocurrir fallas o discrepancias con el funcionamiento teórico que se ha planteado, lo cual es de esperarse cuando se construye un prototipo en algún proyecto. Tenemos que apuntar que nos hubiera gustado construir el diseño, pero el trabajo restante será tema de proyectos futuros.

Entre las carencias del documento está el que no profundiza en el análisis de los estándares europeo y japonés para justificar el por que de nuestra elección de forma mas detallada, debido a que el trabajo en si ya es muy extenso. Otra razón es que para ello se requeriria realizar pruebas con los tres estándares y el grupo de trabajo no tiene la capacidad económica para adquirir los equipos necesarios.

Recientemente el grupo universitario de trabajo de investigación en TV digital ha adquirido un equipo transmisor bajo la norma ATSC.

TESIS CON
FALLA DE ORIGEN



TESIS CON
FALLA DE ORIGEN

289

BIBLIOGRAFÍA

- "A guide to MPEG fundamentals and Protocol Analysis"; Tektronix.
- Álvaro Arriagada Escudero et al. "FEC (Forward Error Correction) y Código Reed-Solomon". Departamento de Ingeniería Eléctrica- Universidad de Concepción. 2001
- ATSC Standard: Digital Audio Compresión (AC-3), Revisión A, A/52A, Agosto 2001.
- ATSC Standard: Digital Televisión Standard, Revisión B, A/53, A/53A, A/53B, Agosto, 2001
- "ATSC theory and fundamentals"; Rhode and Schwarz Training Center, 2002.
- Basilio, B; et. al; " Propuesta para la migración a un sistema de transmisión de televisión digital terrestre en TV UNAM. Tesis para obtener el título de ingeniero en telecomunicaciones; UNAM; México, 2002.
- Cristóbal Jara Rodríguez et. al. "Fundamentación del Código de Trellis". ". Departamento de Ingeniería Eléctrica- Universidad de Concepción. 2001
- Davis, M.; "The AC-3 multichannel coder"; Audio Engineering Society, 95th Convention, 1995.
- Delgado Gutiérrez Alejandro. "FLUJOS DE PROGRAMA Y DE TRANSPORTE MPEG-2. APLICACIÓN A DVB". Universidad politécnica de Madrid. España. 2002
- "Digital Audio Compresión (AC-3), Revision A"; Estándar ATSC A/52A; ATSC; E.E.U.U. 2001.
- "Digital Television Standard, Revision B with Amendments 1 and 2"; Estándar ATSC A/53B; ATSC; E.E.U.U. 2003.
- ORFANIDIS, S.; "Introduction to Signal Processing"; Prentice Hall; E.E.U.U., 1996.
- Todd, C; et. al. ; "AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"; Audio Engineering Society, 96th Convention, 1994.
- Vernon, Steve; "Design and implementation of AC-3 coders"; IEEE Tr. Consumer Electronics, Vol. 41, No. 3, August 1995.

Bibliografía

- <http://www.cybercollege.com/span/typ009.htm>
- <http://www.ufasta.edu.ar/amura/radio/typ042.htm>
- http://www.filmonline.com.ar/nuevo/notas/hdtv/hdtv00_00index.htm
- <http://www.canalmp.com/main.php3?pl=busca-result&formulari=noticias&pubID=1&listados=1&num=10&days=&campo=&categoria=&palabra=tv+digital>
- <http://www.asenmac.com/tvdigital/usa.htm>
- http://www.tektronix.com/Measurement/App_Notes/DigitalTV/25W-7203-3.pdf
- <http://www.digitaltelevision.com/weblinks.shtml>
- <http://www.videopopular.es/Editorial99.htm>
- <http://mhp.det.uvigo.es/mhp/>
- libro: <http://www.digitaltelevision.com/publish/dtvbook/toc.shtml>
- <http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/>
- <http://mhp.det.uvigo.es/mhp/>
- <http://www.sonicolor.es/catalogo/antenas/paginas/antenas/antenas3.html>
- <http://www.tienda-red.com/consulta.phtml?familia=MATERIAL%20PARA%20TELEVISION>
- <http://www.isesatv.com/ecommerce/index.html>
- <http://www.engel.es/pdf/spanish/Ip0405f.pdf>
- <http://coaxicom.com/>
- <http://www.indeca.com.ar/coax75/>
- <http://www.costonet.com.mx/main.asp?Clave=ELECAB03005>
- http://catalogo.syscom.com.mx/conectores_RF.htm
- <http://www.diotronic.com/bnc.htm>
- <http://www.lombiz.com/gamaterrestre.html>

