

2/
203.



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE INGENIERIA

"ANALISIS Y EVALUACION DE LAS NUEVAS
TECNOLOGIAS DE TRANSMISION DE
VIDEO DIGITAL"

263549

T E S I S

QUE PARA OBTENER EL TITULO DE:
INGENIERO MECANICO ELECTRICISTA

AREA ELECTRICA-ELECTRONICA

P R E S E N T A N ;

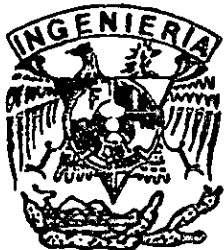
JUANA LETICIA ALVARADO HERNANDEZ

FERNANDO BECERRIL ORTA

JUAN EFRAIN JAIME LOA

EDUARDO PLAZA LOPEZ

JOSE GUADALUPE RUIZ CRUZ



DIRECTOR DE TESIS: DR. VICTOR GARCIA GARDUÑO

CIUDAD UNIVERSITARIA.

1998.



Universidad Nacional
Autónoma de México



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.

AGRADECIMIENTOS

*Gracias a ti, **Pepe**, mi esposo, compañero y amigo por apoyarme y ayudarme a salir adelante, porque sin tu fortaleza e integridad no habría sido posible terminar este trabajo que culmina con uno de nuestros más grandes anhelos desde que nos conocimos.*

***José**, te agradezco que me enseñaras a ser inmensamente feliz y también, que me enseñaras que no es tan fácil conservar esa felicidad aunque se tenga todo para lograrlo, a conocer la vida como realmente es con tristezas, desengaños, amargura y soledad, pero también con grandes alegrías y recompensas como lo es nuestro hijo, gracias.*

*Para ti **Pepito**, mi chiquito querubín, gracias por tu amor y cariño incondicional que me ha hecho sentir el más grande amor que Dios me podría haber regalado, amor que me ha impulsado a terminar este trabajo y salir adelante a buscar nuevos horizontes, porque el hecho de tenerte a mi lado todo lo borra y sé que Dios está entre nosotros brindándonos su amor.*

Mamá y Papá, Gracias a ustedes este trabajo lo puedo presentar, porque es parte de su esfuerzo, cariño y dedicación hacia sus hijos.

Mamá Conchita, para ti todo mi amor y agradecimiento eterno, por apoyarme y ayudarme a salir adelante cuando más lo he necesitado, porque gracias a tú inmenso amor de madre no he desfallecido, porque a una madre no se le puede ocultar jamás cuando un hijo está sufriendo o cuando es feliz y tú siempre has estado a mi lado brindándome todo tu amor y apoyo sin cuestionar jamás nada, además me has enseñado que el amor que se le tiene a los hijos es tan grande que nos señala el camino siempre hacia adelante, cuídate mucho porque tú eres nuestro apoyo, fortaleza y ejemplo.

Papá, gracias por brindarme el apoyo y la confianza para realizar uno de los anhelos más importantes de mi vida, porque afortunadamente podré buscar otras alternativas en mi vida gracias a esta oportunidad que tú me otorgaste, gracias por haber creído en mí y quererme como lo has hecho.

Bertha, gracias por tu invaluable apoyo como hermana y amiga, ya que me has prestado tu ayuda y compartido conmigo buenos y malos momentos, como tu hermana mayor espero que este trabajo te haga reflexionar y sigas el ejemplo para poder brindar a Saulito tu pequeño hijo, nuevas y mejores oportunidades.

Saúl, pequeño sobrino, casi mi hijo, quien me enseñó con sus primeros pasos por la vida a luchar por lo que se quiere y a dar amor sin reservas, te agradezco que con tus pequeños detalles me hicieras reflexionar sobre que significa tener un hijo, y ahora que **Pepito** es una hermosa realidad espero aprender poco a poco a través de ustedes dos a amar y disfrutar cada momento de la vida tratando de hacer felices a los que me rodean.

Beto, te agradezco que me mostrarás que el pequeño hermano, también es capaz de enseñar, apoyar y comprender cuando lo necesitamos, nuevamente gracias y espero que Dios con su infinita sabiduría te guíe para que logres salir adelante con esos dos pequeñitos que te presto para que los eduques dentro de un ambiente de amor y seguridad.

Deseo dar las gracias a la Universidad casa donde me he formado como profesionista y que me mostró el camino con bases sólidas de valores e integridad universitaria, a todos mis compañeros de estudios con quienes compartimos muchísimos buenos y agradables momentos, otros también difíciles y algunos tristes pero eso sirvió para afianzar más nuestra amistad, a mis profesores y ahora también amigos quisiera brindarles mi agradecimiento y reconocimiento por transmitirme y compartir sus conocimientos y experiencia para poder buscar un lugar dentro del ambiente profesional, un especial reconocimiento al Ing. Esteban Salinas Elorriaga por haberme brindado su amistad y mostrarme que la carrera de ingeniería era realmente mi vocación.

J. Leticia Alvarado Hernández

AGRADECIMIENTOS

A Leticia mi esposa y amiga, mil gracias por apoyarme y del lado mío luchar por superar día con día las adversidades. Junto contigo, puedo finalmente decir que concluye una de las etapas más determinantes de mi vida.

A mi hijo José Alberto, quien sin saberlo, con su cariño y sus pequeñas manos, me ha dado el impulso final para terminar con una de mis metas más importantes.

A mis Padres Roberto y Leonor, para quienes no existen las palabras suficientes para agradecerles su apoyo y ejemplos de valor, entereza e integridad ante la vida; y en cuyas manos puedo finalmente poner el fruto de muchos años de esfuerzo; no solamente el mío, sino el de ustedes.

A mis hermanos Carlos, Roberto, Rosa María, Angel, Adriana Lourdes, Estela, Ma. del Carmen, Luis Enrique, Alberto y Magdalena (q.e.p.d.) por su incondicional apoyo y aliento en mis empeños. Vaya una gratitud muy especial a mi abuelo Toribio (q.e.p.d.), quien siempre fue para nosotros un ejemplo de esfuerzo ante la vida.

A mis familiares y amigos, en quienes siempre encontré apoyo y palabras de ánimo. Y de manera muy particular a la Universidad Nacional Autónoma de México y a sus profesores, quienes sin egoísmos compartieron conmigo sus conocimientos e invaluable experiencia.

José Guadalupe Ruiz Cruz

AGRADECIMIENTOS:

A mis padres, Manuel Plaza León y Guadalupe López Gutiérrez por haberme brindado la vida y la oportunidad de cursar una carrera. Siempre me apoyaron en mis estudios y me impulsaron a seguir adelante.

A mis hermanos Olga Guadalupe, Sergio, Luis Alejandro y Dolores Cecilia por su interés en mis estudios y su gran calor humano, gracias al cual nuestra familia siempre se ha mantenido unida y en un ambiente propicio para el estudio.

A Beatriz Sánchez Arellano por su gran apoyo y comprensión durante toda mi carrera. Anteponiendo ante todo mis estudios. Gracias Amor. " No digas quien eres tus hechos lo dirán ".

A la Facultad de Ingeniería de la UNAM y a todos sus profesores por su tiempo y dedicación que ellos de alguna manera me dedicaron.

En especial al Dr. Víctor García Garduño por su gran dedicación en esta Tesis.

A todos ustedes Gracias.

A MIS PADRES, por haberme hecho lo que soy y llevarme de la mano hasta este momento. Finalmente hemos llegado.

A MI PAPÁ, tenacidad, amor por la vida, carácter,..... ejemplo que trascenderá en mí.

A MI MAMÁ, amor pleno y apoyo incondicional. Aquí está el resultado.

A PATY, amor y compañía HOY y POR SIEMPRE.

A LICHA, siempre mi mejor amiga. Ahora sigo tu ejemplo.

A MI FAMILIA, quien me apoyó a culminar esta gran etapa de mi vida.

A TODOS los que de una u otra forma me apoyaron incondicionalmente para lograr este éxito.

Fernando Becerril Orta

Agradecimientos:

A mis padres:

Porque ellos nos trajeron de pequeños a mis hermanos y a mí desde Jerecuaro(Guanajuato) a la ciudad de México para aprender a leer y escribir, principalmente a mi madre por su apoyo y paciencia a lo largo de mi formación profesional, "gracias mamá" por la entrega hacia tu familia que nunca ha sido lo suficientemente valorada y porque siempre has querido lo mejor para todos nosotros.

A mis hermanos:

Por el cariño que siempre me han brindado, por su ayuda y apoyo, así como también por alentarme a seguir estudiando, principalmente a mis hermanos mayores:

Esther Jaime Loa
María Jaime Loa
Reyna Jaime Loa
José Luis Jaime Loa
Rosalinda Jaime Loa
Jesús Jaime Loa
Etc...

Al Dr. Víctor García Garduño:

Por su colaboración y asesoría en la realización de este trabajo

A mis compañeros:

José, Leticia, Fernando y Eduardo, por el esfuerzo realizado para la terminación de este trabajo

ÍNDICE

PARTE I. INTRODUCCIÓN AL VÍDEO DIGITAL.

1.- Representación de video digital.	1
1.1 Video analógico.	1
1.2 Video digital.	13
1.3 Etapa de adquisición: cámaras de vídeo.	32
Bibliografía.	56
2. Técnicas de Compresión de Video	58
2.1 Compresión de imágenes fijas	58
2.1.1 Codificación predictiva	58
2.1.2 Codificación por transformada coseno discreta.	67
2.2 Compresión de Imágenes en Movimiento	69
2.2.1 Codificación DCT/DPCM.	69
2.2.2 Compresión por compensación de movimiento.	75
2.2.3 Compresión por <i>Block Matching</i>	85
Bibliografía.	93
3.- Las Normas de Compresión de Video.	95
3.1 JPEG: Imágenes fijas	95
3.2 H.261: Videoconferencia.	104
3.3 H.263: Codificación a baja tasa de transmisión.	115
3.4 MPEG: Imágenes en movimiento.	129
3.4.1 MPEG1: Almacenamiento C.D.	134

3.4.2 MPEG2: Radiodifusión	144
3.4.3 MPEG4: Aplicaciones multimedia.	148
Bibliografía.	158

PARTE II. LOS SISTEMAS DE VÍDEO DIGITAL.

1.- Televisión. Alta Definición.	160
1.1 Definición y estándar para HDTV.	162
1.2 Algunos tipos de sistemas híbridos	165
1.3 Sistema MUSE	167
1.4 Sistema HD-MAC	169
Bibliografía.	178
2.- Videoconferencia	180
Bibliografía.	202
3.- Comunicación multimedia vía N-ISDN.	204
3.1 Estándar H.320: Video teleconferencia.	208
3.2 H.221: Teleservicios audiovisuales	213
3.3 H.242: Establecimiento de comunicación entre termi- nales audiovisuales por canales digitales de hasta 2Mbps.	218
3.4 H.230: Señales de control en indicación para sistemas audiovisuales.	224
Bibliografía.	231
4.- Comunicaciones multimedia vía PSTN y Radio Móvil.	233
4.1 H.223: Protocolo de multiplexación para comunica- ciones multimedios a baja velocidad binaria	238
4.2 H.245: Protocolo de control para comunicaciones	

	multimedios.	244
4.3	H.324: Transmisión en línea de señales no telefónicas.	258
4.4	Hacia los sistemas del futuro.	272
	Bibliografía.	289
5.-	Transmisión multimedia vía B-ISDN y LAN.	290
5.1	H.321: Adaptación de las terminales videotelefónicas H.320 a los entornos de RDSI-BE.	290
5.2	H.310: Sistemas y terminales para comunicaciones audiovisuales de banda ancha.	299
5.3	H.322: Sistemas y equipos terminales videotelefónicas para las redes de área local que dan un servicio con calidad garantizada.	312
5.4	H.323: Sistemas y equipos videotelefónicos para redes de área local que proporcionan una calidad de servicio no garantizada.	316
	Bibliografía.	346
6.-	Codificación de vídeo con VLSI.	347
6.1	De algoritmos a arquitecturas VLSI.	348
6.2	Estimación de arquitecturas alternativas.	354
6.3	Funciones específicas de las arquitecturas.	355
6.4	Sistema multiprocesador programable.	360
PARTE III. CONCLUSIONES.		
	Conclusiones.	366

INTRODUCCIÓN

Los sistemas y servicios de telecomunicación se han multiplicado de manera espectacular en los últimos años, de manera que actualmente nadie puede concebir la existencia de una empresa sin la disponibilidad de estos servicios; su evolución no sólo será el fruto de lo que pueda ofrecer la tecnología, sino también, y de forma determinante, de los usuarios y de lo que demande y admita el mercado. Atendiendo el punto de vista de los usuarios y del mercado, aún se encuentra muy lejos el objetivo final de cualquier operador de servicios de telecomunicación que consiste en ser *“capaz de facilitar la comunicación entre las personas y las máquinas de forma fácil y segura desde cualquier lugar, en cualquier momento, y con cualquier medio o combinación de medios”* (voz, datos, textos e imágenes) a un costo razonable.

La vida cotidiana se encuentra fuertemente asociada a dos servicios de telecomunicación, que son el paradigma de este conjunto de servicios, el teléfono y la televisión.

La última década, ha sido un período de tiempo donde el desarrollo e interés por las imágenes han experimentado un notable crecimiento que ha afectado a todos los sectores sociales y profesionales. Para el futuro se prevé la existencia de infraestructuras únicas, capaces de soportar todo tipo de servicios desde la telefonía hasta la distribución de televisión de alta definición por cable, de forma que el usuario se pueda conectar a todo tipo de terminal, para poder utilizar cualquier tipo de servicios. El vídeo juega un papel muy importante dentro de estos servicios, papel que no ha sido aún muy explotado y desarrollado, pero que debido al notable crecimiento del interés y la necesidad que ha despertado en el ámbito profesional y general, es por lo que se presenta en los capítulos siguientes de la tesis.

La presente tesis desarrollada en forma conjunta con el P.A.T. (Programa de Apoyo a la Titulación), está dividida en tres partes. En un primer plano se presentan los aspectos teóricos básicos del vídeo, es decir los fundamentos sobre los cuales descansa el resto del documento, o dicho de otra manera, los aspectos teóricos acerca de la adquisición de imágenes y se analizará el concepto básico de lo que hoy se conoce como vídeo digital, revisando de manera sencilla los conceptos de muestreo y velocidad de muestreo de las señales digitales, así como los métodos de codificación para la transmisión de tales señales; y en una segunda parte se revisarán algunas de las técnicas y normas de compresión de imágenes fijas y en movimiento.

En la segunda parte de este documento, se analizarán de manera breve algunas de las normas de transmisión de vídeo en diversos

medio, tales como: redes ISDN, de banda ancha, de banda angosta, redes PSNT, TVHD, etc., con la finalidad de mostrar el estado del arte en lo que se refiere a protocolos y medios de transmisión de vídeo digital.

Y por último, en la tercera sección se presentan algunas perspectivas del presente trabajo.

PARTE I

INTRODUCCIÓN AL VÍDEO DIGITAL

CAPÍTULO 1

REPRESENTACIÓN DE VÍDEO DIGITAL

1.1 VÍDEO ANALÓGICO

Los primeros métodos utilizados en la transmisión y grabación televisiva fueron analógicos, y los formatos de las señales estaban determinados por los requisitos del tubo de rayos catódicos (TRC) como elemento de visualización. En un sistema analógico, la información se transmite mediante alguna variación infinita de un parámetro continuo como puede ser el voltaje en un cable o la intensidad de flujo en una cinta.

Las señales de video son formas de onda eléctricas bidimensionales (una onda eléctrica es bidimensional en el sentido que lleva un voltaje que varía con respecto al tiempo) que permiten transmitir imágenes en movimiento de un lugar a otro. Con el fin de transmitir información de una imagen tridimensional a través de un cable bidimensional, es necesario recurrir a un proceso de exploración. En lugar de intentar

transmitir el brillo de todas las partes de una imagen a la vez, la exploración transmite el brillo de un sólo punto que se desplaza con el tiempo. [Watkinson, 1994]

En la transmisión de la imagen, el tubo de cámara convierte la imagen óptica en una señal de vídeo. El tubo de cámara es un TRC con una placa de imagen fotoeléctrica y un cañón electrónico confinado por una envoltura de vidrio en la cual se ha practicado el vacío (actualmente se utiliza un circuito integrado denominado *CCD Charge Coupled Device*) que sustituye con ventaja al tubo de cámara, y el tubo de imagen es muy parecido al TRC utilizado en un osciloscopio, la placa frontal de vidrio tiene un revestimiento fluorescente en su superficie interior, el cuello estrecho contiene el cañón electrónico, cuando el haz de electrones incide en la pantalla de fósforo, ésta emite luz.

Exploración horizontal y vertical

La exploración consiste en barridos horizontales rápidos combinados con barridos verticales más lentos, de manera que la imagen queda explorada en líneas. Al final de cada barrido vertical, o cuadro, el proceso vuelve a repetirse.

Como resultado de la exploración vertical, todas las líneas son ligeramente inclinadas hacia abajo. Los monitores para ordenadores exploran de esta manera, pero en la mayoría de los sistemas de radiodifusión el proceso de exploración consiste en un entrelazado 2:1. En el procedimiento de exploración que ha sido universalmente adoptado se emplea la exploración horizontal de líneas pares e impares entrelazadas. Las especificaciones de la exploración de la Comisión Federal de Comunicaciones (FCC) en Estados Unidos para difusión de televisión dan un formato estándar de la exploración que incluye un total de 525 líneas horizontales en una trama o cuadro

rectangular cuya relación de aspecto es 4:3 (4 de ancho por 3 de alto). Las tramas se repiten al ritmo de 30 por segundo con dos campos entrelazados en cada trama. Primero se exploran todas las líneas impares, punto por punto de izquierda a derecha y línea por línea desde la parte superior a la inferior, omitiendo las líneas pares. Después de este ciclo de exploración vertical, una rápida retraza vertical hace que el haz electrónico de exploración vuelva a la parte superior del cuadro. Luego son exploradas punto por punto de izquierda a derecha y línea por línea desde la parte superior a la inferior todas las líneas pares, que fueron omitidas en la primera exploración. El tiempo de retorno es lo que se llama retraza o retorno, durante las retrazas horizontal y vertical no es explorada ninguna información de imagen. Cada grupo de líneas pares o de líneas impares es lo que se llama un campo. Este proceso de entrelazado se muestra en la figura 1.1.1.

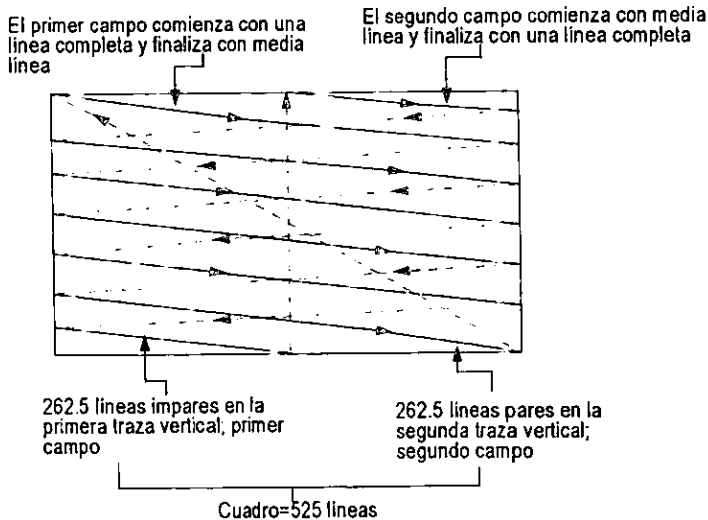


Figura 1.1.1 Entrelazado 2:1

Frecuencia de exploración vertical y horizontal.

La velocidad de repetición de los campos es de 60 por segundo, ésta es la frecuencia de exploración vertical, ya que durante un periodo de cuadro de 1/30 segundos son explorados dos campos, ésta velocidad es lo suficientemente rápida para eliminar el problema de parpadeo con los altos niveles de iluminación que se producen en la pantalla del tubo de imagen y también ésta velocidad es igual a la frecuencia de línea de 60 Hz. En los países en que la frecuencia de la línea de energía es 50 Hz, la velocidad de cuadro es 25 Hz, y por consiguiente, la frecuencia de campo es 50 Hz.

El número de líneas de exploración horizontal de un campo es la mitad del total de 525 líneas de un cuadro completo, esto es 262.5 líneas horizontales por campo. El tiempo que corresponde a un campo es 1/60 segundos y cada campo contiene 262.5 líneas, el número de líneas por segundo o frecuencia horizontal es:

$$262.5 \text{ líneas/campo} \times 60 \text{ campos/segundo} = 15,750 \text{ líneas/segundo}$$

El tiempo durante el cual se realiza la exploración de una línea horizontal incluyendo traza y retraza es: $1/15,750 = 63.5 \text{ ms}$.

Si la frecuencia de cuadro se mantiene constante y si hubiese más líneas, el tiempo de exploración sería más corto, y las frecuencias de vídeo serían más altas ($f=1/T$). En el sistema de 525 líneas, la frecuencia más alta de vídeo está limitada a 4 Mhz aproximadamente, a causa de la restricción de 6 Mhz para los canales de difusión comercial de televisión.

Sincronización horizontal y vertical.

El procedimiento de exploración requiere el uso de impulsos de sincronismo, para ello se transmiten con la información de imagen impulsos rectangulares que se utilizan para controlar la exploración en la cámara y en el receptor respectivamente. Se transmite al final de cada línea (al final de la traza o principio de la retraza) un impulso rectangular de sincronismo horizontal, por cada línea horizontal a fin de conservar sincronizada la exploración horizontal, y al final de cada campo (comienzo de la retraza vertical) se transmite un impulso rectangular de sincronismo vertical para sincronizar el movimiento de exploración vertical, estos impulsos de sincronización se transmiten como parte integrante de la señal de vídeo, pero son enviados durante los periodos de borrado en que no se transmite información alguna. Cada impulso de sincronización horizontal es insertado en la señal de vídeo dentro del tiempo del impulso de borrado horizontal. También es insertado cada impulso de sincronización vertical en la señal de vídeo dentro del tiempo del impulso de borrado vertical.

En resumen, un impulso de borrado comienza poniendo la señal de vídeo en el nivel de negro; luego ocurre la señal de sincronización que inicia la retraza en la exploración. Esta secuencia se aplica a las retrazas horizontal y vertical. Los impulsos de sincronización y borrado tienen siempre la misma frecuencia que la exploración.

Empezando a la izquierda en la figura 1.1.2, las dos últimas líneas de exploración horizontal de la parte inferior de la trama están representadas con los impulsos necesarios de sincronismo y borrado. Inmediatamente después de la última línea visible, la señal de vídeo alcanza el nivel de negro a consecuencia del impulso de borrado vertical como preparación para la retraza vertical. El periodo de borrado vertical comienza con un grupo de seis impulsos igualadores o

ecualizadores que están espaciados a intervalos de media línea. Con dos impulsos en el tiempo de una línea completa horizontal, los impulsos igualadores se repiten a razón de $2 \times 15,750 = 31,500$ Hz. Los grupos de impulsos igualadores inmediatamente antes y después del sincronismo vertical se utilizan para mejorar la sincronización vertical en los campos pares e impares para obtener un buen entrelazado. Luego sigue el impulso de sincronismo vertical almacenado o fraccionado que realmente produce el retorno vertical en los circuitos de exploración, las hendiduras se producen a intervalos de media línea, también tienen una frecuencia de 31,500 Hz. Realmente, los impulsos igualadores y los almacenados verticales son impulsos similares pero de polaridades opuestas. A continuación del impulso de sincronismo vertical hay otro grupo de seis impulsos igualadores y un tren de impulsos horizontales. El tiempo de borrado horizontal, es del 16% del tiempo de exploración horizontal de una línea y el tiempo de borrado vertical es del 8% del tiempo de exploración vertical de un campo. [Grob, 1990]

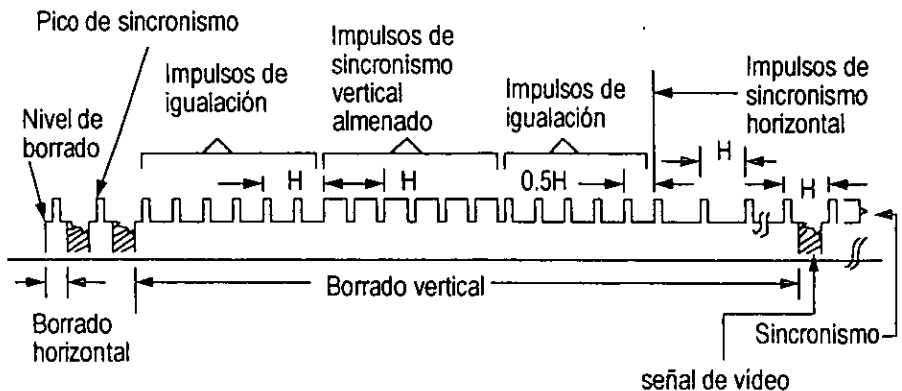


Figura 1.1.2 Detalle de los impulsos de sincronismo y de borrado de campo sucesivos en la exploración vertical. Obsérvese que todos tienen la misma amplitud, pero su anchura o su forma de onda es diferente. El tiempo de H es $1/15,750$ segundos.

La figura 1.1.3 Muestra algunos de los tipos básicos de video analógico de color. Debido a que las cámaras de color prácticas normalmente tienen tres sensores individuales, uno para cada color primario, habrá un sistema RGB (Red, Green, Blue o Rojo, Verde, Azul) en alguna de las etapas de los mecanismos internos de la cámara.

En televisión de circuito cerrado las señales de rojo, verde y azul son la única información necesaria para reproducir la imagen, también se emplean en los monitores para diseño asistido por ordenador (CAD). Las señales vídeo R, G y B no son compatibles para los receptores en blanco y negro, porque cada una sólo contiene una parte de la información de la imagen. Son todas necesarias, pero habría que utilizar demasiado ancho de banda para las tres señales separadas.

Por tanto, las señales video de color se codifican, combinándolas en proporciones determinadas para proveer la misma información de video en forma diferente. El resultado de la codificación es la formación de dos señales separadas: la señal de crominancia, o croma C, para el color y la señal de luminancia, o brillo Y, para la información de blanco y negro.

En el receptor, el tubo de imagen en color necesitará también las señales vídeo R, G y B que corresponde a los fósforos de color de la pantalla, la señal C es decodificada por demodulación. Luego la salida detectada es combinada con la señal de luminancia para recuperar las señales de video originales de rojo, verde y azul para el tubo de imagen en color. Todos los colores pueden ser producidos como mezcla de rojo, verde y azul. [Watkinson, 1994]

Ahora se considerará más detalladamente cómo se produce la señal en el transmisor.

En el transmisor la matriz combina las tensiones R, G y B en proporciones determinadas para formar tres señales video que son las elegidas para la difusión. Una señal contiene la información de brillo Y (compatible con el monocromo), las otras dos señales contienen el color (R-Y y B-Y).

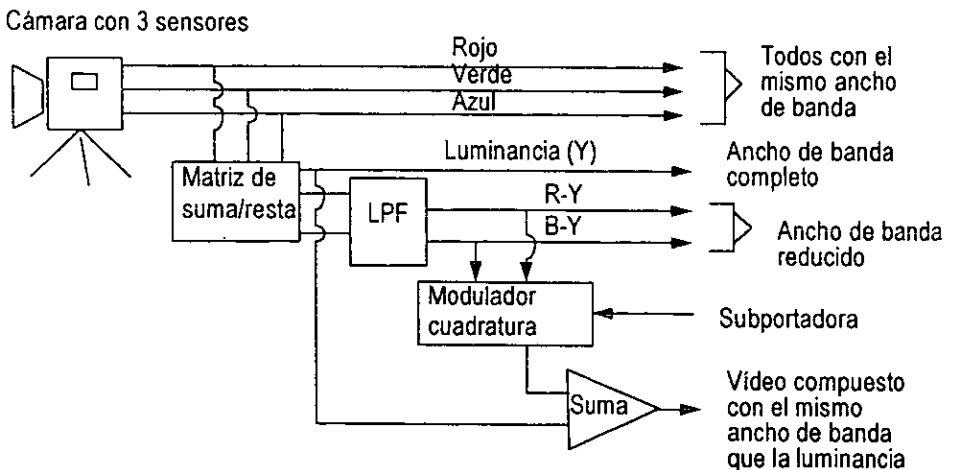


Figura 1.1.3 Las señales roja, verde y azul salen de la cámara, siendo necesario un ancho de banda completo. Si se obtiene una señal de luminancia mediante una suma ponderada de R, G y B, necesitará el ancho de banda total (normalmente de 0 a 4 Mhz), pero las señales de diferencia de color R-Y y B-Y necesitan menos ancho de banda (de 0 a 0.5 Mhz). La combinación de R-Y y B-Y en un sistema de modulación de subportadora permite la transmisión de color en el mismo ancho de banda que el monocromo.

Es posible ahorrar ancho de banda si se trabaja con diferencia de color. El ojo humano se basa en el brillo para obtener todos los detalles, y se necesita mucha menos resolución en la información de

color, ya que no es igual de sensible a los tres colores primarios; la señal de luminancia es una suma ponderada constituida de:

$Y=30\%$ de vídeo R+ 59% de vídeo G+ 11% de vídeo B.

Las señales R-Y y B-Y se usan para modular la frecuencia subportadora de 3.58 Mhz. De esta manera la información de color es transmitida como modulación en una subportadora, generalmente se denomina señal C para crominancia o croma.

El valor de 3.58 Mhz se elige como una frecuencia de vídeo alta, para separar la señal de crominancia con respecto a las videofrecuencias más bajas de la señal de luminancia. Además, la frecuencia alta da por resultado una baja visibilidad de cualquier interferencia de croma en la señal de luminancia, y en el extremo opuesto, la frecuencia de la señal C no puede estar demasiado cerca de 4.5 Mhz, a fin de evitar las interferencias con la señal de sonido.

Para la emisión televisiva de color compatible con monocromo en un solo canal, la subportadora que lleva dos señales de diferencia de color de banda restringido ha de ser invisible en la pantalla de televisión monocromático. A un sistema de color basado en subportadora se le suele denominar como vídeo compuesto.

El uso de solamente las bandas laterales de modulación, sin la propia señal portadora, es lo que se llama "transmisión con portadora suprimida". La finalidad de suprimir la señal subportadora es reducir la interferencia en 3.58 Mhz que puede producir puntos finos en la pantalla. Con transmisión de portadora suprimida, el receptor debe tener un circuito oscilador de 3.58 Mhz que genere la señal subportadora, con el fin de detectar la crominancia, se transmite una

muestra que consiste en un "burst" (tren o ráfaga) de sincronismo de color de 8 a 11 ciclos de la señal subportadora de 3.58 Mhz en el umbral posterior de cada impulso de borrado horizontal, su objetivo es sincronizar el oscilador de color de 3.58 Mhz del receptor.

La figura 1.1.4 muestra los detalles del burst de sincronismo de color de 3.58 Mhz transmitido como parte de la señal de video compuesto total. El burst y la señal C son ambos de 3.58 Mhz, pero el burst sólo esta presente durante el tiempo de borrado en el cual no hay información de imagen. La señal C está presente durante el tiempo de traza visible de la información de color en la imagen. La presencia o ausencia del burst determina de qué manera reconoce el receptor de color si un programa es de color o monocromático.

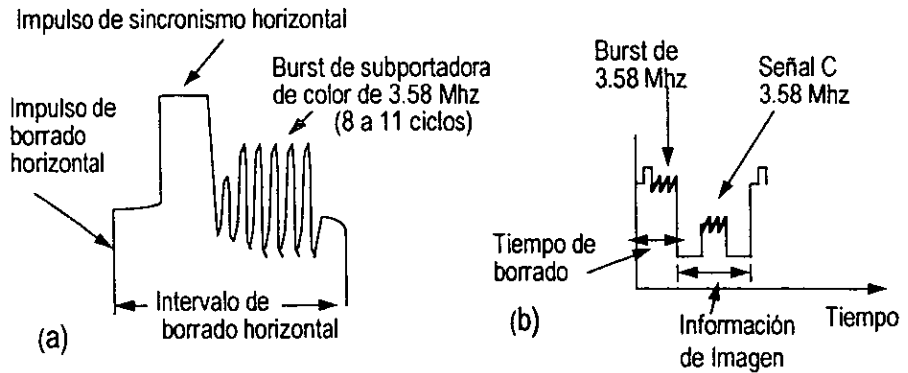


Fig. 1.1.4 (a) *Burst de sincronismo de color en el umbral posterior de cada impulso de sincronismo horizontal.* (b) *Comparación del burst y de la señal C, ambos de 3.58 Mhz pero en instantes diferentes.*

La señal C con la información de color y la señal de luminancia Y son ambas acopladas a la sección del sumador. Esta etapa combina la señal Y con la señal C de 3.58 Mhz para formar la señal de video total.

Esta señal es transmitida al receptor por modulación de amplitud de la onda portadora de imagen, en el canal de 6 Mhz asignado a la estación.

En la difusión de televisión, se utiliza modulación de amplitud (AM) para la señal de imagen, y modulación de frecuencia (FM) para la señal de sonido asociado. La AM es mejor para la señal de imagen a causa de que las imágenes fantasmas resultantes de la recepción por diversos caminos son menos perceptibles. Con AM los fantasmas se mantienen inmóviles, pero con FM fluctúan en la imagen.

Las dos características principales de la señal portadora de imagen transmitida son:

1. Polaridad negativa de modulación. Entre las ventajas tiene que se necesita menos potencia para la transmisión, y que se tiene el pico de sincronismo como referencia para la intensidad de la portadora.
2. Transmisión de banda lateral residual. La señal en AM de la imagen no se transmite como señal de doble banda lateral normal. Una parte de la banda lateral inferior se suprime por filtrado y se mantiene un residuo de bandas laterales. El propósito es reducir el ancho de banda, específicamente, se utiliza un canal de difusión de televisión de 6 Mhz en lugar de 8 Mhz, o más, que serían necesarios para dobles bandas laterales con modulación de la banda base de video (de 0 Hz para corriente continua hasta 4 Mhz).

Para la señal de sonido en TV se usa FM con el fin de aprovechar las ventajas de menos ruido e interferencia, con una excursión de ± 25 Khz, y se utiliza una portadora 4.5 Mhz más alta que la portadora de imagen, ambas en el canal estándar de televisión de 6 Mhz.

La figura 1.1.4 ilustra como se adaptan las señales de imagen y sonido en el canal de 6 Mhz. La frecuencia portadora de imagen no está en el centro del canal a causa de la transmisión de la banda lateral residual. Obsérvense los siguientes espaciados para la frecuencia portadora:

1. La portadora de imagen P está 1.25 Mhz por encima del extremo inferior del canal.
2. La portadora de sonido S está 4.5 Mhz por encima de la portadora de imagen, o S está 0.25 Mhz por debajo del extremo superior del canal.
3. La subportadora de color C está 3.58 Mhz por encima de la portadora de imagen, como modulación de video en la banda lateral superior. [Grob,1990]

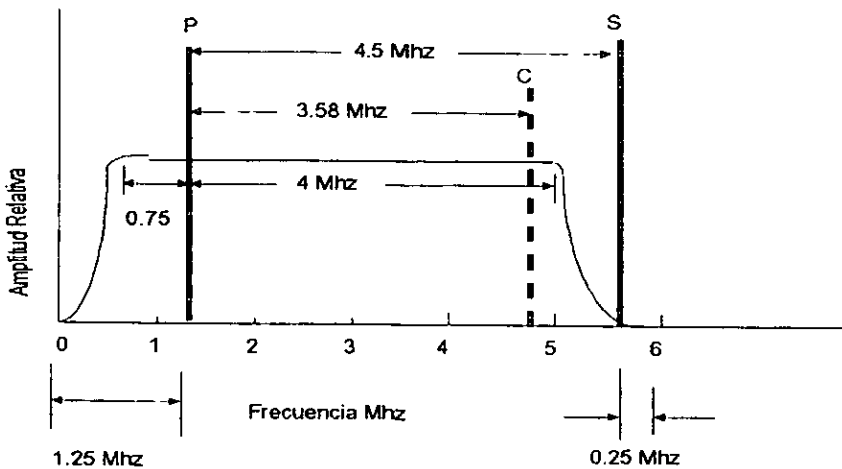


Figura 1.1.5 Separación de frecuencia, en general, para cualquier canal de 6 Mhz.

Los sistemas *PAL* (*Phase Alternate Line*); desarrollado en Alemania a partir del sistema *NTSC*, y adoptado por otros países y *SECAM* (*Système Electronique Couleur Avec Mémoire*); desarrollado en Francia y adoptado por otros países, comparten el mismo estándar de exploración de 625/50, pero los métodos para transmitir el color son totalmente incompatibles. El sistema *NTSC* (*National Television Systems Commitee*); desarrollado e introducido en los Estados Unidos de América al principio de los años cincuenta, y posteriormente adoptado en Canadá, Japón y México explora a 525/60, el sistema de color tiene unos cuantos puntos en común con el sistema *PAL*. Los tres estándares utilizan un entrelazado de 2:1.

Es característico de los sistemas analógicos el hecho de que las degradaciones no pueden ser separadas de la señal original, al final de un sistema determinado, la señal estará formada por la suma de todas las degradaciones introducidas en cada etapa por las que haya pasado, esto limita el número de etapas por las que una señal puede pasar sin que quede inutilizable. El ruido no puede detectarse en un sistema analógico, se tiene la sospecha de que existe ese ruido pero ¿cómo se puede saber qué proporción de la señal recibida corresponde al ruido y cuál es la señal original?. Si la función de transferencia de un sistema no es lineal, se producirán distorsiones, pero las formas de onda distorsionadas aún serán válidas; un sistema analógico es incapaz de detectar distorsiones. [Watkinson, 1994]

1.2 VIDEO DIGITAL

Uno de los conceptos vitales para comprender lo que es el vídeo digital; es que no es más que un medio alternativo para transportar formas de onda de vídeo. Un sistema de vídeo digital ideal, tiene las mismas características que un sistema analógico ideal: ambos son

totalmente transparentes y reproducen la forma de onda aplicada originalmente sin error. Si se comparan los sistemas tanto analógicos como digitales de alta calidad, con las mismas señales, es fácil darse cuenta que tan transparente puede ser un equipo moderno. Así como no existe en la naturaleza nada ideal, los equipos de vídeo, tanto analógicos como digitales, carecen de la misma condición de idealidad.

A pesar de que existen muchas formas en las cuales se pueden representar digitalmente formas de ondas de vídeo, existe un sistema conocido como modulación por codificación de pulsos¹, el cual es ampliamente utilizado en la actualidad. La figura 1.2.1 nos muestra como trabaja la modulación por codificación de pulsos.

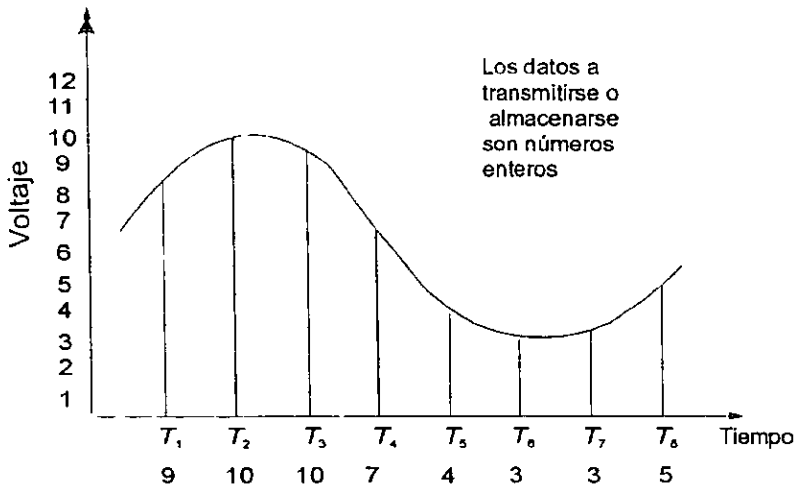


Figura 1.2.1 - Muestreo por PCM

En lugar de que el eje del tiempo sea continuo, éste se representa de manera discreta o escalonada. La forma de onda no se transporta por

¹ Pulse Code Modulation (PCM)

una representación continua, sino con mediciones que se llevan a cabo en intervalos regulares. Este proceso se llama muestreo y la frecuencia con la que se efectúa la toma de las muestras (muestreo), se denomina frecuencia de muestreo o relación de muestreo F_s . Es importante hacer notar que el muestreo es un proceso analógico. Cada muestra varía de la misma manera en que lo hace la forma de onda original. Los dispositivos de muestreo analógicos son bien conocidos en el campo del vídeo. Para completar la conversión a PCM cada muestra se representa con precisión finita por un número discreto en un proceso conocido como cuantificación.

En sistemas de televisión la imagen de entrada que llega al sensor de la cámara, será continuo en el tiempo, y continuo en dos dimensiones espaciales correspondientes a alto y ancho de sensor. En sistemas de vídeo analógico, el eje del tiempo es muestreado en cuadros y el eje vertical es muestreado en líneas. El vídeo digital simplemente agrega un tercer proceso de muestreo a lo largo de las líneas.

Existe una conexión directa entre el concepto de muestreo temporal, en donde la entrada cambia con respecto al tiempo a una determinada frecuencia y es muestreada a otra frecuencia; y el muestreo espacial, en donde una imagen cambia en un determinado número de veces por unidad distancia y es muestreada a otro determinado número de veces por unidad de distancia. La conexión entre los dos procesos es el proceso de rastreo². Multiplicando la frecuencia espacial por la velocidad del rastreo, se puede obtener una frecuencia temporal. Si se supusiera un hipotético sensor de imagen que tiene 1,000 sensores discretos a lo largo de un 1 cm, la velocidad de muestreo espacial de este sensor sería por lo tanto de uno por centímetro. Si los sensores se miden secuencialmente durante el rastreo, que toma un milisegundo,

² Del inglés SCANNING

entonces se tendrá una velocidad de muestreo de 1 MHz. En el dominio de lo espacial nos referimos a resolución, de la misma manera que en el dominio del tiempo nos referimos a respuesta en frecuencia. Los dos están ligados por el rastreo.

Aún cuando cualquier velocidad de muestreo suficientemente alta es apropiada para utilizarse con vídeo, es común establecer la velocidad de muestreo como un múltiplo de la velocidad de líneas. Así pues, las muestras serán tomadas siempre en la misma posición en cada línea. Con una velocidad de muestreo como esta, una imagen digital monocromática será un arreglo rectangular de puntos en donde la brillantez se almacena como un número. Estos puntos son conocidos como unidades de imagen³. Este arreglo generalmente se dispone de manera matricial, con filas y columnas con un determinado espacio entre los pixeles. Si estos pixeles se ubican dentro del arreglo suficientemente cerca uno del otro, entonces es posible que un observador pueda percibir la imagen como una imagen continua.

Obviamente entre mejor se defina el espaciamiento entre los pixeles, se obtendrá una mejor resolución de imagen, pero se necesitará almacenar una mayor cantidad de información en la misma proporción en que se mejore la resolución de la imagen.

Si se desea transmitir una imagen a colores, entonces cada imagen consistirá de 3 capas sobrepuestas de muestras, una por cada uno de los componentes. Así pues, el pixel ya no corresponde a la unidad mínima de representación de imagen, ya que este se substituye por un vector que describe de alguna manera la brillantez, el matiz y la saturación de cada punto de la imagen.

³ Picture Cells, generalmente abreviados como **pixeles**, aún cuando algunas veces la abreviación es mas corta y son denominados como **pels**.

En los sistemas RGB, los pixeles contienen tres números unipolares que representan la proporción de cada uno de los tres colores primarios en ese punto de la imagen.