-
- <http://www.tecatel.com/productos/detalle.asp?cod=111&pos=0>
 - <http://www.mecxico.com/>
 - <http://www.antenasmalaga.com/productos/ZZB.pdf>
 - <http://www.amphion.com/video.html>
 - <http://www.oren.com/pdf/OR51210DataSheet.pdf>
 - <http://www.semiconductors.philips.com/catalog/219/283/27117/42500/27140/27144/index.html#27144>
 - <http://my.semiconductors.com/acrobat/literature/9397/75005383.pdf>
 - http://www.platan.ru/pdf/dist/philips/TDA8960_1.pdf
 - http://www.semiconductors.philips.com/markets/consumer/pc_tv/atsc/
 - http://www.videon-central.com/products/product_detail_11.html
 - <http://www.ati.com/technology/hardware/nxt2000/>
 - <http://e-www.motorola.com/collateral/MCT2100AI.pdf>
 - <http://e-www.motorola.com/brdata/PDFDB/docs/MCT4000AI.pdf>
 - <http://e-www.motorola.com/brdata/PDFDB/docs/DSP56364.pdf>
 - <http://e-www.motorola.com/brdata/PDFDB/docs/MPC850TS.pdf>
 - <http://e-www.motorola.com/brdata/PDFDB/docs/DSP56362P.pdf>
 - [http://whatis.techtarget.com/definition/0,,sid9_gci332161,00.html\(8-VSB\)](http://whatis.techtarget.com/definition/0,,sid9_gci332161,00.html.....(8-VSB))
 - <http://www.aniret.org.mx/pdf/articulos/dtv.pdf>
 - http://www.xilinx.com/esp/prof_brdocst/collateral/dvbatpsc.pdf
 - <http://www.bestbuy.com/HomeAudioVideo/Televisions/hr/hr7Televisions.asp?cat=24&m=1>
 - <http://www.itu.int/ITU-D/tech/digital-broadcasting/kiev/Presentations/graves/ATSCpart1a.pdf>
-

Bibliografía

- http://support.xilinx.co.jp/esp/prof_brdcst/collateral/settopbox.pdf
- <http://www.cabletoday.com/ct2/archives/0500/0500feat10.htm>
- http://www.sonypro.com.mx/info_tec/pdfdoc/mpegaplicacion.pdf
- <http://www.primedtv.com/index.html>
- http://www.atsc.org/faq/faq_general.html
- http://www.atsc.org/faq/faq_atscrelated.html
- http://www.atsc.org/faq/faq_psisip.html
- http://www.atsc.org/document_map/receivers.htm
- http://www.4i2i.com/reed_solomon_codes.htm
- <http://www.canal9-mendoza.com.ar/TELEVISIONDIGITAL.htm>
- <http://www.servisystem.com.ar/ATSC/dtv8.html>
- http://cva.stanford.edu/ee482s/projects/group2_proposal.pdf
- <http://www.com.uvigo.es/asignaturas/scvs/trabajos/curso0203/dvb-s/DVB-S.pdf>
- www.tektronix.com/video_audio/
- <http://www.iso.org/iso/en/aboutiso/introduction/index.html>
- <http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/>
- http://www.fuac.edu.co/autonoma/pregrado/ingenieria/ingelec/proyectosgrado/compresvideo/compression_audio.htm
- <http://www.dolby.com/tech/desgnac3.pdf>
- <http://www.dolby.com/tech/ac3flex.html>
- <http://www.dolby.com/tech/ac-3mult.html>
- <http://www.dolby.com/tech/L.mn.0002.5.1Guide.s.pdf>
- http://www.atsc.org/standards/a_52a.pdf

- http://www.upv.es/satellite/trabajos/Grupo8_99.00/mpeg.html#Errores
- <http://personales.mundivia.es/jtoledo/angel/error/error32.htm>
- <http://www.upcnet.es/~erm10/dvbs/dvbs.htm>
- <http://www.rares.com.ar/PDF/1007.pdf>
- <http://dis.eafit.edu.co/cursos/st780/material/medios/fund-compresion.pdf>
- <http://trabajospracticos.4mg.com/cod/secc2.htm>