Para producir imágenes en movimiento, lo único que se necesita es simplemente contar con un procedimiento en donde el valor de cada pixel se pueda actualizar periódicamente. Esto resulta en un arreglo tridimensional en donde dos de los ejes son espaciales y el tercero es temporal.

Los mayores esfuerzos que se realizan en materia de nuevos desarrollos de convertidores analógico-digital (ADC) están encaminados a eliminar la inestabilidad de tiempo en el reloj de muestreo, de tal manera que las muestras se tomen cada una en un intervalo de tiempo constante. Si existe un error en el tiempo base, la muestra tomada llegará con alteraciones de tiempo, y si no se controla esto adecuadamente, será fácilmente visible. Una manera de solucionar este problema de inestabilidad en el tiempo base, es almacenar temporalmente las muestras recibidas de manera irregular en memoria, para después, y posteriormente leerlas, utilizando un reloj estable generado localmente. Este proceso se llama corrección de tiempo de base y es ampliamente utilizado por prácticamente todos los sistemas de video digital. De esta manera, el error de tiempo de base no simplemente se reduce, sino que se elimina totalmente.

Existe una gran preocupación entre aquellos que no están familiarizados con principios digitales, ya que se piensa que existe algún tipo de pérdida de información, ya que no se sabe que sucede entre las tomas de las muestras. Esta apreciación sería válida en un sistema que tenga un ancho de banda infinito, pero no existe señal

analógica que pueda tener un ancho de banda infinito⁴. Cuando una señal tiene un ancho de banda finito la velocidad a la que cambia es limitada, y el modo en que lo hace, entonces puede llegar a ser predecible.

Previamente se definió que cada muestra es discreta. La longitud de la muestra, que será proporcional al voltaje de la señal de video, se representa por un número entero. A este proceso se le denomina cuantificación y el resultado es una aproximación, pero la magnitud del error se puede controlar hasta que éste es despreciable. Así como existen muchas otras mediciones que se aproximan a medidas adecuadas para el caso de señales de video, también es posible encontrar una resolución apropiada para cuantificarla, y una vez identificada, no es conveniente el incrementar la resolución identificada. La ventaja de utilizar números enteros para la cuantificación es que de alguna manera se controla que estos no se disparen demasiado. Si se puede transportar un número entero de un lugar a otro sin error numérico, éste no sufre de ningún cambio. Es por ésto que al representar numéricamente las formas de onda de video, se logra que la información original tal y como se expresó, sea mucho más resistente a alteraciones indeseables.

Esencialmente, el video digital transporta la forma de onda original de manera numérica. El número de muestras es una representación del tiempo, que a su vez corresponde a una representación de posición en la pantalla, y el tamaño de la muestra es (en casos de luminiscencia), una representación de la brillantez de un determinado punto de la imagen. De hecho la sucesión de muestras en un sistema digital corresponde a la representación de la forma de onda original.

⁴ Todas las fuentes de señales analógicas, como las cámaras, tienen un límite de resolución o respuesta de frecuencia, como de hecho lo tiene la misma visión humana.

Como ambos ejes (número de muestras y tamaño de la muestra) de la forma de onda representada digitalmente son discretos, la forma de onda se puede recuperar de una manera muy precisa. Si se necesitara de mayor precisión, es evidente que se necesitaría de un mayor número de muestras.

En términos sencillos, una forma de onda de vídeo es grabada en una grabadora digital como si se hubiera medido a intervalos regulares el voltaje con un medidor digital y las lecturas se hubieran escrito en un rollo de papel. La velocidad a la que se tomaron las mediciones y la precisión del medidor son los únicos factores que pueden determinar la calidad, ya que una vez que un parámetro se expresa como un número discreto, se pueden transmitir series de dichos números sin ningún cambio. La calidad es determinada, solamente, por la precisión de la conversión y es independiente de la calidad de la trayectoria de la señal.

Como se mencionó previamente, uno de los grandes retos en tecnología de vídeo digital, es el evitar al máximo la posibilidad de error, y una de las maneras más eficientes de lograr este propósito es utilizar, para la representación de las señales de vídeo, un sistema numérico que nos permita minimizar la posibilidad de alteración de la información. Para esto, en la transmisión de vídeo digital, se utiliza el sistema binario, ya que, al tener solamente dos dígitos (0 y 1), existe muy poca posibilidad de error además de que tiene la gran ventaja de que se pueden transmitir eficientemente, ya que los dos dígitos de este sistema se representan con estados de "encendido" y "apagado"⁵.

⁵ A estos dos dígitos binarios (Binary digITS) se les denomina como BITS.

En el sistema decimal, los dígitos de un número (contando a partir de la derecha o el número menos significativo) representan unidades, decenas, centenas, etc. En el sistema binario los bits representan 1, 2, 4, 8, 16, etc. Un número binario con múltiples dígitos es comúnmente llamado palabra, y el número de bits de la palabra es denominado longitud de palabra. Una palabra de 8 bits se llama byte que es una contracción de "by eight". Las capacidades de memorias y medios de almacenamiento se miden en bytes. Las direcciones de memoria son por sí mismos números binarios y la longitud de palabra determina el número de direcciones disponibles. Este número se determina elevando el número 2 a la potencia del número de la palabra. Así pues una palabra de 4 bytes tiene 16 combinaciones y puede tener hasta 16 locaciones de memoria disponible. Una palabra de 10 bits tiene 1024 combinaciones posibles, que es cercano al 1000. En terminología digital, 1K = 1,024 de tal manera que un kilobyte de memoria contiene 1,024 bytes.

En un sistema de vídeo digital, el número entero, que representa el tamaño de la muestra, se expresa en binario. Las señales enviadas tienen dos estados y cambian a tiempos predeterminados de acuerdo con un reloj preestablecido. Si el ruido degrada una señal binaria, entonces el receptor rechazará la señal, ya que éste evalúa la señal, únicamente en función de si está por arriba o por abajo del umbral. Este proceso es denominado como *slicing*.

La señal se puede transportar por un canal con un ancho de banda finito y esto limita la tasa de cambio de la señal. El ruido añadido a la señal de transferencia (*sloping*) puede cambiar el tiempo en el cual el *slicer* evalúa que el nivel pasó a través del umbral. Este efecto también se puede eliminar retemporizando la salida del *slicer*. A pesar de las múltiples etapas a través de las cuales pasa la señal digital binaria,

ésta "saldrá" siendo la misma señal, pero solamente después de pasar por todas estas etapas.

Las muestras de video que son representadas por números enteros se pueden transportar de un lugar a otro por medio de este esquema, y si el número se recibe correctamente, la pérdida "en el camino" será nula.

Existen dos métodos en los que se pueden utilizar señales binarias para transportar muestras: cuando cada uno de los dígitos de un número binario se transporta en un cable independiente se denomina como transmisión en paralelo. Usar cables en paralelo puede resultar incómodo, particularmente cuando se utiliza una longitud de palabra grande, en este caso, se puede utilizar un cable monofilar en donde todos los dígitos de cada muestra se envían consecutivamente de manera serial. Esta es la definición de modulación por codificación de pulsos (PCM). Utilizando este método, queda claro que la frecuencia del reloj ahora tiene que ser mayor que la velocidad de muestreo. Mientras que la transmisión de video utilizando este tipo de esquemas tiene las ventajas de que se puede eliminar el ruido y el error de tiempo de base; pero tiene la desventaja de que para una señal de alta calidad con tonalidades de color se requieren de aproximadamente 200 millones de bits por segundo.

Una vez revisadas las ventajas que ofrece la transmisión, grabación y reproducción de video utilizando un sistema como el binario, es necesario establecer las principales razones por las que se está utilizando el video digital cada vez con mayor frecuencia. Por supuesto cada actividad tiene su punto de vista sobre las ventajas de grabar, transmitir y reproducir video digital, pero las principales podrían ser básicamente dos: la calidad de la reproducción de un sistema de video digital bien diseñado es independiente del medio de transmisión

dependiendo solamente de la calidad del proceso de conversión, y el hecho de que la conversión de vídeo al dominio de permite grandes ventajas que no son posibles con señales analógicas.

También tiene la ventaja de que cuando una grabación digital se duplica, la información del original aparece idéntica en la copia: realmente no es una copia, podemos denominarla como un clon. Las grabaciones digitales se pueden copiar de manera prácticamente indefinida sin pérdida de calidad.

Al ser grabada la información de manera digital, ésta es almacenada como números discretos, lo que nos permite empacar dicha información sin pérdida de calidad. Por supuesto, es evidente que habrá algunos bits que se reciban con error, pero a pesar de estos errores debidos a ruido o pérdida de información, el sistema de corrección de errores puede recuperar la información original sin ningún problema. Las grabaciones digitales necesitan muchos menos espacio que una grabación analógica con la misma calidad o aún con una calidad superior.

El hecho de que el vídeo digital sea un tema amplio, no necesariamente tiene que ser un tema difícil. Cada proceso se puede dividir en procesos menores, cada uno de los cuales es relativamente fácil de seguir. La principal dificultad del estudio es apreciar en donde embonan cada uno de estos pequeños pasos dentro de todo el esquema. La figura 1.2.2 nos muestra un sistema de vídeo digital mínimo.

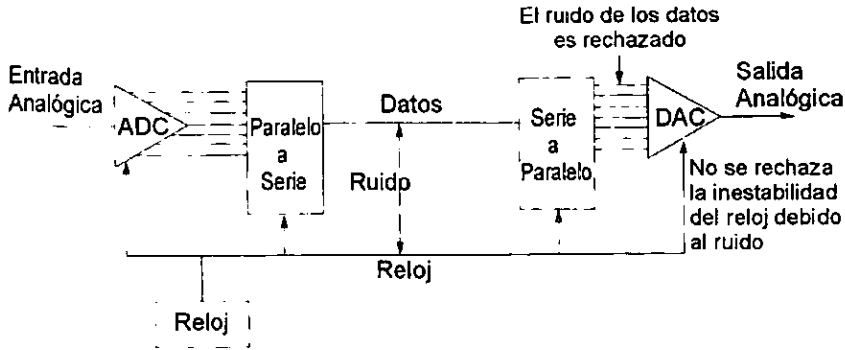


Figura 1.2.2. – Sistema de vídeo digital mínimo.

Este sistema básico, no es más que un enlace punto a punto que transmite vídeo analógico. Este consiste en un convertidor analógico-digital y un dispositivo que permite la conversión de información en paralelo a información en serie para en un extremo, y en el otro extremo un dispositivo que convierte de información en serie a información en paralelo, así como un convertidor digital-analógico. Este sistema, tal como esta expresado se puede considerar ideal, pero en un sistema real existen, como vimos anteriormente, muchos factores, tales como la inestabilidad del reloj y el ruido, que pueden hacer que la información que se recibe no sea la correcta y que el deterioro de la imagen sea perceptible.

El problema de la inestabilidad se corrige con la inclusión de un PLL⁶ como se muestra en la figura 1.2.3.

⁶ Phase Locked Loop

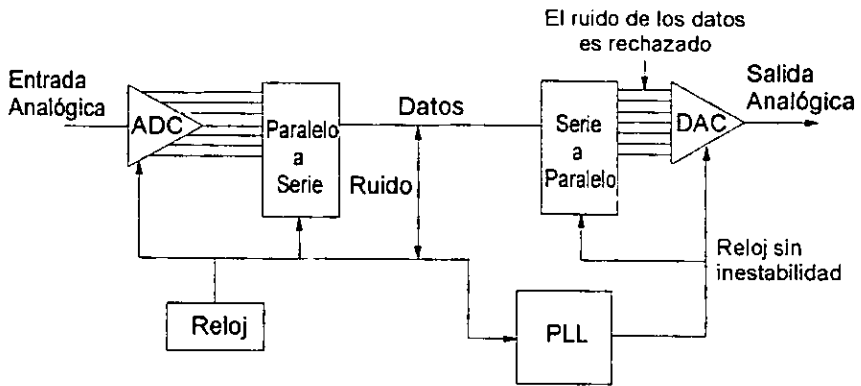


Figura 1.2.3 – Sistema de video digital con corrección de inestabilidad.

El PLL es un oscilador que se sincroniza a sí mismo a la frecuencia promedio de reloj, y que filtra todas aquellas inestabilidades instantáneas del reloj. Con este ajuste en el sistema, se logra que las muestras se alimenten al convertidor con un intervalo regular, con lo que el deterioro se reduce considerablemente.

Existe una variación del sistema anterior como se muestra en la figura 1.2.4, en el que se adicionan algunas memorias de acceso aleatorio (RAM).

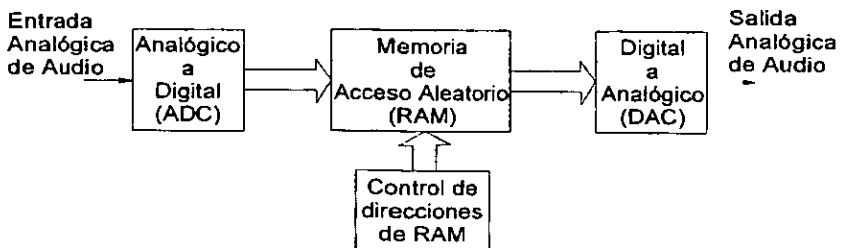


Figura 1.2.4 – Sistema de video digital con memoria.

La operación de este dispositivo se determina por la manera en la cual se pueden controlar las direcciones de memoria. Si la dirección de memoria se incrementa en uno cada vez que se toma una muestra del convertidor analógico - digital, ésta se almacena en la RAM, con ello se puede realizar una grabación por un corto periodo de tiempo hasta que ésta memoria se llene. La grabación se puede reproducir repitiendo la secuencia de las direcciones con la misma velocidad del reloj, pero ahora leyendo la información de la memoria hacia el convertidor digital a analógico. Este procedimiento se denomina como almacenamiento de cuadros. Si se incrementa la capacidad de la memoria el dispositivo se puede utilizar como un sistema de grabación. A una velocidad de 200 millones de bits por segundo cada cuadro necesita un megabyte de memoria, de aquí que éste tipo de dispositivos se restringen a un muy corto periodo de grabación/reproducción. Una posibilidad para mejorar el tiempo de reproducción es utilizar la reducción de datos y en algunos casos, cuando las imágenes corresponden a patrones predefinidos, se puede ocupar una memoria solamente de lectura (ROM) en lugar de la RAM.

Si se utiliza la memoria RAM de diferente manera, ésta se puede escribir y leer al mismo tiempo. Entonces, el dispositivo se convierte en un sincronizador que permite el intercambio de video entre dos sistemas. El proceso de controlar la relación entre las direcciones, propicia que la memoria RAM tenga un retraso variable. Las direcciones generadas por los contadores generalmente llegan a un límite y regresan a cero una vez que alcanzaron su máximo nivel de conteo al finalizar cada uno de los cuadros. Es por esto, que el espacio de la memoria pareciera ser circular, ya que la lectura y la escritura de direcciones se realiza en uno y otro lado alrededor del "círculo". Si la dirección de la lectura es cercana a la dirección de la

escritura, el retraso es corto. Si está exactamente frente de la dirección de escritura entonces el retraso es máximo.

Para eliminar los errores de temporización estática como los retrasos de propagación en grandes sistemas, se puede utilizar un sincronizador con menos de un cuadro de memoria RAM.

Existe la posibilidad de poder manipular ampliamente las imágenes con una unidad de efectos de vídeo digital⁷. Esta amplia gama de efectos se puede dividir en varias categorías. Algunos efectos tales como la posterización o el cuadriculado son efectos relativamente simples de realizar. Algunos efectos más complejos tales como manipular el tamaño y la posición de la imagen se logran por medio de la transferencia de la información de los píxeles de un cuadro de almacenamiento a otro. La lectura y escritura son diferentes debido a que el valor del píxel varía en la pantalla.

Cuando se convierten las muestras, el convertidor analógico-digital debe trabajar a una velocidad constante de reloj, y a la salida del convertidor se obtiene una cadena de muestras ininterrumpida durante un periodo determinado. La compresión de tiempo, permite que la cadena de muestras se pueda partir en pequeños bloques para un manejo más adecuado.

Este procedimiento, se logra con un sistema como el que se muestra en la figura 1.2.5.

⁷ Digital Video Effects (DVE)

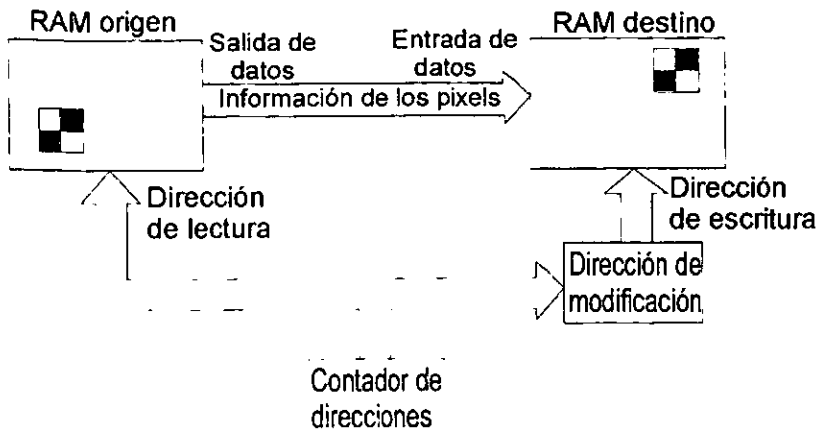


Figura 1.2.5 – Operación básica de una unidad de Efectos de Video Digital.

En un sistema como el que se muestra, el convertidor analógico - digital alimenta a un par de memorias RAM. Al momento en que el convertidor analógico - digital esta escribiendo información en una de ellas, el convertidor digital - analógico esta leyendo la información de la otra memoria. En cuanto la primera memoria RAM se llena, la salida del convertidor analógico - digital se cambia a la entrada de la otra memoria RAM, de tal manera que no existe pérdida de muestras. Entonces, la primera memoria RAM se puede leer con una velocidad de reloj mayor que la velocidad de muestreo. Así pues, la lectura de la memoria toma menor tiempo de lo que se tardó en escribirse, y la salida del sistema se detiene hasta que se llena la segunda memoria. Cuando esto sucede, las muestras están comprimidas en tiempo. En lugar de tener una cadena de muestras ininterrumpida, que es muy difícil de manipular, ahora se tienen bloques de la cadena de muestras acomodadas con las pausas convenientes entre ellos. Durante el tiempo en que el sistema se detiene mientras espera poder leer la

segunda memoria, se pueden realizar una gran cantidad de procesos. En todos los tipos de grabación, la compresión en tiempo de las muestras permite tiempo para la sincronización de patrones, codificación y corrección de errores de la información a grabarse.

En algunos sistemas digitales, al comprimirse la información en tiempo, permite dejar parte de la pista de grabación disponible para información de audio, y de esta manera, se puede utilizar la misma pista de grabación para vídeo y audio, utilizando el mismo equipo para la grabación de ambos.

Cualquier proceso de compresión en tiempo se puede revertir con una expansión de tiempo. Las muestras se escriben dentro de la memoria RAM a la velocidad del reloj de entrada, pero se pueden recuperar a la velocidad de muestreo estándar. La compresión en tiempo es totalmente indetectable, a menos que exista una falla de diseño del sistema.

Adicionalmente, la compresión en tiempo, permite que varios canales de audio se puedan transportar junto con el vídeo a través de un mismo cable.

En la transmisión de vídeo digital, como en cualquier transmisión de información, la sincronización es un posible problema para la correcta recepción de la información. En el campo de lo digital, este problema se soluciona fácilmente, ya que cada punto de recepción de información puede tener su propio corrector de tiempo de base.

Existen muchos procedimientos que permiten asegurar que la información que se recibe, corresponda con la información original, pero todos ellos con el fin de que a la información que se reciba se

pueda, si es necesario, corregir o si así lo ameritan disimular los errores.

Las grabaciones analógicas son mucho más susceptibles que las grabaciones digitales de sufrir ruido y hasta pérdida de información, debido a que la información en una grabación digital es información binaria. Un bit es correcto o incorrecto pero no existen etapas intermedias. Todas aquellas pequeñas cantidades de ruido se eliminan, pero es inevitable que algunos impulsos de ruido fuera de frecuencia causen que algunos de los bits queden en error. Algunas alteraciones causan que una gran cantidad de bits queden con información incorrecta.

La "visibilidad" de un bit en error depende de cuál de los bits de la muestra esta involucrado. Si el bit que contiene el error es el bit menos significativo de una muestra, entonces el error se puede "maquillar" totalmente y nadie podrá percibirlo. Por el contrario, si es el bit más significativo, entonces nadie dejará de percibir el error.

En código binario un bit tiene solamente 2 estados posibles. Si este es incorrecto, solamente es necesario invertirlo de estado y se corregirá el problema. De esta manera el proceso de corrección es sencillo y perfecto. El problema principal es la identificación del bit incorrecto. Esto se hace codificando los datos y agregando bits de redundancia. Evidentemente entre mayor falla exista, se necesitará agregar mayor redundancia. La cantidad de redundancia es proporcional a la cantidad de falla que se puede manejar. Dentro de los límites adecuados las muestras se pueden regresar exactamente a su valor original. De esta manera las muestras corregidas son indetectables. Si la cantidad de error excede la cantidad de redundancia entonces no es posible la corrección y, entonces se utiliza la técnica de disimulo o

encubrimiento de errores. El encubrimiento de errores es un proceso en donde el valor de una muestra faltante se estima de aquellos que están cerca. El valor estimado de la muestra no es necesariamente exacto al del original, y es por ello que en algunas circunstancias el disimulo puede ser audible y/o visible, especialmente si éste se presenta de manera frecuente.

El disimulo es posible solamente si se arregla, previamente a la grabación, la secuencia de muestras. El arreglo consta en la separación de las muestras en dos grupos, uno con las muestras numeradas como nones y otro con las muestras numeradas como pares. Estos dos grupos, se distribuyen a través de dos canales diferentes de la cinta, de tal manera que si existen errores incorregibles, solamente afecten a uno de los grupos. En la reproducción las muestras se vuelven a combinar, en su secuencia natural y ahora el error se ha dividido, y por lo tanto se puede estimar de los valores más cercanos.

Estos son solamente algunas de las técnicas para asegurar la integridad de los datos, aún cuando existen muchas otras que se ocupan de manera frecuente para la grabación y reproducción de vídeo digital.

Así pues, a pesar de que por muchos años las transmisiones analógicas de televisión han dado un buen servicio, éstas son extremadamente ineficientes debido a que las formas de onda transmitidas son directamente compatibles con los dispositivos de tubos de rayos catódicos y no se transmite nada durante los periodos sin información. Usando la compresión por tiempo, la reducción de datos y las técnicas de modulación digital, se puede obtener la misma calidad de imagen con solamente una fracción del ancho de banda

que ocupa una señal analógica. El incremento en el uso del espectro por otros sistemas, tales como la telefonía móvil, asegura que el futuro de las transmisiones de televisión y radio será digital.

Si la televisión de alta definición logra llegar a los hogares, esta lo hará a través de transmisión digital con reducción de datos, ya que el ancho de banda que se requeriría sería de otra manera poco económico. Además de no perjudicar el espectro, la transmisión digital es resistente a recepción con interferencia y ofrece una muy alta calidad de imagen a través de este sistema.

La mayoría de las ventajas que se atribuyen al vídeo digital también son aplicables al audio y es por ello que es hasta épocas realmente recientes se ha transformado.

El vídeo digital está mucho más allá que la simple comparación favorable con el vídeo analógico. Los aspectos más importantes son las tremendas posibilidades que no están disponibles en la tecnología analógica. Corrección de errores, reducción de datos, estimación de movimiento e interpolación, son tareas difíciles y aún más, prácticamente imposibles de corregir en el dominio de lo analógico, pero directas y "simples" en el dominio de lo digital. Una vez convertida, la imagen original no es más que una serie de números, y éstos se pueden almacenar, transmitir y procesar de muchas maneras diferentes.

La industria de la computación ha invertido décadas en perfeccionar máquinas para el almacenamiento, la transmisión y el procesamiento de cadenas de números a altas velocidades y sobre todo a un bajo costo. Las computadoras pueden generar imágenes artificiales para sistemas de diseño asistido por computadora y simuladores.

Actualmente, se ha enfocado el esfuerzo principalmente en mejorar la resolución y eliminar el parpadeo y algunos otros problemas a través de técnicas de compensación de movimiento, por lo que las transmisiones de vídeo digital pueden mejorar substancialmente con todas estas nuevas técnicas. Las grabaciones se pueden almacenar en unidades de disco de las computadoras ya sean magnéticos u ópticos que tienen cabezas en movimientos radiales, lo que permite un rápido acceso aleatorio a la información.

Las redes de comunicaciones a través de cable y satélite por todo el mundo se están incrementando utilizando transmisión digital, y se puede transmitir un paquete de información de vídeo digital a través de ella, tan fácil como transmitir un mensaje por fax o como una transacción bancaria. Proporcionando los números originales y transmitiéndolos en el orden correcto y a la velocidad correcta al destino final, para conversión a imagen, no importa cual sea el medio de transmisión, la información se transmitirá correctamente y con una alteración de la información prácticamente nula.

Los mundos del vídeo digital, el audio digital, las comunicaciones y la computación están estrechamente relacionados y es aquí en donde está el potencial real. El tiempo en el cual la televisión era un tema exclusivo de especialistas y estaba aislado de otras disciplinas es historia, la tecnología digital se está encargando de ello.

1.3 ADQUISICIÓN: CÁMARAS DE VIDEO

1.3.1 Cámara Vidicón.

"El *vidicón* es por su constitución el tubo de imagen más sencillo, es un tubo de cámara ampliamente utilizado, es substancialmente más

pequeño que el *orticón* de imagen, altamente adaptable a aplicaciones industriales y para equipo portátil de campo en que los requisitos de resolución son menos severos. Requiere una iluminación de escena aproximadamente ocho veces mayor que la del *orticón* de imagen, pero parece que su utilidad será cada vez mayor en aplicaciones en las que no sean serios inconvenientes la poca resolución y la necesidad de mayor iluminación." [DillenBurger, 1992]

El frente del tubo está compuesto por tres capas distintas que, enumeradas hacia el cañón de electrones, son las siguientes: la placa frontal de vidrio, el electrodo de señal (que es una película conductora que cubre la superficie interior de la placa frontal) y una capa fotoconductora depositada sobre el electrodo de señal.

Como capa fotosensible se utiliza un semiconductor cuya resistencia ohmica varía con la iluminación. Por lo tanto, se hace uso del efecto fotoeléctrico interno. Sobre un disco plano, situado al final de un tubo de cristal de 1" de diámetro, va colocada una capa de óxido de estaño o indio. Sobre ésta capa se coloca por rociado una capa muy fina de selenio o de sulfuro de antimonio.

La imagen se forma ópticamente sobre la capa fina del semiconductor. La imagen tiene normalmente una superficie de 9,5 x 12,7 mm². El lado de la capa del semiconductor que está apartado del camino de la luz se explora con un haz de electrones muy fino. Delante de la capa del semiconductor a una distancia de 1 a 2 mm se encuentra una malla muy fina con la mayor transparencia posible, que está unido con el cilindro anódico.

La magnitud del diámetro del haz de electrones se determina por medio de un diafragma pequeño de 40 a 50 μ de diámetro que se

encuentra en el ánodo del sistema. El haz de electrones se enfoca en el plano del semiconductor por medio de una bobina que produce un campo magnético homogéneo, y que se extiende a lo largo de todo el tubo. Entre la bobina de enfoque y la corona de cristal del tubo se encuentra una serie de bobinas de desviación que sirven para orientar el haz de electrones.

“La capa fotoconductora consiste en un mosaico de elementos que no son conductores en la oscuridad pero que se hacen eléctricamente conductores cuando están iluminados. Su grado de conductividad depende de la intensidad de iluminación. Cuando se proyecta la imagen óptica de la escena sobre la cara frontal, los elementos iluminados se hacen más conductores, y experimentan una variación de tensión, aproximándose a la carga de la placa de señal. Desde el momento en que esta placa de señal es positiva, se crea una imagen de carga en la placa de señal, carga que es la imagen o reproducción eléctrica de la imagen óptica.” [Shure, 1993]

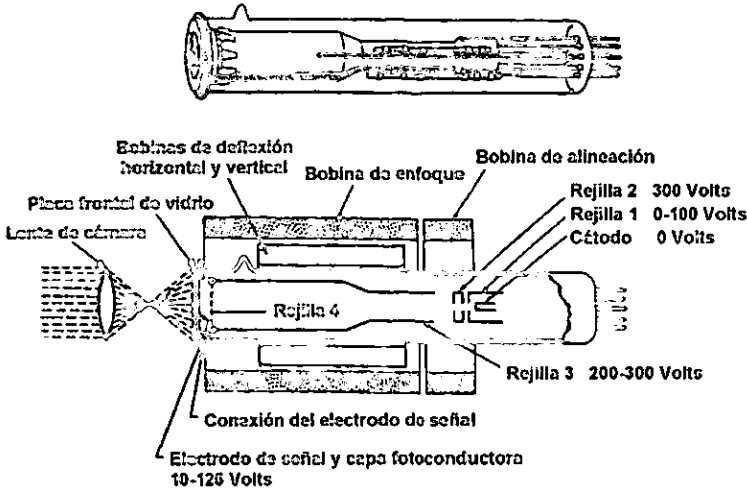


Figura. 1.3.1.1 Tubo vidicon y su estructura.

El cilindro anódico tiene en la mayoría de los casos una tensión de 300 Volts aproximadamente. Actualmente se utilizan también tensiones mucho mayores, tanto para el ánodo como para el cilindro anódico, con lo que se consigue una nitidez de imagen mayor.

Cuando los electrones del haz llegan a la capa fotoconductora, algunos son absorbidos para neutralizar las cargas positivas existentes allí debidas a la imagen, y los electrones excedentes se dispersan. La variación de la absorción de electrones produce en la resistencia de carga una corriente de la correspondiente intensidad. La caída de la tensión de señal en esta resistencia sirve como tensión de salida de cámara, la cual es transferida a los amplificadores a través del condensador de acoplo.

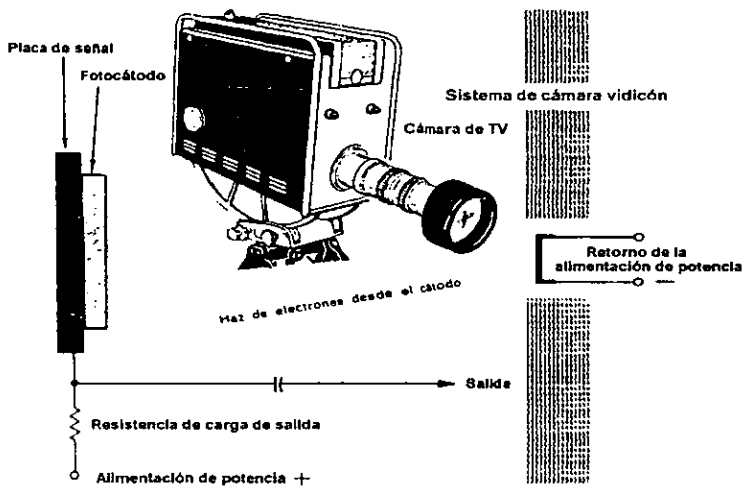


Figura 1.3.1.2 Sistema de cámara vidicon.

Inconvenientes del Vidicon

A pesar de las ventajas que reportan su pequeño tamaño y sencilla construcción, lo cual caracteriza al tubo vidicon, los productores de televisión no suelen ser partidarios de reemplazar con él al orticon de imagen a causa del detrimento de la calidad de la imagen que implica su uso. Ya se ha mencionado su pobre resolución y la necesidad de mayor intensidad de iluminación de la escena pero no son éstas las únicas desventajas del uso del vidicon. Su velocidad de respuesta, llamada "retraso" en condiciones de iluminación normal es tan lenta que las imágenes de los objetos móviles resultan a menudo borrosas y como empañadas.

Otra característica desfavorable del vidicón es su relativamente intensa corriente oscura, es decir, su señal aparece aunque no incida luz en el tubo. El principal efecto de la corriente oscura es producir sombreados desiguales de la imagen a causa de que la señal no es uniforme sobre toda el área de la imagen explorada y de que además depende del ajuste de tensión de la pantalla así como de la temperatura.

Por otra parte, la quemadura suele ser un problema clave en las cámaras equipadas con vidicón. Este efecto suele ser producido por la exposición prolongada del tubo a objetos inmóviles o estacionarios brillantes, que algunas veces dejan una imagen permanente en la capa fotoconductor lo cual hace necesario reemplazar el vidicón.

No obstante, hasta no hace mucho tiempo, en ciertas condiciones de campo, la necesidad de compatibilidad y simplicidad del equipo accesorio hizo que fuese imperativo el uso de los vidicones a pesar de sus inconvenientes.

1.3.2 Cámara de C.C.D.

El C.C.D. (*Charge Coupled Device*) es un circuito integrado de Silicio que detecta luz. Y lo podemos definir como ***“un dispositivo que contiene una superficie fotosensible hecha de silicio. Cuando un fotón cae sobre esta superficie, se libera una carga eléctrica, que se almacena en una región del dispositivo. La cantidad de carga almacenada es proporcional al número de fotones que golpean la superficie, es decir su respuesta es lineal”***. [Morison, 1991]

“Cuando la imagen de un objeto conformado por un conjunto de haces luminosos que inciden sobre el C.C.D, la luz se transforma en señales

eléctricas, por lo que el nombre de C.C.D. describe la forma en que la señal del C.C.D. es leída, (la distribución de cargas acumuladas se lee por medio de transferencia de las mismas a lo largo del arreglo)." [Berry, 1988]

Un detector de C.C.D. es un arreglo de elementos que tiene propiedades capacitivas. Un electrodo metálico está separado de un semiconductor tipo *p* por una capa aislante de dióxido de silicio, lo que conforma un arreglo tipo MOS (*Metal Oxide Semiconductor*), como se muestra en la figura 1.3.2.1.

"Los C.C.D.'s son muy sensibles y almacenan del 20% al 40% de todos los fotones que llegan a su superficie, además que los C.C.D.'s son sensibles a la radiación roja y a la infrarroja cercana, mientras que las placas fotográficas trabajan mejor en la región azul del espectro". [Morrison, 1991]

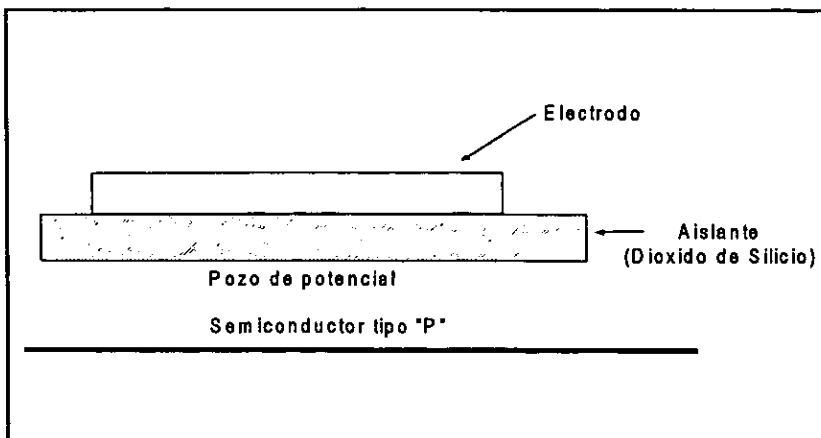


Figura 1.3.2.1 Capacitancia tipo MOS (*Metal Oxide Conductor*)

La aplicación más sencilla de un C.C.D. es una cámara para la detección de imágenes en el foco de un telescopio. Estas cámaras de C.C.D han superado a las placas fotográficas en muchas características. Algunas ventajas de los C.C.D.'s sobre otros dispositivos de detección son:

- Son dispositivos muy pequeños y compactos.
- Tienen una excelente sensibilidad a un amplio intervalo de longitudes de onda.
- Tienen una buena estabilidad y linealidad.
- Tienen compatibilidad inmediata con las computadoras. No se necesita el revelado de películas, ya que la salida del C.C.D. es una secuencia de pulsos eléctricos listos para ser transferidos a una computadora.

Las principales desventajas que los C.C.D s presentan ante otros detectores se explican a continuación:

- La imagen de sesgo.
- Diferencias físicas entre los pixeles, generando variaciones de sensibilidad de pixel a pixel.
- Efecto de corriente oscura, menor a la vidición
- Gran tamaño y peso de la cámara de enfriamiento del C.C.D.
- Ruido en la lectura.

Operación Básica del C.C.D.

Para mover los electrones almacenados en el arreglo C.C.D, sólo es necesario elevar y disminuir los voltajes de los electrodos en una forma irregular (por medio de pulsos de reloj proceso conocido como *clocking*), y los electrones serán "empujados" hacia el siguiente pozo después de cada pulso de reloj.

Si se considera la estructura de la figura 1.3.2.2 en la que se tienen 10 elementos de un C.C.D. y en donde cada tres electrodos se conectan entre si. Si en un tiempo $t = t_1$ los voltajes aplicados a los electrodos son $\phi_1 = +V$, $\phi_2 = \phi_3 = 0$ entonces se forman pozos de potencial bajo los electrodos 1, 4, 7 y 10, como lo muestra la figura 1.3.2.2.b, (los puntos simbolizan electrones que emigran dentro de los pozos de potencial). En un tiempo después, $t = t_2$, el voltaje ϕ_2 cambia a $+V$, pero ϕ_1 y ϕ_3 mantienen sus valores anteriores. Por lo tanto el perfil de potenciales es alterado, como lo muestra la figura 1.3.2.2.c. Ahora las cargas almacenadas son compartidas por dos electrodos adyacentes, debido a la difusión de los electrones desde el pozo original hacia el nuevo pozo creado. En un tiempo $t = t_3$, $\phi_1 = 0$ y ϕ_2 y ϕ_3 se mantienen, el perfil de potencial en este tiempo t_3 se muestra en la figura 1.3.2.2.d.

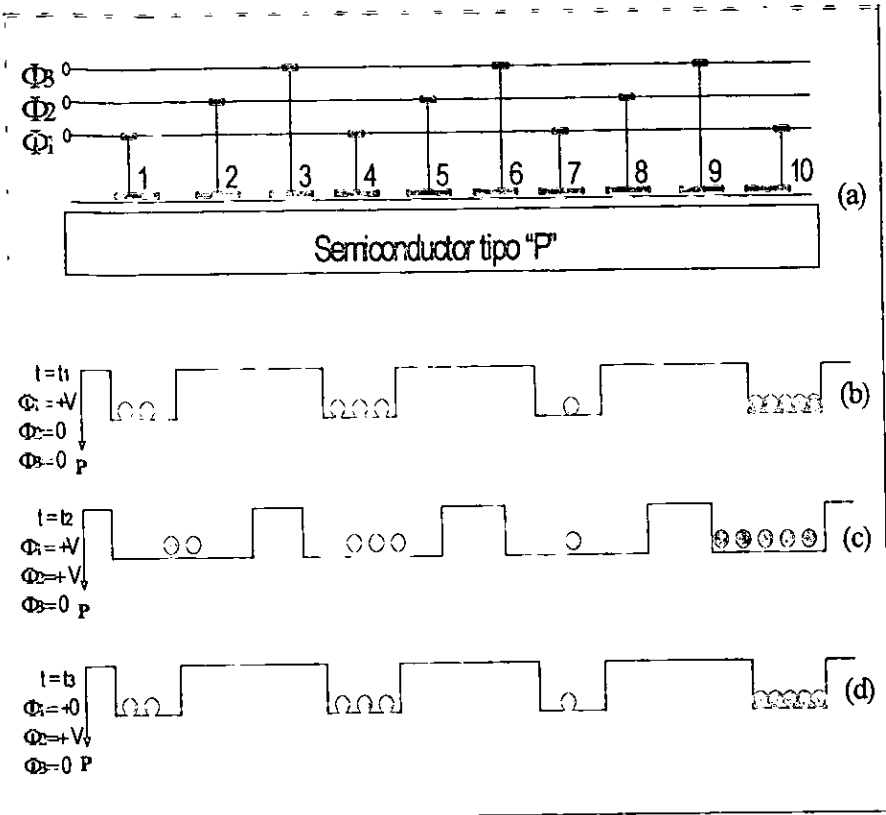


Figura 1.3.2.2 Transferencia de cargas en el C.C.D., los puntos representan cargas almacenadas.

La secuencia descrita representa la transferencia de cargas desde un electrodo hacia el siguiente en un registro de corrimiento del C.C.D. Ya que son necesarios tres voltajes, se requieren tres fases en los pulsos de reloj. Las formas de onda ϕ_1 , ϕ_2 y ϕ_3 necesarias en la transferencia se muestran en la figura 1.3.2.3

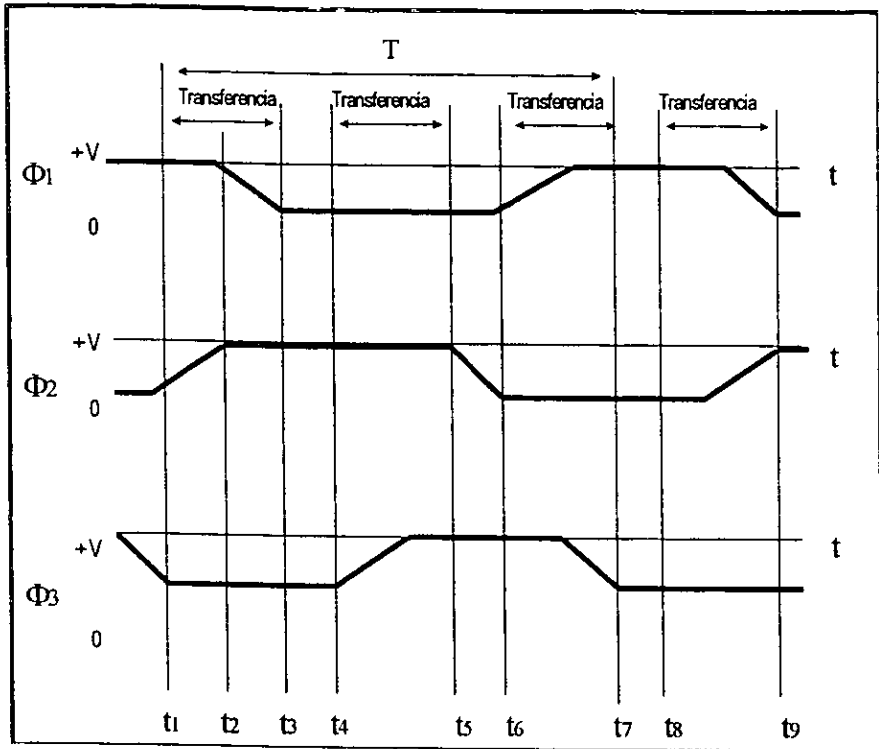


Figura 1.3.2.3 Formas de onda de las tres fases de excitación para el C.C.D.

La primera transferencia toma lugar entre t_1 y t_3 , la segunda entre t_4 y t_5 y la tercera entre t_6 y t_7 . Por lo tanto para cualquier ciclo de período T , se desarrollan tres transferencias. En los intervalos entre cada transferencia (por ejemplo t_3 y t_4), los voltajes de fase se mantienen constantes y los perfiles de potencial no se alteran.

Electrodos por Pixel

Observando la figura 1.3.2.2.b o 1.3.2.2.d, se puede observar que si una carga es almacenada debajo de un electrodo, ninguna información puede ser almacenada bajo los siguientes dos electrodos. En otras

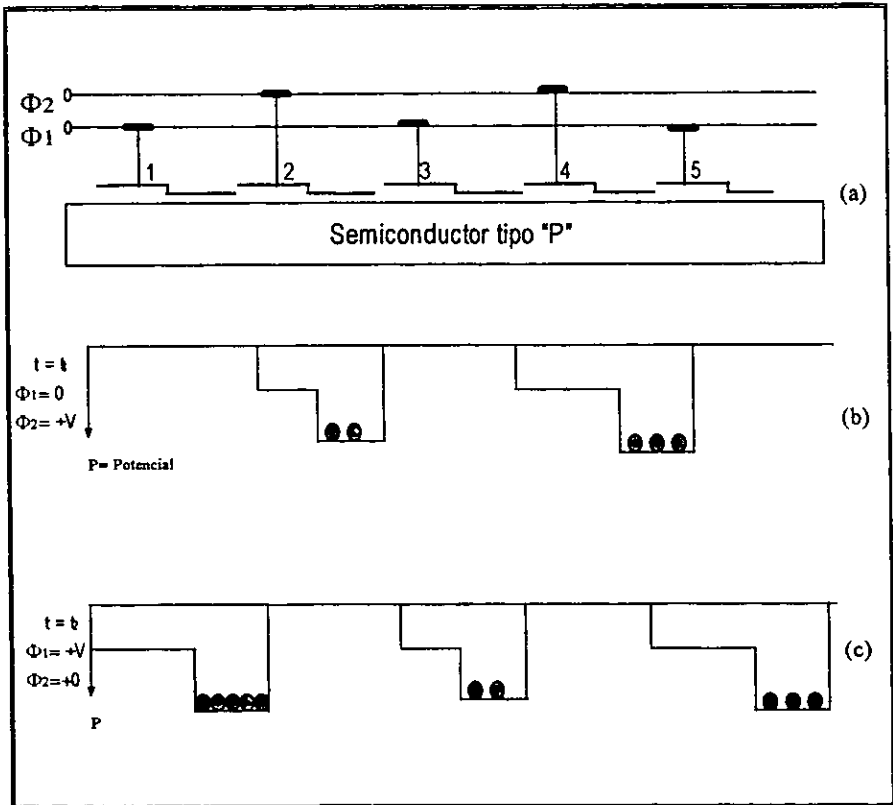
palabras, una celda de almacenamiento consiste de tres electrodos, donde solamente una carga es almacenada. Para el C.C.D. de la figura 1.3.2.2 los **electrodos por pixel** son tres ($E / P = 3$, donde E = Electrodos y P = pixel). Ya que la información es leída al final del registro de corrimiento, por ejemplo en el electrodo 10 de la figura. 1.3.2.2, se requieren tres corrimientos antes de que la siguiente carga (almacenada bajo el electrodo 7) sea leída. Debido a que en el periodo T toman lugar tres transferencias, la información debe ser leída (o escrita) una sola vez por ciclo.

C.C.D. de Dos Fases

Las estructuras de electrodos metálicos no planos pueden utilizar dos fases, según lo muestra la figura 1.3.2.4.a. La mitad derecha de cada electrodo se encuentra sobre una capa delgada de óxido y la mitad izquierda sobre una capa de óxido más gruesa que la del lado derecho. Debido a esta estructura las líneas de fuerza penetran más profundamente en el lado derecho del material. De ahí que el pozo de potencial tiene la misma forma que los electrodos. Los electrodos alternados se conectan juntos, resultando un sistema de dos fases, y las formas de onda se muestran en la figura 1.3.2.5.

Los perfiles de energía potencial se muestran en la figura 1.3.2.4. a/c para los tiempos t_1 y t_2 . En el tiempo $t = t_1$, $\phi = 0$ y $\phi_2 = +V$, por que no existe un pozo potencial bajo el electrodo 1 y si bajo el electrodo 2.

En $t = t_2$, $\phi = +V$ y $\phi_2 = 0$, se obtiene un perfil de la figura 1.3.2.4.c. En el intervalo de t_1 a t_2 las cargas son transferidas un electrodo a la derecha. Entre t_3 y t_4 , toma lugar la segunda transferencia.



1.3.2.4 Transferencia de cargas en el C.C.D. de dos fases.

Ya que existen dos electrodos por pixel, ($E / P = 2$) la celda de almacenamiento contiene dos electrodos, y la información debe leerse una sola vez por periodo T , en el intervalo t_2 a t_3 o de t_4 a t_5 , el cual también es llamado *intervalo de entrada/salida I/O*.

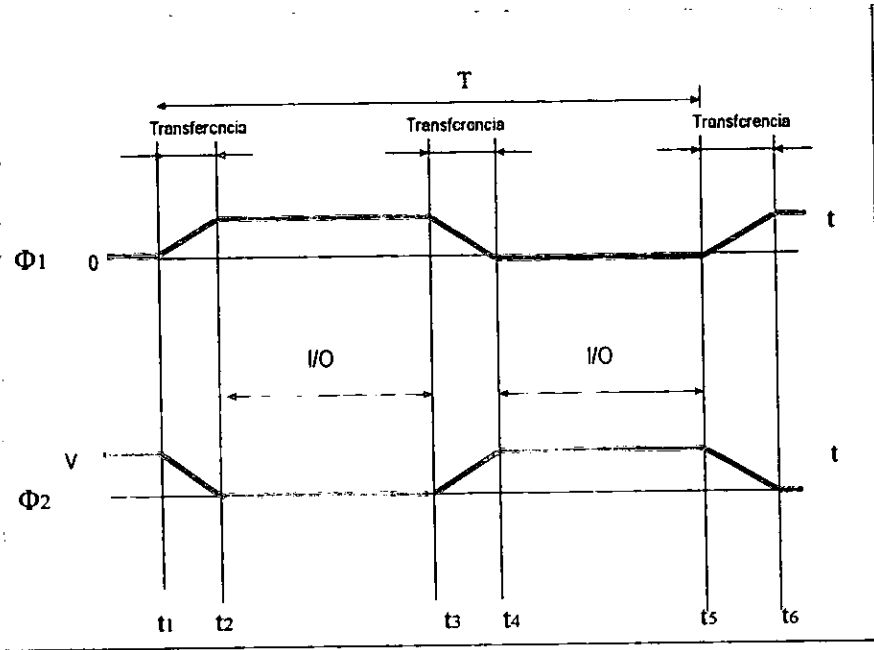


Figura 1.3.2.5 Formas de onda en el C.C.D. de dos fases.

Frecuencias de Operación.

La operación de un C.C.D. en un estado constante no es posible, debido a que las cargas generadas por elevación de temperatura bloquean los pozos de potencial vacíos, cambiando su estado de energía. Este fenómeno llamado efecto de corriente oscura, limita el intervalo de frecuencias mínimas de pulsos de reloj de 10 KHz a 1 MHz.

La máxima potencia disipada por las capacitancias del C.C.D., determina un intervalo de frecuencias máximas para los pulsos de reloj, entre 1 MHz y 30 MHz. Además no es posible trabajar las celdas a muy altas frecuencias ya que no habría tiempo suficiente para que

las cargas se transfieran completamente de un pozo de potencial hacia otro, y las cargas almacenadas se perderían.

Eficiencia Cuántica.

La cantidad observada en el C.C.D. es un flujo de fotones de luz, pero la cantidad detectada es un voltaje (V_0) que es amplificado y digitalizado. Si son absorbidos N_p fotones, entonces son detectados $(QE) N_p$ electrones por segundo, lo que corresponde a un voltaje de:

$$V_0 = A_{SDFC} ((QE) N_p e) / C \quad \text{Volts a la salida del C.C.D.}$$

Donde:

QE	Es la eficiencia cuántica
e	Es la carga del electrón, 1.6022×10^{-19} Coulombs
C	Es la capacitancia en el nodo de detección
A_{SDFC}	Es la ganancia del amplificador de salida

La principal característica atractiva de los detectores C.C.D. es su alta eficiencia cuántica, que es de 20% a 40%. [Smith, 1995]

Área total, Transferencia de Área y Transferencia de Líneas.

A pesar de que todos los C.C.D.'s operan según los mismos principios, la arquitectura de cada C.C.D. es diferente, como se muestra en la figura 1.3.2.6

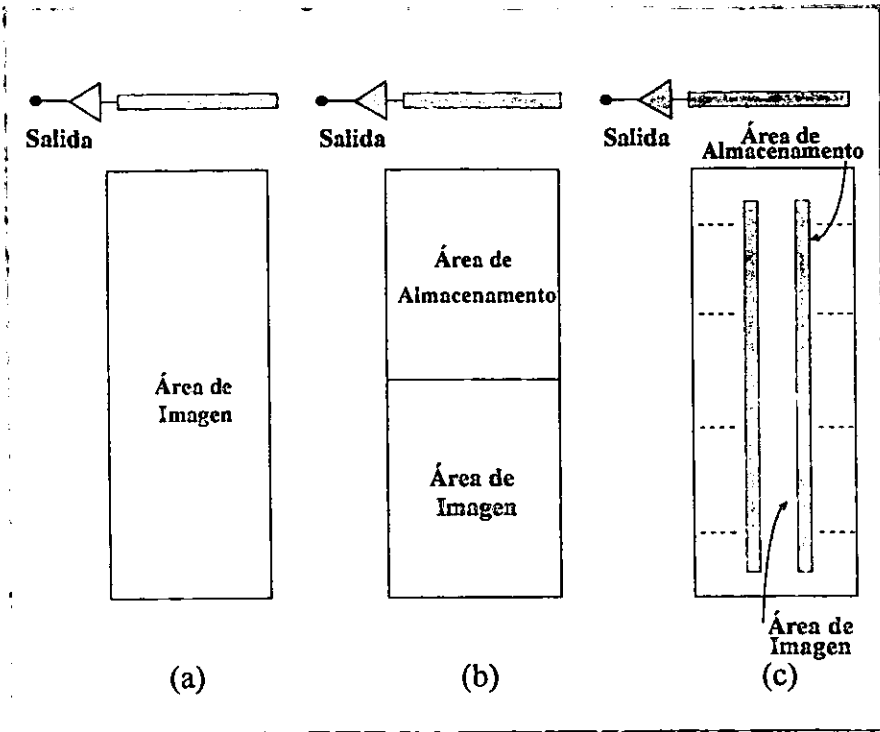


Figura 1.3.2.6 Dispositivos de (a) Área total, (b) Transferencia de Área y c) Transferencia de líneas.

La manera más sencilla de organizar un C.C.D. consiste en un dispositivo llamado *C.C.D. de Área Total*, en el cual se lee una sola vez la imagen entera, sin embargo existe un problema asociado con este tipo de dispositivo: la luz continúa cayendo sobre el circuito integrado al mismo tiempo que es leído. Esto produce un efecto de manchas en la imagen.

Para impedir este inconveniente se diseñó el dispositivo llamado *C.C.D. de transferencia de área* en el que la mitad inferior del arreglo

se cubre con un máscara opaca por medio de pulsos de reloj muy rápidos la mitad superior del arreglo se transfiere a la mitad inferior cubierta, evitando el manchado de la imagen. Estos C.C.D.'s son los más adecuados para las aplicaciones astronómicas, especialmente porque la máscara que cubre el área del almacenamiento puede ser retirada para obtener un C.C.D. cuyo tamaño es el doble del tamaño original.

Para aplicaciones en las que el C.C.D. debe trabajar a altas velocidades, los diseñadores idearon un C.C.D llamado *Transferencia de Líneas*. En este diseño, renglones alternados del arreglo del C.C.D, se cubren con máscaras opacas. Para leer la imagen todos los paquetes de carga en el área de la imagen se transfieren hacia las líneas cubiertas para después leer las cargas de la misma manera que en un arreglo de área total.

Salidas del C.C.D.

Como ya se indicó anteriormente, la distribución de cargas que representan la imagen pueden transferirse verticalmente con el envío de pulsos de reloj en los tres electrodos, en el caso de un C.C.D. de tres fases. Para un C.C.D. de tamaño finito, las cargas que son transferidas de un renglón a otro son apiladas en el extremo final del C.C.D. Para evitar ese apilamiento en el extremo final del C.C.D. se encuentra un renglón de pixeles llamado *registro de salida*, en cuyo extremo se encuentra un amplificador de salida. Los electrodos del registro de salida se ordenan en un ángulo recto con respecto al área de la imagen del C.C.D. y se transfieren las cargas en forma horizontal. Ya que el registro de salida es un solo renglón usualmente se llama *registro serie*, mientras que el área principal del C.C.D se forma por *registros paralelos*. La figura 1.3.2.7 muestra la organización de los registros en serie y paralelo.

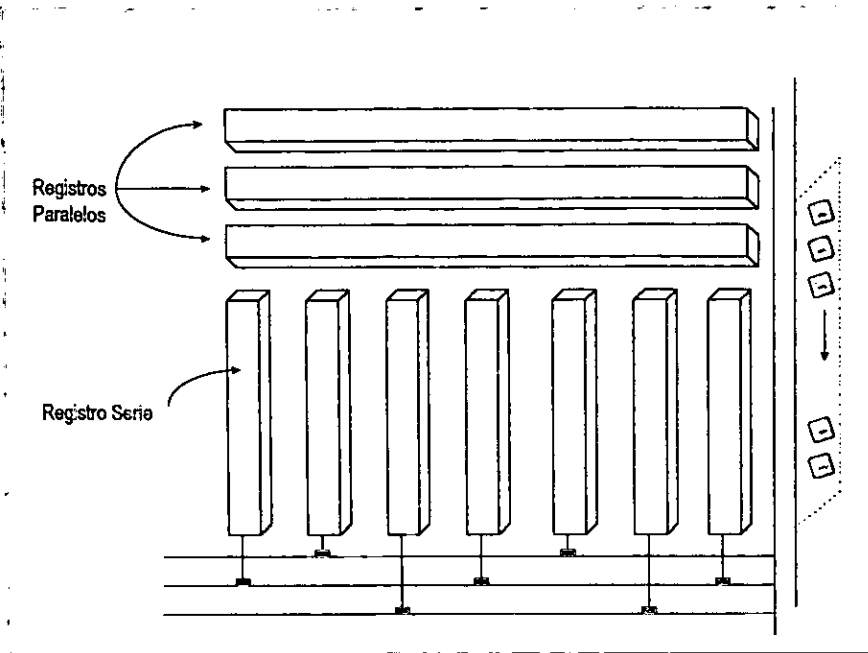


Figura 1.3.2.7 Vista esquemática del registro serie y de los registros paralelos.

La secuencia total de pulsos de reloj necesarios para la transferencia de carga es:

- Un corrimiento vertical de un renglón de la imagen completa en los registros paralelos.
- Un corrimiento horizontal de todos los pixeles en el registro serie. De esta manera cada carga del renglón se transfiere al amplificador de salida, un pixel a la vez, donde la carga se detecta y se convierte en un voltaje que puede medirse y almacenarse.
- Otra transferencia vertical, depositando el siguiente renglón de la imagen en el registro serie.
- Transferencia horizontal.

Y el proceso anterior se repite hasta que todos los renglones del C.C.D. hayan sido depositados en el registro serie y transferidos hacia el amplificador de salida.

En general, los pixeles asociados con el registro de salida son más grandes que los pixeles asociados con los registros paralelos, para asegurar mayor capacidad de almacenamiento, y por tanto menor riesgo de saturación (condición en la que los pozos de potencial se encuentran llenos de carga hasta el borde).

Detección de Carga

Cuando los paquetes de carga del arreglo de C.C.D, al finalizar la transferencia, llegan al registro serie, los electrones deben convertirse en una señal eléctrica, esto se lleva a cabo en una estructura llamada *nodo de detección de carga*.

Adquisición de la Imagen

La adquisición de la imagen del C.C.D. se puede realizar de diferentes de formas:

- Adquisición Normal
- Combinación interna
- Combinación externa
- Por muestreo Doble

Adquisición Normal

La forma más sencilla de adquisición de una imagen consiste en la transferencia de cada registro paralelo hacia abajo, y después de una transferencia de registro serie hacia el nodo de detección. En este método cada pixel del C.C.D. contiene la corriente oscura, el ruido de

inicialización, el voltaje de referencia y la carga que corresponde a la imagen detectada.

Combinación interna

La combinación interna consiste en combinar dos líneas del arreglo C.C.D, transfiriendo dos veces los registros paralelos hacia el registro serie, para después transferir la línea combinada hacia el nodo de detección. A pesar de que la combinación interna permite una lectura rápida, debido a que varios paquetes de carga se combinan en un solo paquete, se pueden llegar a saturar los registros del arreglo C.C.D.

Combinación externa

La Combinación externa se realiza por el corrimiento de los registros paralelos 1, 4, 7, 10,.... Hacia el registro serie, y después hacia el nodo de detección. Posteriormente se realiza el recorrido de los registros paralelos 2, 5, 8, 11,.... Hacia el registro serie y luego hacia el nodo de detección. Y de igual forma los registros paralelos 3, 6, 9, 12,.... De esta manera, si tuviéramos un arreglo de pixeles de 755 x 242 la información se combina de tal forma que se escribe el valor de cada pixel en la posición correspondiente a un arreglo de 252 x 242 pixeles de (755 x 242).

Por muestreo Doble

En este método el ruido es eliminado al sustraer la señal detectada, la señal de voltaje de referencia. Para esto, se debe realizar un doble muestreo, primero de la imagen detectada y después del voltaje de referencia, en un intervalo de tiempo de 20 μ s, para después obtener la diferencia de ambas señales.

Imagen de Sesgo (bias)

Cuando se lee la carga de un C.C.D. formada por un tiempo de exposición de cero segundos, aparece una imagen en la pantalla de despliegue, con niveles de intensidad de luz diferentes de cero, a esta imagen se le llama *imagen de sesgo*. La imagen de sesgo de un arreglo C.C.D. se debe a un diferente voltaje de cada pixel en el arreglo, sin la presencia de luz. Este efecto del C.C.D. se debe a los defectos de fabricación del dispositivo.

Una imagen de sesgo puede obtenerse al cerrar el obturador de la cámara de C.C.D. y fijar el tiempo de exposición en cero segundos, aunque existe otra forma de obtener esta información, por medio de una técnica denominada *sobrelectura*. En esta técnica el número de pulsos de reloj enviados para la lectura de las cargas del arreglo es mayor al número normal de pulsos requeridos para la lectura vertical y horizontal del C.C.D. real (por ejemplo 522 x 522 pulsos en lugar de 512 x 512). De esta forma las señales leídas del arreglo que caen fuera del área física del mismo contienen únicamente señales de sesgo.

Al sustraer la imagen de sesgo de la imagen deseada, es posible remover las irregularidades producidas por el C.C.D., ya que sin la sustracción de la imagen de sesgo, estos defectos serán parte de la imagen capturada. Es recomendable esperar al menos cinco minutos después de haber encendido el sistema de enfriamiento antes de tomar la imagen de sesgo.

Campo Plano

Un C.C.D. no tiene una respuesta uniforme a la luz que cae sobre su superficie, ya que cada pixel tiene diferente sensibilidad a la luz. Estas variaciones se deben a diferencias físicas entre los pixeles (como

defecto de fabricación) y a efectos de atenuación óptica (como partículas microscópicas de polvo en la superficie del C.C.D.). A pesar de que estas variaciones son muy pequeñas deben reducirse ya que producen imagen ruidosa.

Cuando se expone el C.C.D. a la luz, el brillo sobre el área total del circuito integrado es constante (lo cual se llama campo plano) la imagen no es uniforme, debido a que algunos pixeles son menos sensibles que los otros. Por lo que las imágenes de campo plano representan las variaciones de sensibilidad de pixel en el circuito integrado del C.C.D.

Para capturar una imagen de campo plano se debe apuntar la cámara de C.C.D. a un área iluminada uniformemente, como una hoja de papel o el interior de un domo alumbrado.

Es posible eliminar los efectos que estas variaciones de sensibilidad producen en la imagen deseada, al dividir la imagen detectada entre la imagen de campo plano.

Corriente Oscura y Enfriamiento

El movimiento aleatorio de los átomos del cristal de silicio originado por la temperatura normal del ambiente que lo rodea, genera la energía necesaria para crear pares de electrón - hueco con la ausencia de luz. A temperatura normal del ambiente que rodea al C.C.D, la corriente oscura es de 100,000 electrones por segundo en cada pixel, lo que significa que los pozos del potencial del arreglo de C.C.D. se saturarán solo por el efecto de la corriente oscura, en cuestión de segundos.

Para exposiciones que exceden aproximadamente 30 segundos de duración, es necesario capturar la imagen de datos térmicos y

substraerla de la imagen deseada, es indispensable que los tiempos de exposición para la imagen térmica y la imagen deseada sean de la misma duración.

"En el silicio, la corriente oscura disminuye una tercera parte de su valor inicial al reducir la temperatura 10° C. Debido a que el efecto de corriente oscura disminuye rápidamente con un decremento de la temperatura; la solución a este efecto consiste en enfriar el C.C.D. a una temperatura de 135° K, es decir -138° C, con esto es posible tener una corriente de 30 electrones por hora en cada pixel en lugar de 100,000 electrones por segundo en cada pixel." [McLean, 1989]

Aparentemente existe una desventaja en el enfriamiento del C.C.D., porque el C.C.D. debe encapsularse en una cámara de vacío bastante voluminosa, cientos de veces más grande y más pesada que el circuito integrado C.C.D., la figura 1.3.2.8 muestra un esquema de la cámara de enfriamiento.

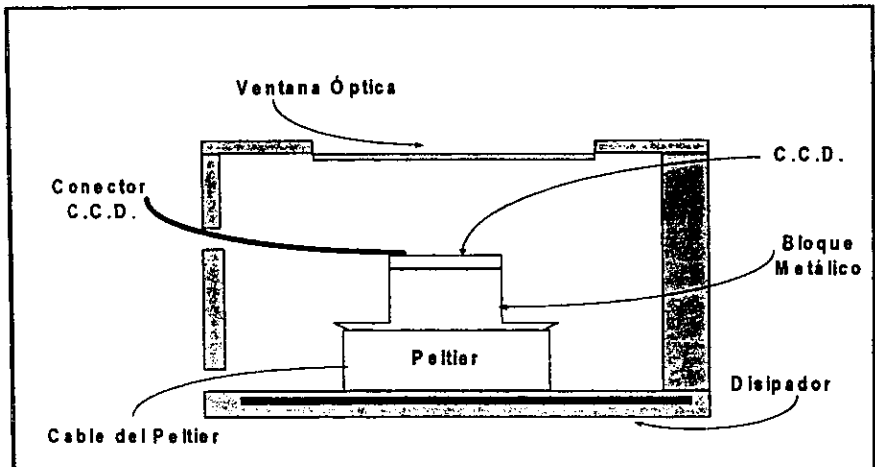


Figura 1.3.2.8 Cámara de enfriamiento del C.C.D.

Pixeles Defectuosos y Eficiencia de Carga

Si los pozos de potencial están llenos hasta el borde, la carga total almacenada puede distorsionar el nivel de los voltajes de reloj afectando la transferencia de cargas. Además, la carga de cada pixel puede desbordarse hacia los pixeles adyacentes, continuando la saturación en toda una línea de pixeles, manifestándose como rayas en la imagen del C.C.D.

La eficiencia de la transferencia de cargas disminuye cuando un grupo de cargas bloquean la transferencia de cada pixel. Esto se manifiesta visualmente como una cola en un punto de la imagen.

La eficiencia de la transferencia de cargas se fija en un valor de 0.99999 por transferencia, Esto significa que después de una transferencia, el 99.999% de la carga original fue transferida exitosamente, después de una segunda transferencia el 99.999% de esa carga transferida, es decir el 99.998% de la carga original, y finalmente después de 1,000 transferencias el 99.00% de la carga del pixel inicial ha sido transferida a una nueva posición, mientras que el 1% de la carga es distribuida entre los pixeles intermedios.

Ruido en la Lectura

Para leer la imagen almacenada, en los electrodos se aplican pulsos de voltaje que varía entre niveles altos y bajos. De esta manera la carga total de la imagen se transfiere a lo largo del arreglo del C.C.D. Los pequeños errores introducidos por el circuito en la lectura se conocen como ruido de lectura, que corresponde a una carga de 5 a 15 electrones.

Bibliografía.

1. Cheshire, David
Manual del Vídeo Aficionado
Ed. Copula Imagen, Ediciones CEAC
2. Watkinson, John
The Art of Digital Video
Ed. Focal Press. 2da., Edición 1994, p. 238-244
3. Dillenburger Wolfgang.
Técnica de televisión Tomo I
Ed. Urmo, Bilbao, España, 1992 p. 169-170
4. Alexander Schure
Televisión Básica vol. I, 2da Edición,
Ed. Marcombo. Boixereau Editores, 1993, p.54-56
5. Abell, George O.; Morrison,
David; Wolf, Sidney C. Exploration of the Universe.
Ed. Saunders College Publishing, Philadelphia. E.U.A., 1991, p. 148
6. Berry, Richard; Kanto, Veikko; Munger, John.
The C.C.D. Camera Cookbook.
Ed. Willmann-Bell. Berlin Alemania, 1988, p.12
7. Smith Robert C.
Observational Astrophysics.
Ed. Cambridge University Press. Cambridge Gran Bretaña, 1995, p. 64
8. McLean, Ian S.
Electronic and Computer-Aided Astronomy.
Ed. John Willey & Sons. New York, E.U.A., 1989, p. 119

9. Arch C. Luther
Principles of Digital Audio and Video
Archtec House, Inc., 1997

CAPÍTULO 2

TÉCNICAS DE COMPRESIÓN DE VÍDEO

2.1 COMPRESIÓN DE IMÁGENES FIJAS

2.1.1 Codificación Predictiva

Como con la codificación por transformada, la codificación predictiva es una de las técnicas que explota la memoria de la fuente. La idea básica es codificar la diferencia entre el valor de la fuente y una predicción de este valor obtenida de los valores fuente previamente recibidos. Así pues, el objetivo principal de la codificación predictiva es el eliminar la redundancia mutua entre los píxeles sucesivos y codificar únicamente la información novedosa.

Los valores de la fuente o los valores pasados pueden ser valores tanto escalares como vectoriales. Si la fuente tiene memoria, la predicción residual tendrá menor variación que la fuente misma. Para el mismo número de bits, cuando se codifica el error cuadrático medio,

la diferencia de información será menor que el que se obtendría si se codificara la fuente directamente.

Generalmente, una secuencia de datos tiene cierta dependencia o redundancia de una muestra con la siguiente. Consideremos una secuencia muestreada $u(m)$, que sea codificado hasta $m=n-1$ y en donde $u'(n-1)$, $u'(n-2)$,... son valores de una secuencia reproducida (decodificada).

En el punto en que $m=n$, cuando se recibe $u(n)$, una cantidad $\bar{u}'(n)$ (una estimación de $u(n)$) se predice de las muestras decodificadas previamente $u'(n-1)$, $u'(n-2)$,..., esto es:

$$\bar{u}'(n) = \Psi(u'(n-1), u'(n-2), \dots)$$

en donde $\Psi(\cdot, \cdot, \dots)$ representa la regla de predicción.

Cuando se transmite la muestra $u(n)$, se tiene la ventaja de que las muestras transmitidas previamente pueden contener cierta parte de la información de la misma. De acuerdo a esto, la cantidad $u'(n)$, un estimado de $u(n)$, se predice de las muestras transmitidas previamente de acuerdo a la regla de predicción, y la secuencia de predicción de error se define entonces como:

$$e(n) \triangleq u(n) - \bar{u}'(n)$$

Con esto, es suficiente cuantificar $e(n)$ en lugar de $u(n)$ para transmitirlo.

Si $e'(n)$ es el valor cuantificado de $e(n)$, el valor reproducido de $u(n)$ esta dado por:

$$u'(n) = \bar{u}'(n) + e'(n)$$

El proceso de codificación continúa de manera recurrente de esta forma. Este método es denominado como modulación de codificación de pulsos diferencial¹ (DPCM) o PCM diferencial. La figura 2.1.1.1 nos muestra un sistema completo codificador-decodificador DPCM.

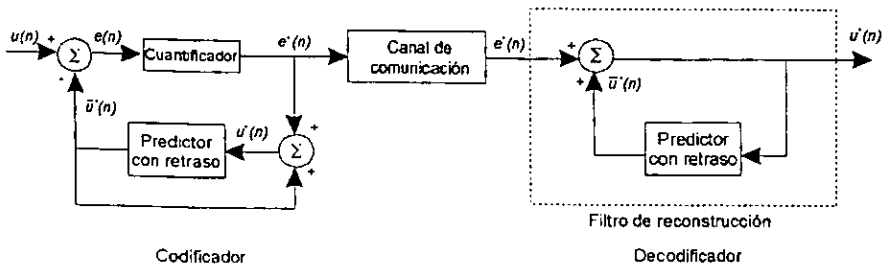


Figura 2.1.1.1 – Codificador-Decodificador DPCM.

Es importante hacer notar que el codificador tiene que calcular la secuencia producida $u'(n)$. El decodificador es simplemente el loop de predicción del codificador.

Con un predictor razonable, el valor cuadrático medio de la señal diferencial $e(n)$ es mucho menor que de $u(n)$. Esto significa, que $e(n)$ requiere menos bits de cuantificación que $u(n)$.

Un aspecto importante en DPCM es que establece que la predicción se basa en la salida (las muestras cuantificadas), en lugar de basarse en la entrada (las muestras sin cuantificar). Como resultado, el predictor debe quedar en el loop de realimentación alrededor del cuantificador, de tal manera que el error del cuantificador, en un paso dado, sea

¹ Differential Pulse Code Modulation (DPCM).

realimentado a la entrada del cuantificador en el siguiente paso. Esto tiene un efecto estabilizador que previene la tendencia y la acumulación del error en la señal reconstruida $u'(n)$.

Por otro lado, si la regla de predicción se basa en las entradas previas como se muestra en la figura 2.1.1.2 (a), el error de reconstrucción de la señal dependerá de todos los errores de cuantificación generados en la secuencia de predicción de error por adelantado $\varepsilon(n)$.

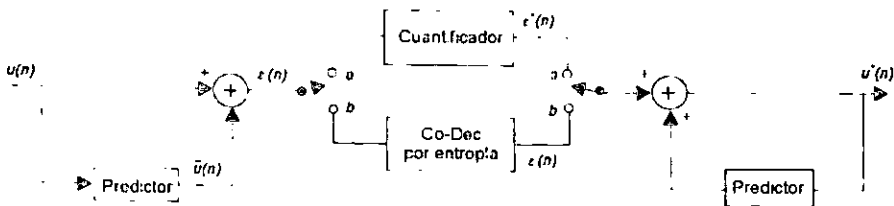


Figura 2.1.1.2 – Codificador por adelantado (a) con distorsión (b) con menos pérdidas.

En procesamiento digital, la secuencia de entrada $u(n)$ se digitaliza generalmente en la fuente misma por un número suficiente de bits². Entonces, se puede considerar a $u(n)$ como una secuencia entera. Forzando que las salidas del predictor sean valores enteros, la secuencia de predicción de error también tomará valores enteros y puede ser codificada entrópicamente sin distorsión. Esto proporcionará un codificador-decodificador predictivo con menos distorsión (figura 2.1.1.2 (b)), en donde, la mínima proporción que puede lograr será igual a la entropía de la secuencia de predicción de error $\varepsilon(n)$.

Así pues, en un sistema DPCM, los principales componentes son el predictor y el cuantificador, y podemos establecer que el desempeño

² Por lo general 8 para imágenes.

de los codificadores predictivos dependen del diseño del predictor y del cuantificador. Por simplificación, el predictor se diseña sin considerar los efectos del cuantificador. Entonces el predictor DPCM se diseña como

$$u^*(\bar{n}) = \sum_k a(k) u^*(n-k)$$

Modulación Delta

La modulación delta (DM) es el más sencillo de los codificadores predictivos. Este utiliza una función de retraso de un paso como predictor y un bit de cuantificación dando una representación de un bit a la señal. Así pues,

$$\bar{u}^*(n) = u^*(n-1), \quad e(n) = u(n) - u^*(n-1)$$

Un sistema práctico DM, que no requiere de muestreo de la señal de entrada, se muestra en la figura 2.1.1.3.

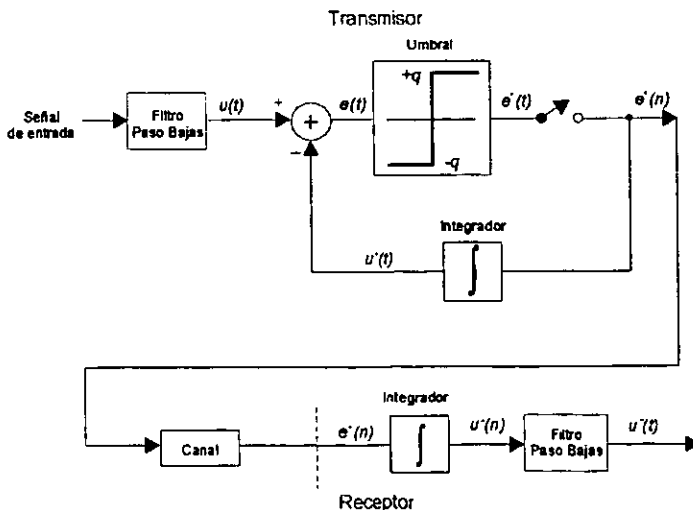


Figura 2.1.1.3 – Modulación Delta (un sistema práctico).

El predictor integra la salida cuantificada, que es una secuencia de pulsos binarios. El receptor es un integrador simple. La figura 2.1.1.4, nos muestra señales típicas de entrada-salida de un modulador delta.

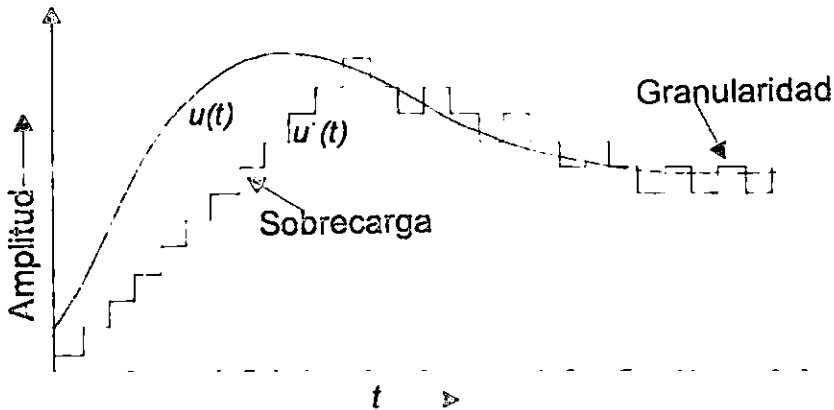


Figura 2.1.1.4 – Modulación Delta (Señales de entrada y salida típicas)

Las limitantes esenciales de la modulación delta son:

- 1 Atenuación por sobrecarga.
- 2 Ruido por granularidad.
- 3 Inestabilidad a los errores del canal.

La atenuación por sobrecarga ocurre siempre que existe un gran “brinco” o una discontinuidad importante en la señal, a lo cual el cuantificador no puede responder mas que en varios escalones delta.

El ruido por granularidad es la apariencia de escalones en la salida cuando la señal es prácticamente constante.

Estos dos errores se pueden compensar en cierto grado con un proceso de filtrado paso bajas en las señales de entrada y la salida. La atenuación por sobrecarga también se puede reducir incrementando la velocidad de muestreo, lo que reducirá las diferencias entre los píxeles. Por otro lado, una mayor velocidad de muestreo propiciará que sea más difícil la compresión del archivo. Una alternativa para reducir la granularidad manteniendo la simplicidad es utilizar un modulador delta de 3 estados.

Existe otro método llamado DPCM línea por línea, en el cual cada línea rastreada de la imagen es codificada independientemente por la técnica DPCM.

DPCM Bidimensional

Todas las técnicas anteriormente expresadas se han expuesto en una sola dimensión, pero éstas se pueden extender a dos dimensiones, en las cuales, para imágenes se utiliza para la predicción la información de los 4 "vecinos" más cercanos como se muestra en la figura 2.1.1.5.

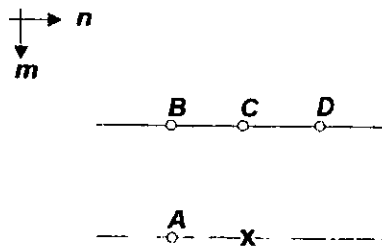


Figura 2.1.1.5 – Píxeles utilizados en una predicción bidimensional.

Para un buen funcionamiento de este tipo de técnicas, el cuantificador debe diseñarse para limitar los tres tipos de degradación: granularidad, atenuación por sobrecarga y movimiento en los extremos. Establecer

niveles muy toscos, causan la granularidad en regiones planas de la imagen. La atenuación por sobrecarga ocurre con bordes de alto contraste en donde el error de predicción excede los valores máximos del cuantificador, resultando con bordes borrosos. El movimiento en los extremos es causado por bordes poco nitidos, en donde los píxeles reproducidos en líneas rastreadas adyacentes tienen diferentes niveles de cuantificación. Una solución para minimizar estos efectos es incrementar el número de niveles de cuantificación, y usar un codificador entrópico para sus salidas. Esto incrementa el rango dinámico y la resolución del cuantificador.

Cuando se implementa físicamente un sistema DPCM bidimensional, generalmente el predictor se simplifica para minimizar el número de multiplicaciones por paso.

La selección de la regla de predicción también está influenciada por la respuesta del filtro para la reconstrucción de errores de canal.

Por lo general, DPCM es sencillo y se puede implementar fácilmente el hardware para aplicaciones en tiempo real. Sus mayores ventajas son la sensibilidad a variaciones en estadísticas de imagen y a errores de canal. Las técnicas adaptativas se pueden utilizar para mejorar el rendimiento de compresión de DPCM.

Técnicas adaptativas

El rendimiento de DPCM se puede mejorar adaptando las características del cuantificador y del predictor a variaciones en las estadísticas locales de la información de la imagen. Las técnicas adaptativas utilizan un intervalo de características de cuantificación y/o predictores de los cuales se selecciona el "óptimo" de acuerdo a las propiedades de la imagen local. Para eliminar el exceso debido al

procedimiento de adaptación, los píxeles codificados previamente se utilizan para determinar el modo de operación del codificador adaptivo. En ausencia de errores de transmisión, esto permite al receptor seguir la misma secuencia de decisiones hecha en el transmisor. Los predictores adaptivos son diseñados generalmente para mejorar la calidad subjetiva de la imagen, especialmente en los bordes. Una técnica popular es usar varios predictores, cada uno de los cuales se desempeña bien si la imagen es correlacionada en una dirección específica.

Otros métodos

A velocidades de transmisión de bits muy bajas, el desempeño de la técnica DPCM se deteriora rápidamente.

Una de las razones es que el predictor y el cuantificador, que fueron diseñados independientemente, no trabajan en sus niveles óptimos de funcionamiento. De esta forma las entradas sucesivas al cuantificador pueden tener una correlación significativa, y el predictor puede no ser lo suficientemente bueno. Existen dos métodos que pueden mejorar el funcionamiento y que son:

- 1.- Codificación predictiva retardada.
- 2.- Cuantificación de un vector predictivo.

En el primer método, se genera un código de árbol por un filtro de predicción excitado por los diferentes niveles de cuantificación. Como los píxeles sucesivos se codifican, el predictor selecciona la trayectoria en el árbol (en lugar del valor de una de las ramas, como DPCM), de tal manera que el error cuadrático medio se minimiza. Los retrasos se introducen al predictor para permitir el desarrollo del árbol con trayectorias suficientes hacia adelante.

En el segundo método, las entradas sucesivas al cuantificador se ingresan en un registro alterno, cuyo estado se utiliza para definir el valor de salida del cuantificador. De esta manera es como la salida del cuantificador depende de las salidas previas.

2.1.2 Codificación por Transformada Coseno Discreta.

La *transformada Coseno Discreta (DCT)* para compresión de imágenes es una técnica de codificación con pérdida. Los bloques de píxeles son de tamaño de 8 x 8 píxeles, éstos son procesados con el objeto de obtener su representación en el dominio de la frecuencia, y se logra por medio de la Transformada Coseno Discreta. Donde los componentes de frecuencia con valor mínimo son descartados, quedando únicamente los componentes de mayor energía.

Aplicaciones.

Las técnicas de compresión con pérdidas de imagen son usadas para reducir la cantidad de información necesaria para representar una imagen, sacrificando calidad.

Implementación

1. Se divide la imagen en bloques de tamaño de 8 x 8 píxeles.
2. Se aplica la Transformada Coseno Discreta a cada bloque.
3. Se descartan los coeficientes de baja energía (ver figura 2.1.2.1).

La DCT para compresión de imágenes causa degradación en la intensidad de los bloques, que se incrementa proporcionalmente con la tasa de compresión. Las técnicas de compresión de imágenes por la

DCT típicamente proporcionan tasas de compresión aproximadas de 10:1 sin degradación importante.

La Transformada Coseno Discreta se define como:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos[(2x+1)u\pi/2N]$$

para $u = 0, 1, 2, \dots, N-1$. De igual manera, la DCT Inversa está definida como:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos[(2x+1)u\pi/2N]$$

para $x = 0, 1, 2, 3, \dots, N-1$. En ambas ecuaciones α está definida como:

$$\alpha(u) \begin{cases} \sqrt{1/N} & \text{para } u = 0 \\ \sqrt{2/N} & \text{para } u = 1, 2, \dots, N-1 \end{cases}$$

La Transformada Coseno Discreta correspondiente en dos dimensiones es:

$$C(u) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos [(2x + 1)u\pi / 2N] \cos[(2y + 1)v\pi / 2N]$$

para $u, v = 0, 1, 2, \dots, N-1$. De igual manera, la DCT Inversa está definida como:

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u) \alpha(v) C(u,v) \cos[(2x+1)u\pi/2N] \cos [(2y+1)v\pi/2N]$$

para $x, y = 0, 1, 2, \dots, N-1$. En ambas ecuaciones α está definida como:

$$\alpha(u), \alpha(v) \begin{cases} \sqrt{1/7N} & \text{para } u, v = 0 \\ \sqrt{2/7N} & \text{para } u, v = 1, 2, \dots, N-1 \end{cases}$$

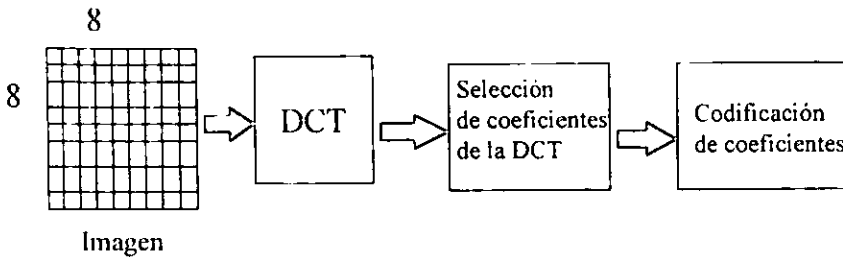


Figura 2.1.2.1 Esquema general de compresión por DCT.

2.2 COMPRESIÓN DE IMÁGENES EN MOVIMIENTO

2.2.1 Codificación DCT/DPCM

Los sistemas PCM son aquellos en los que se usa un código digital bastante directo para transmitir información. Hay algunos sistemas PCM alternos con características lo bastante distintos para justificar una notación especial.

La codificación de imagen predictiva puede proporcionar una excelente razón de compresión. Las técnicas predictivas trabajan con el principio de que se puede predecir la brillantez del píxel basándose en la

brillantez del píxel anterior. Empleando éste código predictivo, se codifica solamente la información diferente entre dos píxeles adjuntos.

En la transmisión de mensajes que tienen valores de muestra repetidos (es decir un valor medio), la transmisión repetida representa un desperdicio de capacidad de comunicación porque hay muy poco contenido de información en los valores repetidos. Una forma de superar esta situación es enviar sólo las diferencias entre valores de muestra sucesivos, codificadas en forma digital. Esto se conoce como *Modulación de Código de Pulsos Diferencial*, abreviada como *DPCM* (*differential pulse-code modulation*). [Wayne Tomasi, 1994]

La *DPCM* se emplea generalmente en la transmisión de información de imágenes muestreadas porque una porción muy apreciable del código asignado a cada nivel describe simplemente el nivel de fondo promedio. Así una imagen cuantificada de 6 bits (es decir, 64 niveles de brillantez por elemento de imagen) puede transmitirse con una comparable fidelidad usando *DPCM* de 4 bits. En la figura 2.2.1.1 se muestran los diagramas de bloques de un modulador y un demodulador *DPCM*.

El esquema de compresión *DPCM* trabaja sobre una imagen entera, píxel, por píxel. El primer píxel del ángulo superior izquierdo, no cambia los residuos, éste es codificado exactamente con la brillantez original.

El proceso lo mueve hacia el segundo píxel en la línea. El valor de brillantez del píxel precedente es sustraído de la corriente de brillantez del píxel, por lo que el resultado de la sustracción es el nuevo valor de codificación para el segundo píxel en la imagen. Este proceso se repite a través de la línea 0 de inicialización de la siguiente línea, el proceso inicia hacia abajo.

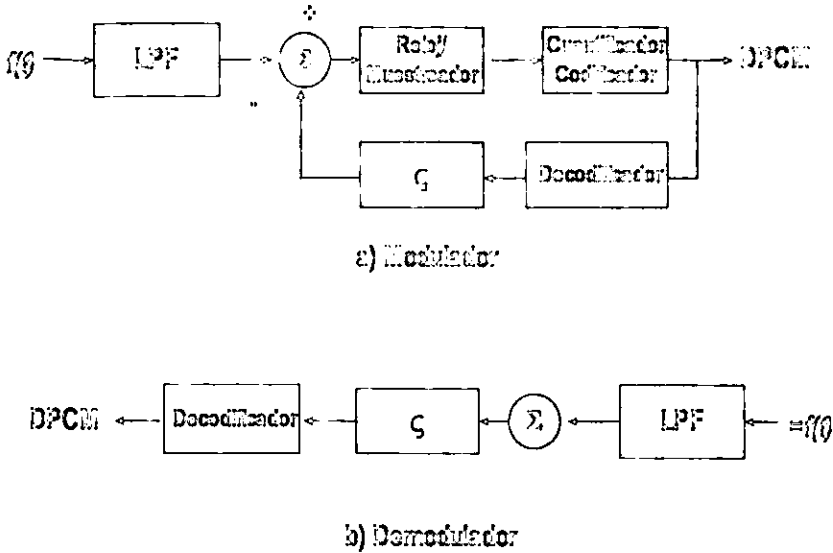


Figura 2.2.1.1 Modulación de Código de Pulsos Diferencial.

El proceso continúa hasta que la imagen es codificada totalmente en DPCM. Las pérdidas de la compresión y descompresión DPCM se ilustran en la tabla 2.2.1.

	76	0	0	0	0	1	0	0	0	0	1	0
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
x →												
Y	76	76	76	76	76	78	79	79	79	79	80	80
↓												

Imagen Original

Tabla 2.2.1.1 Código de Operación DPCM sin pérdidas.

Por ejemplo, si el primero de cinco píxeles de una línea de imagen tiene valores de brillantez de 23, 48, 76, 56 y 86, se puede asumir que la imagen se origina con 8 valores de brillantez, estos valores del código *DPCM* se muestran en la Tabla 2.2.1.2.

No. de Pixel	Original (Valores de 8 Bits)	Código DPCM (Valores de 6 bits)
Pixel 1	23	23
Pixel 2	48	$48 - 23 = 25$
Pixel 3	76	$76 - 48 = 28$
Pixel 4	56	$56 - 76 = -20$
Pixel 5	83	$83 - 56 = 27$
	40 Bits	30 Bits

Tabla 2.2.1.2 Código *DPCM* sin pérdidas.

La brillantez del 5º Pixel original se comprime desde valores de 5 x 8 bits (40 Bits), hasta la diferencia de valores de 5 x 6 Bits (30 Bits).

La operación *DPCM* trabaja asumiendo que los valores de los píxeles adyacentes son similares o altamente correlacionados, como resultado, estas diferencias serán usualmente valores pequeños.

Observando los valores del ejemplo anterior, ninguno es más grande que 31 o menor que -32, estos diferentes valores pueden ser codificados empleando números de 6 Bits en lugar de números de 8 Bits, produciéndose una tasa de compresión de $8/6 = 1.3333 : 1$. Si todos los valores tienen una diferencia de inferior a los 16 Bits, por lo que supuestamente números de 4 bits serán necesarios, con un rendimiento de 2 : 1 en la tasa de compresión.

En el ejemplo anterior, empleando un código de 6 Bits se puede dar la condición donde la diferencia entre los pixeles adyacentes es tan grande como 31 o menor que - 32. Aunque esto puede ser raro, es probable que ocurra en una imagen con una distribución normal de brillantez.

Un código de 6 Bits puede solamente representar la diferencia de valores entre 31 y - 32. Si la diferencia entre éstos dos pixeles es mayor que éstos límites, deberán ser realizados con un código de diferente valor. Comúnmente un código (tal como -32 en este caso), será reservado para indicar que es un caso de sobreflujo. Cuando éste código es empleado, indica que es absoluto en vez de diferencial, y que los siguientes valores de brillantez son valores de 8 Bits. Como ahora se codifica la corriente de brillantez en valores de 8 Bits, en vez de valores diferenciales de 6 Bits, revisando el ejemplo anterior, pero con una condición de sobreflujo, se muestra en la Tabla 2.2.1.3.

La brillantez del 5° pixel original, se da comprimiendo valores de 5 x 8 Bits (40 Bits) y para valores diferenciales de 5 x 6 Bits más 1 x 8 Bits del valor absoluto de brillantez da un total de $30 + 8 = 38$ Bits.

No. de Pixel	Original (Valores de 8 Bits)	Código DPCM (Valores de 6 bits)	
Pixel 1	23	23	
Pixel 2	48	$48 - 23 = 25$	
Pixel 3	86	$86 - 48 = 38$	(Sobreflujo)
		> -32	y 86 en 8 Bits
Pixel 4	56	$56 - 76 = -30$	
Pixel 5	83	$83 - 56 = 27$	
	40 Bits	38 Bits	

Tabla 2.2.1.3 Código DPCM sin pérdidas con un caso de Código de Sobreflujo.

La técnica de compresión *DPCM*, es semejante a otros métodos de compresión y es confiable bajo la estadística de que la brillantez del pixel en una imagen "normal", debido a que es estadísticamente raro que la diferencia de valores pueda exceder el límite del código seleccionado. En este caso 6 Bits, por lo tanto el promedio total de la tasa de compresión será generalmente buena.

La imagen comprimida en *DPCM* se descomprime por repetición, sumando la corriente de diferente valor del pixel y la brillantez del último pixel computado. Si el valor de la corriente, es un valor de sobreflujo, entonces la brillantez del pixel es el siguiente valor absoluto de brillantez del 8 Bit.

La operación de compresión *DPCM* trabaja mejor sobre imágenes que no tienen un número extraordinario de oscilaciones de gran brillantez entre pixeles adjuntos. Cuando se aplica a imágenes normales, el código *DPCM* puede proporcionar alrededor de 2:1 de tasa de compresión. Para imágenes de largas series de pixeles de valores constantes, las tasas de compresión se incrementan significativamente.

Entre las desventajas de la *DPCM* está el hecho de que si se comete un error, se mantiene una polaridad incorrecta hasta que se corrija. Además éstos sistemas adolecen de una posible sobrecarga por tasa de elevación debido a las operaciones de diferenciación y truncamiento. Por ejemplo, si dos muestras de imagen adyacentes difieren en más de ± 7 niveles (en el código de 4 bits uno se reserva para el signo), el sistema podría enviar sólo ± 7 , y la sobrecarga resultante podría causar un error en la reconstrucción. Este problema puede reducirse anteponiendo un filtro lineal, (por ejemplo un filtro RC) al codificador para limitar la máxima tasa de elevación o descenso de

una señal, (por supuesto ello también reduce la resolución de la imagen reconstruida). Es posible detectar la sobrecarga y variar el filtrado cuando ésta aparece (es decir el filtrado no lineal). Más aún, puede hacerse que el procedimiento se ajuste en función de los datos de entrada.

Desde una perspectiva más amplia, el propósito de la trayectoria de retroalimentación del codificador *DPCM* es predecir el valor de la muestra siguiente. El sentido de esto, es que si el sistema puede hacer una buena predicción del siguiente valor muestreado, no hace falta que el codificador envíe tanta información, (es decir, el codificador ha eliminado alguna redundancia en los datos).

Normalmente se utilizan como algoritmo de predicción los llamados *Sistemas Lineales de Predicción*, estos emplean líneas de retardo en sus derivaciones de ganancia ajustable para suprimir redundancia de los datos antes de la transmisión, pero a costa de cierta complejidad adicional.

Un sistema de codificación que es especialmente simple de construir es aquel que utiliza un sistema *DPCM* con un solo bit de salida que indica el signo de la diferencia entre muestras, este sistema *PCM* se conoce como modulación delta y se abrevia *DM*. [Storer, 1992]

2.2.2 Compresión por compensación de movimiento. [Stremler, 1992]

La forma en la cual el movimiento del ojo evita el desalineamiento es fundamental para la calidad de percepción de las imágenes de televisión. Muchos procesos necesitan manipular imágenes en movimiento de ésta misma manera para evitar las obvias dificultades de procesamiento con respecto a una trama fija de referencia.

Procesos de éste tipo son referidos como compensación de movimiento, contando con un proceso por separado el cual tiene medido el movimiento. Estimación de movimiento es literalmente un proceso, el cual analiza campos sucesivos y determina como los objetos se mueven de un campo al siguiente. Es una importante y capaz tecnología porque la función es comparable a la acción del ojo humano.

A fin de ver porque la compensación de movimiento es esencial para ciertos procesos, es necesario considerar la manera en la cual varios sistemas de formación de imagen manejan el movimiento. Esto puede hacerse mostrando la posición de imágenes sucesivas sobre el eje del tiempo. De hecho hay una considerable variedad de formas en la cual el eje del tiempo es manejado, y con esto guiar algunas potenciales incompatibilidades.

La figura 2.2.2.1(a) muestra el más simple eje del tiempo, el de la cinta, donde tramas enteras son simultáneamente expuestas o muestreadas, a 24 Hz típicamente.

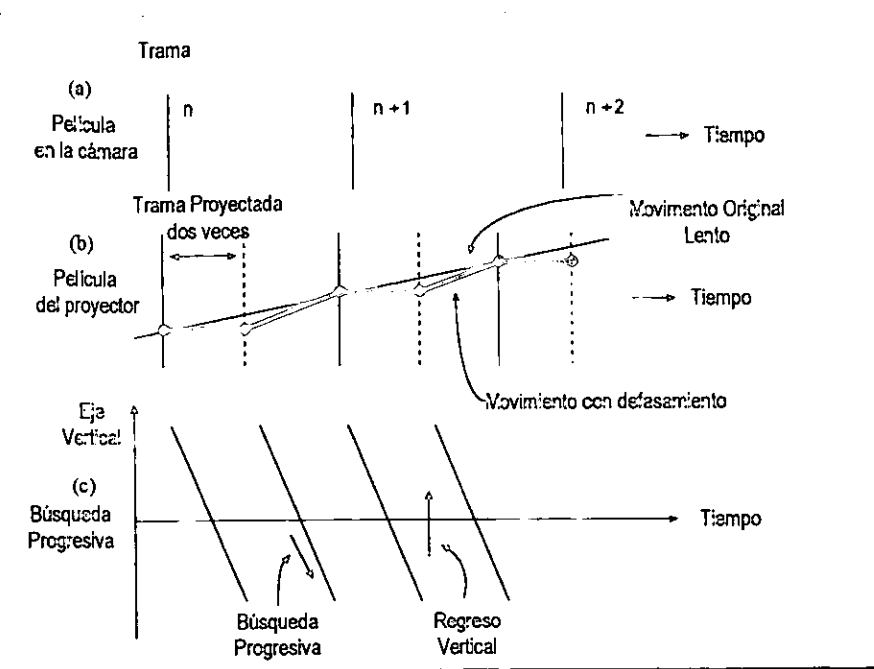


Figura 2.2.2.1 Características de espacio-tiempo de una película

El resultado es que la imagen está efectivamente en ángulos rectos al eje del tiempo. Durante la filmación, alguno de los periodos de trama es requerido para el corrimiento de la cinta, y el obturador se cierra mientras esto ocurre. Por lo tanto, la apertura temporal o exposición es un poco más corta que el periodo de trama y cuando son proyectadas, cada trama de la cinta es generalmente proyectado dos veces produciendo un parpadeo con una frecuencia de 48 Hz. El resultado con un objeto en movimiento es que el movimiento no es apropiadamente representado y se observa una sobreexposición. La figura 2.2.2.1(b) muestra el origen de la sobreexposición.

La figura 2.2.2.1(c) muestra el eje del tiempo en búsqueda progresiva de cámaras de televisión (por ejemplo, tubos de cámaras) y pantallas de rayos catódicos (CRT). La búsqueda vertical toma una parte substancial del periodo de trama y así la imagen se inclina con respecto al eje del tiempo. Como la cámara y la pantalla tienen la misma inclinación, esta inclinación no tiene efecto en la representación del movimiento.

En las cámaras CCD, la imagen es muestreada al mismo tiempo que la cinta, y así los campos están en ángulos rectos al eje del tiempo. Si es proyectada en un CRT, el eje del tiempo es deformado; como que las cosas son proyectadas más tarde que cuando fueron enviadas a la parte baja de la pantalla. Los objetos moviéndose transversalmente tienen velocidad vertical dependiente, como se muestra en la figura 2.2.2.1(e).

El mismo efecto es evidente si la cinta es proyectada sobre un CRT vía telecine. Las tramas de la cinta son muestreadas, pero en el CRT es barrido. En el telecine cada trama resulta en dos campos de 50 Hz estándar y así resulta también en defasamiento. En telecines de 60 Hz resultan tramas extrañas en dos campos, e igualmente resultan en tres campos, la bien conocida caída de 3:2.

La representación de movimiento (o carencia de éste) como se muestra en la figura 2.2.2.2; siendo exigentes, las cámaras CCD y principalmente los telecines, en algún grado no son compatibles con las pantallas de barrido CRT.

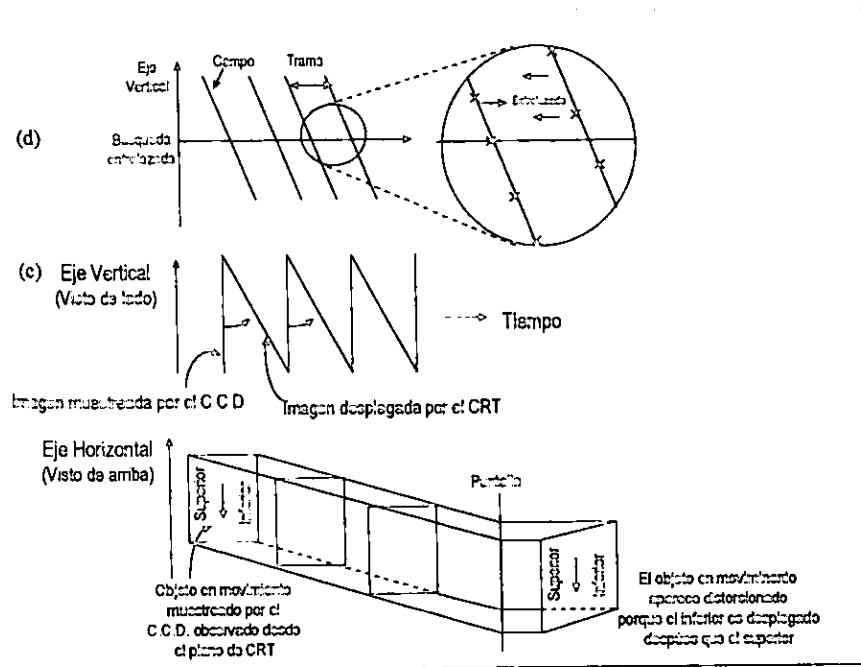


Figura 2.2.2.1 (d) El resultado de la búsqueda entrelazada. (e) Desplegando muestras de imágenes CCD sobre una búsqueda CRT resulta en distorsión en la presencia de movimiento.

La operación de convertidores estándar puede seguirse considerando la localización de entrada y salida de los campos sobre el eje del tiempo, de la misma manera que en el párrafo anterior. Los cuatro convertidores de campo convencionales se interpolan a través del tiempo, usando datos de pixel en orden desde los cuatro campos de entrada para registrar los probables valores de los pixeles en un campo de salida intermedio. Esta apertura temporal grande es requerida para eliminar frecuencias entre las dos porciones de campo, pero el movimiento no se maneja transparentemente.

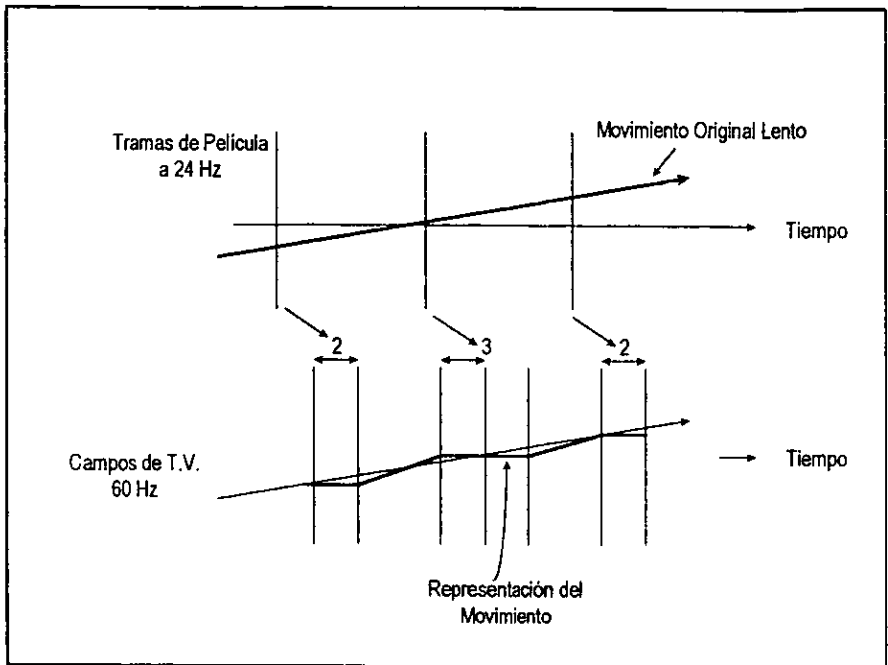


Figura 2.2.2.2 Los telecines deben usar caída de 3:2 para producir tasas de campo de vídeo de 60 Hz.

La figura 2.2.2.3(a) muestra que si un objeto está en movimiento éste puede estar en un lugar diferente en los campos sucesivos, pues interpolando entre varios campos resultan varias imágenes del objeto. La posición de la imagen dominante no se moverá suavemente, un efecto que se aprecia como una sobreexposición.

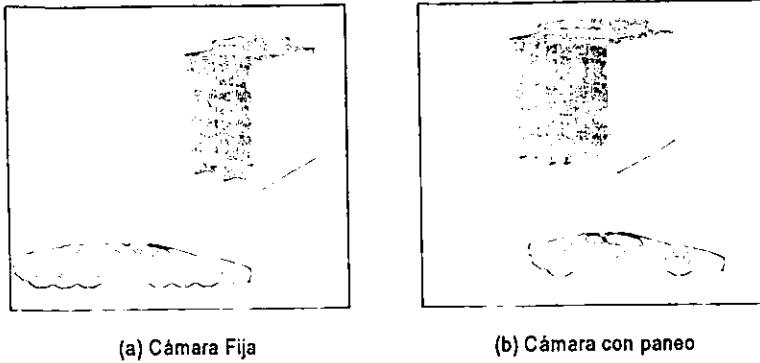


Figura 2.2.2.3 (a) Convertidor convencional de cuatro campos con un objeto en movimiento sufre múltiples imágenes, (b) Si el objeto en movimiento es paneado, la sobreexposición se observa al fondo.

Sin embargo cuando la cámara sigue al objeto en movimiento, será el mismo sitio en campos sucesivos y la figura 2.2.2.3(b) muestra que estará al fondo de tales sobreexposiciones.

Lo anterior puede prevenirse empleando compensación de movimiento, el cual puede realizarse considerando qué ocurre en un volumen de espacio temporal. Un convertidor convencional interpola solamente a lo largo del eje del tiempo. Mientras que un compensador de movimiento estándar puede girar su eje de interpolación fuera del eje del tiempo. La figura 2.2.2.4(a) muestra la entrada de los campos en los cuales el objeto está en movimiento. En la figura (b) se observa

que el eje de interpolación es alineado con el de la trayectoria del objeto en movimiento.

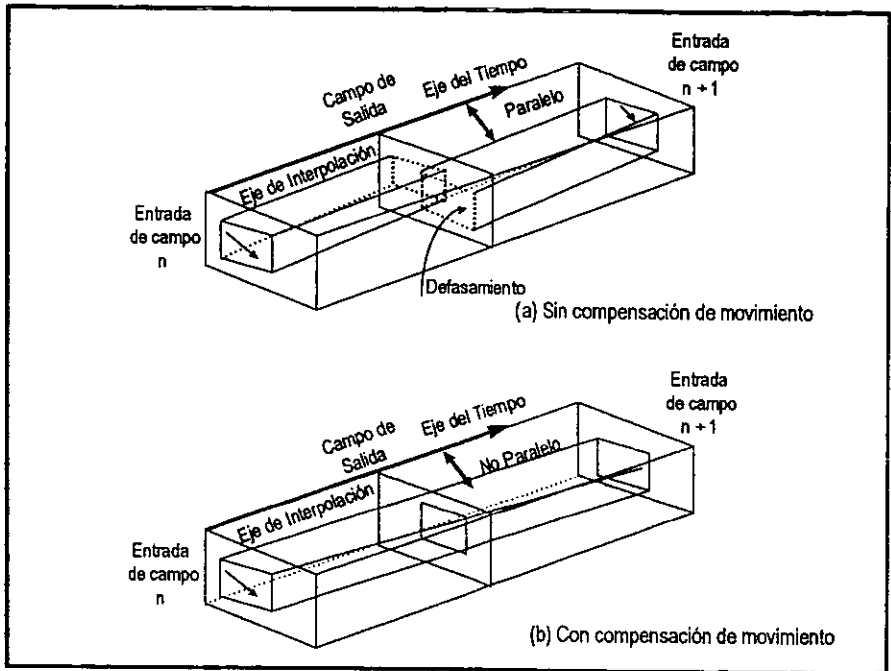


Figura 2.2.2.4 (a) Entrada de campos con objeto en movimiento, (b) Moviendo el eje de interpolación para hacerlo paralelo a la trayectoria del objeto.

Esto tiene un dramático efecto, el objeto ya no se mueve con respecto a su propio eje de interpolación, y así en ese eje ya no se generan frecuencias temporales debido al movimiento y no pueden ocurrir sobreexposiciones temporales. La interpolación correcta a lo largo de los ejes resulta entonces, en una secuencia de campos de salida en los cuales el movimiento está adecuadamente representado. La compensación de movimiento es una importante capacidad

tecnológica porque puede ser aplicada a mucho más que estándares de conversión.

La reducción de ruido en las señales de vídeo se logra combinando tramas sucesivas en el eje del tiempo, de modo que el contenido de la imagen es firmemente reforzado, mientras que no lo es el elemento al azar de la señal debido al ruido. La reducción del ruido se incrementa con el número de tramas sobre las cuales está integrado el ruido, pero el corrimiento de la imagen se puede prevenir mediante una combinación de tramas. Si la estimación de movimiento es posible, la imagen de un objeto en movimiento en una trama en particular puede estar integrada desde la imagen fuente en varias tramas, las cuales fueron superpuestas sobre la misma parte de la pantalla por desplazamientos derivados de la medición del movimiento. El resultado es que se hace posible una mayor reducción del ruido.

En las pantallas de T.V. convencional la porción de campo acotada resulta en un "parpadeo". Si la porción en la cual la pantalla es renovada, y es por ejemplo doblada, interpolando campos extras entre éstos en la señal de entrada el "parpadeo" puede eliminarse. Esta técnica de sobremuestreo es una forma simplificada de los estándares de conversión, los cuales requieren interpolación en el eje del tiempo.

La reducción de datos de vídeo depende de remover la redundancia de la fuente de señal. Mientras que la redundancia puede ser removida desde imágenes individuales, observándose que factores más grandes de compresión son alcanzados solamente por la eliminación de datos que son comunes en imágenes sucesivas, de modo que solamente los datos que son diferentes necesitan ser enviados. El movimiento solamente causa que la imagen se mueva con respecto a la rejilla muestra, causando que todos los valores de la muestra en un área en

movimiento cambien logrando así una reducción efectiva. Con estimación de movimiento, el movimiento de la imagen puede ser cancelado porque la comparación puede ser realizada mejor a lo largo del eje del movimiento que con el eje del tiempo. Mayores factores de reducción de movimiento son posibles entonces porque es necesario enviar solamente los parámetros de movimiento y un número pequeño de diferencias de imagen originales.

Los esquemas de reducción jerárquica de datos fueron propuestos donde un cuadro de diferente resolución puede ser decodificado desde una trama común de datos acorde a la tecnología de pantalla disponible. El submuestreo de imágenes es bastante fácil en sistemas de búsqueda progresiva, pero donde se utiliza entrelazado, el registro de una media imagen se convierte en imposible a menos que se use compensación de movimiento para cancelar el movimiento entre campos desde el proceso de filtrado.

Los aparatos de telecine utilizan un medio que no tiene entrelazado y en el cual la imagen entera es muestreada en un instante. El vídeo es muestreado con búsqueda progresiva y entrelazado así hay considerable disparidad entre las formas en que el eje del tiempo es manejado en los dos formatos. Existe una inevitable sobreexposición en las partes en movimiento de la imagen con telecines de búsqueda directa; particularmente cuando la caída es usada para obtener una tasa de 60 Hz. Actualmente telecines apropiados requieren procesos estándares de conversión en el eje del tiempo el cual puede hacerse correctamente solamente con compensación de movimiento. El resultado es una reducción en la sobreexposición.

En VTR'S que están retrocediendo a una velocidad que no es la correcta, los campos son convencionalmente repetidos u omitidos para

mantener correcta la tasa del campo de salida. Esto resulta en sobreexposición de los objetos en movimiento, particularmente en movimiento lento. Un dramático adelanto en la calidad de cuadros en movimiento lento puede ser obtenida usando un convertidor de estándares de movimiento compensado entre la VTR y el observador en el cual la tasa de entrada del campo es variable. Tal aparato es una de las más estrictas pruebas de estimación de movimiento, así un gran número de campos necesitan ser interpolados entre los campos de entrada y cada uno tendrá sintetizados objetos en movimiento en un diferente lugar. Las inexactitudes en la estimación de movimiento entre dos fuentes de campos serán vistas fácilmente como sean reveladas por muchos periodos del campo de salida en cinta a bajas velocidades. Como tal aparato es destinado a uso en deportes, movimientos entre campos grandes de 24 - 30 pixeles por periodo de campo deben ser esperados y correctamente manejados.

2.2.3. Compresión por Block Matching

En una imagen dada, un bloque de pixeles es seleccionado y tomado como referencia. Si el bloque seleccionado es parte de un objeto en movimiento, un bloque similar de pixeles existirá en la próxima imagen, pero no en el mismo lugar. Con métodos de bloque se asume que el desplazamiento del objeto es constante dentro de un pequeño bloque de B pels de dos dimensiones. La desventaja de esta suposición es que se presentan dificultades en escenas con objetos con movimientos múltiples o en escenas en las que partes diferentes del mismo objeto se mueven con desplazamientos diferentes. Si se disminuye el tamaño del bloque entonces esta suposición se vuelve más válida; sin embargo, el requerimiento de cómputo y transmisión de la información del desplazamiento se incrementa.

Como se muestra en la figura 2.2.3a, en block matching simplemente se mueve el bloque de referencia en la segunda imagen buscando una comparación de los valores de los píxeles. Cuando una réplica es encontrada, el desplazamiento necesario es usado como base para obtener un vector de movimiento.

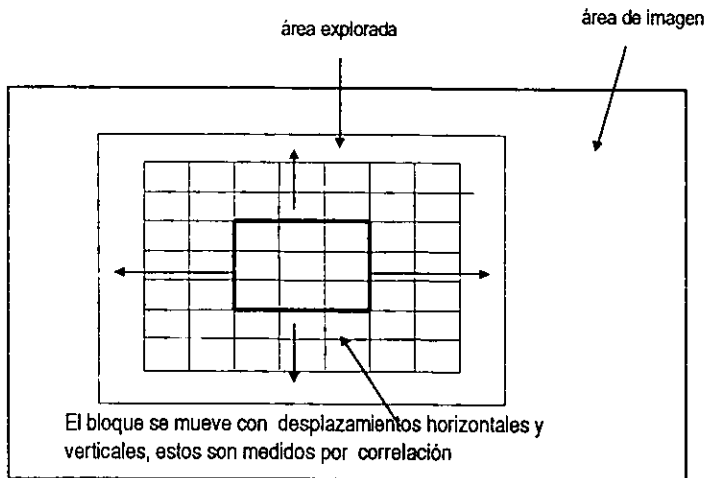


Figura 2.2.3a En block matching el bloque explora todas las posiciones de los posibles movimientos relativos dentro de una área de búsqueda y cada una es medida por correlación.

En las técnicas de block matching, la meta es estimar el movimiento de un bloque de tamaño $(n \times n)$ en la presente trama, en relación con los píxeles de la previa o futura trama. El bloque es comparado con el correspondiente bloque dentro de una área de búsqueda de tamaño $((n+2p) \times (n+2p))$ en la previa (o futura) trama. [Borko,1995]

Esto parece simple en concepto, block matching requiere una enorme cantidad de cálculos porque se debe probar todo movimiento posible alrededor del rango establecido. Así, si el objeto supuesto tiene un cierto alcance de movimiento alrededor del rango de los 16 píxeles, entonces será necesario probar 16 diferentes desplazamientos horizontales por cada una de las 16 diferentes posiciones verticales. A cada posición todos los píxeles en el bloque deben ser comparados con todos los píxeles en la segunda imagen.

Una manera de reducir la cantidad de cálculos es llevar a cabo la comparación en etapas, donde la primera etapa es errónea pero cubre un rango grande de movimiento y la última etapa es correcta pero cubre un rango pequeño. Eventualmente el último estado toma lugar de la predicción deseada.

Block matching puede sólo medir el movimiento al píxel más cercano, pero esto es correcto para muchos propósitos, incluyendo la reducción de datos donde se usa extensivamente. [Watkinson, 1994]

Los desplazamientos de D pueden ser estimados por correlación o técnicas de comparación. Así, se puede escoger D tal que se minimice alguna medida del error de predicción tal como.

$$PE(D) = \sum N(b(z,t) - b(z-D, t-\tau)) \quad (2.2.3.1)$$

Donde $N(\cdot)$ es una distancia métrica tal como una magnitud o el cuadrado de la función. Habiendo encontrado el valor mejor de D , todos los píxeles $z \in B$ se predicen, se codifican y se transmiten usando ese valor de D .

Como un ejemplo considere un bloque de $M \times M$ pels centrado alrededor del pel z_0 en la trama presente al tiempo t . Asumiendo un máximo desplazamiento horizontal y vertical de $d_{\text{máx}}$ pels, la región en la trama previa donde el mínimo de la ecuación (2.2.3.1) tiene que ser encontrado por.

$$z_0 \pm \begin{bmatrix} \frac{1}{2}(M-1) + d_{\text{máx}} \\ \frac{1}{2}(N-1) + d_{\text{máx}} \end{bmatrix} \quad (2.2.3.2)$$

El número de pels en esta región estará dado por

$$(M+2d_{\text{máx}})(N+2d_{\text{máx}}) \quad (2.2.3.3)$$

Así, si el tamaño del bloque es 9×9 y un máximo desplazamiento de $d_{\text{máx}}=10$ es usado, la región de búsqueda de la trama previa sería una área conteniendo 29×29 pels. Un método sencillo de búsqueda estaría en evaluar la ecuación (2.2.3.1) para todo cambio de pels en la dirección horizontal y vertical y escoger el mínimo. Esto requeriría $(2d_{\text{máx}}+1)^2$ evaluaciones de la ecuación (2.2.3.1). Además a menos que se incluyan cambios por pels fraccionarios las distancias también son incluidas. La exactitud con el que D podría ser obtenido se limita a un pel. Se han sugerido varias simplificaciones en la literatura. Un tipo de simplificación es usar el criterio de la ecuación (2.2.3.1). Por ejemplo:

$$PE(z_0, i, j) = \frac{1}{MN} \sum_{m \leq \frac{M}{2}} \sum_{n \leq \frac{N}{2}} (b(z_{mn}, t) - b(z_{m+i, n+j}, t - \tau))^2 \quad (2.2.3.4)$$

$$\text{ó } PE(z_0, i, j) = \frac{1}{MN} \sum_{m \leq \frac{M}{2}} \sum_{n \leq \frac{N}{2}} b(z_{mn}, t) - b(z_{m+i, n+j}, t - t) \quad (2.2.3.5)$$

donde $-d_{\text{máx}} \leq i, j \leq d_{\text{máx}}$

$$z_{mn} = z_0 + [m, n]^{\wedge}$$

La segunda definición de $PE(\cdot)$ en la ecuación (2.2.3.5) tiene la ventaja de que no se requiere ninguna multiplicación, también se observó experimentalmente que la definición precisa de PE no tiene un efecto significativo en la cantidad de exploraciones o la exactitud de estimación de D y, por consiguiente debido a simplicidad, generalmente se prefiere el criterio de ecuación (2.2.3.5).

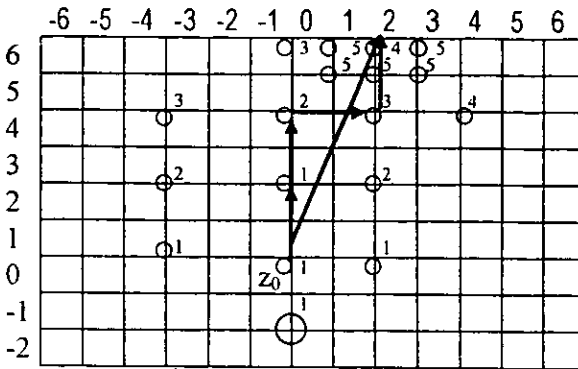
Además de la simplificación del criterio de búsqueda, varios métodos para simplificar el procedimiento de la búsqueda se han desarrollado y evaluado, describimos brevemente tres :

- (a) Búsqueda de 2d-logarítmico
- (b) Tres pasos de búsqueda
- (c) Dirección conjugada modificada.

En todos estos, la meta es requerir tan pocos cambios como sea posible y por consiguiente evaluar PE en tan poco tiempo como sea posible (cada cambio requiere un cálculo de PE). Reduciendo el número de cambios la suposición que se hace normalmente es que $PE(z_0, i, j)$ se incrementa monótonamente como el cambio (i,j) se aleja de la dirección de distorsión mínima.

En una búsqueda 2D-logarítmica, el algoritmo sigue la dirección de distorsión mínima. A cada paso se verifican cinco cambios como se muestra en la figura 2.2.3b. La distancia de búsqueda entre los puntos es disminuida si el mínimo está en el centro de localización, o en el límite del área de búsqueda.

Para el ejemplo mostrado en la figura 2.2.3b, cinco pasos fueron requeridos para encontrar el vector de desplazamiento al punto $(i,j)=(2,6)$.



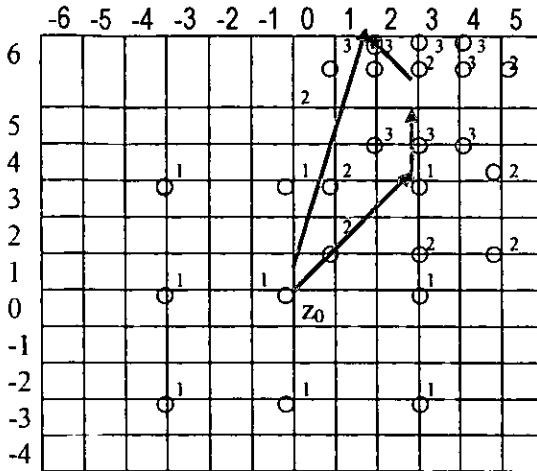
0ª DENOTA LA BÚSQUEDA DE LOS PUNTOS EN N PASOS.

Figura 2.2.3b Ilustración del procedimiento de búsqueda 2D-logarítmico. Los cambios en el área de búsqueda de la trama previa se muestran con respecto al pel z_0 en la presente trama. Aquí la aproximación de los vectores de desplazamiento $(0,2)$, $((0,4), 2,4)$, $(2,5)$, $(2,6)$ son encontrados en los pasos 1,2,3,4 y 5. d_{max} es de 6 pels.

En el procedimiento de búsqueda de tres pasos, ocho espaciamentos burdos de pels alrededor del centro del pel z_0 son probados en el primer paso. En el segundo paso de nuevo se usan ocho pels alrededor del pel de mínimo PE encontrado en el primer paso, pero en este paso el espaciamiento es más corto que en el primer paso.

En la figura 2.2.3c se muestran los puntos $z_0 + (3,3)$ y $z_0 + (3,5)$ como primera y segunda aproximación respectivamente. Este proceso es continuado todavía para otro paso, resultando un vector de desplazamiento final al pel $z_0 + (2,6)$. La búsqueda de dirección conjugada, en su forma más simple, busca la dirección del mínimo $PE(z_0, i, j)$. Que es un procedimiento de dos pasos. En el primer paso,

una búsqueda se lleva a cabo a lo largo de la dirección horizontal. Primero el mínimo de $PE(z_0, -1,0)$, $PE(z_0, 0,0)$ y $PE(z_0, 1,0)$ es calculado. Si $PE(z_0, 1,0)$ es el más pequeño, entonces $PE(z_0, 2,0)$ también es calculado y comparado.



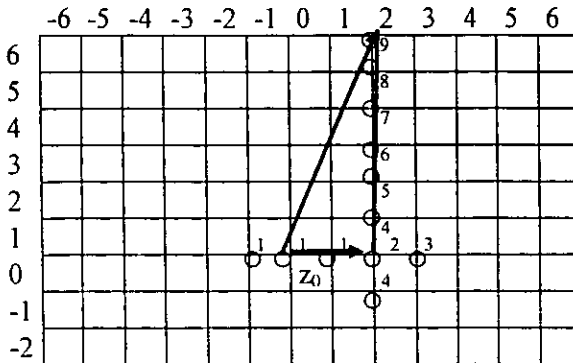
0ª DENOTA LA BÚSQUEDA DEL PUNTO EN EL PASO N

Figura 2.2.3c ilustra el procedimiento de búsqueda de tres pasos. Aquí (3,3), (3,5) y (2,6) son las aproximaciones de los vectores de desplazamiento encontrados en los pasos 1, 2 y 3. d_{max} es 6 pels.

Procediendo de esta manera, el algoritmo determina el mínimo en la dirección horizontal, cuando éste es localizado entre dos valores altos de sus pels vecinos.

En el segundo paso, el procedimiento usado es idéntico para calcular el mínimo en la dirección vertical.

La figura 2.2.3d muestra un ejemplo del método de la dirección conjugada en el cual los resultados de los dos pasos son los pels $z_0 + (2,0)$ y $z_0 + (2,6)$ respectivamente.



0^n DENOTA LOS PUNTOS DE BÚSQUEDA EN EL PASO N

Figura 2.2.3d Ilustración del método simplificado de la búsqueda de dirección conjugada. Aquí, $(2,6)$ es el vector de desplazamiento encontrado en el paso 9. $d_{máx}$ es de 6 pels.

Es obvio que los tres métodos de búsqueda pueden experimentar dificultades en ciertos casos. Se tiene como desventaja que, la complejidad de cálculos del procedimiento de búsqueda está relacionado con el número de cambios para los cuales PE necesita ser evaluado. Se debe notar que la estimación exacta del desplazamiento de los métodos de bloque está limitado generalmente a 0.5 pels. [Netravali, 1988].

Bibliografía.

1. Aml K. Jain
Fundamentals of Digital Image Processing
Prentice Hall (Systems Sciences Series), 1989
2. IEEE Press
Visual Communications Systems
1988
3. Wayne, Tomasi
Electronic Communications Systems Fundamentals Trough Advanced
Segunda Edición, 1994
Ed. Prentice Hall Carer & Technology, E.U.A., p. 636 - 639
4. Storer, James A.
Image and Tex Compression
Kluwer Academic Publishers E.U.A., 1992, p. 196-206
5. Stremler, Ferrel
Introducción a los sistemas de Comunicación, 3ª. Edición
Addison-Wesley Iberoamericana. Wilminngton, Delaware,
E.U.A., 1993, p. 560-566
6. Borko, Furht
A survey of multimedia techniques and standards.
Part 2: Video Compression, Real -Time Imaging 1:pag. 319 - 337
Ed. Cambridge University Press. Cambridge Gran Bretaña, 1995, p. 64
7. Watkinson, John
The Art of Digital Video
Ed. Focal Press. 2da., Edición 1994, p. 244-247

CAPÍTULO 3

NORMAS DE COMPRESIÓN DE VIDEO

3.1 RECOMENDACIÓN JPEG IMÁGENES FIJAS.

JPEG¹ es un mecanismo estandarizado de compresión de imágenes.

JPEG se diseñó para comprimir ya sea imagen a color o en escala de grises. Este sistema de compresión trabaja bastante bien en fotografías y materiales similares, pero no lo hacen tan bien con escritura, caricaturas o dibujos de líneas. JPEG maneja solamente imágenes fijas, pero existe un estándar relacionado llamado MPEG para imágenes en movimiento².

¹ Las iniciales de JPEG provienen del inglés Joint Photographic Expert Group (Unión de Grupos de Expertos en Fotografía), como se llamó originalmente al comité que escribió este estándar.

² Este estándar se verá posteriormente.

JPEG es un "sistema que pierde información", lo que significa que la imagen descomprimida no es lo suficientemente parecida a la imagen original con la que se inicio el proceso. Existen algoritmos de compresión de imágenes que tienen menos pérdida, pero JPEG logra una compresión mucho mayor que la que se logra con métodos con menor pérdida. JPEG es un estándar diseñado para explotar las limitaciones conocidas del ojo humano, haciendo énfasis en el hecho de que los pequeños cambios de color se perciben mucho menos precisos que aquellos pequeños cambios de brillantez. Así pues, JPEG está diseñada para comprimir imágenes que se vean bien a los ojos de los humanos. Por otro lado, si se pretende analizar las imágenes por medio de computadoras o equipos similares, todos éstos pequeños errores generados por JPEG pueden llegar a ser un verdadero problema, aún cuando ellos sean invisibles al ojo.

El estándar de compresión de imágenes JPEG no es simplemente un algoritmo, sino un paquete de técnicas de compresión de imágenes. Bajo este estándar de compresión de imágenes, se pueden "seleccionar" las características de la compresión. Existen 4 modos principales de compresión bajo JPEG:

- 1.- Sin pérdidas. La imagen se reproduce de manera idéntica a la original. Todos los demás modos tienen pérdidas.
- 2.- Secuencial. Este modo codifica la imagen en el orden en que fue "escaneada". Este es el modo en el que se procesan las imágenes generalmente.
- 3.- Progresivo. En este modo, primero se transmite una imagen burda que se puede desplegar rápidamente, y ésta se va definiendo de manera progresiva.

4.- Jerárquico. En este modo, la imagen se codifica de a diferentes resoluciones. El usuario puede seleccionar la resolución de visualización de la imagen.

La figura 3.1.1 nos muestra un diagrama representativo de la arquitectura de JPEG.

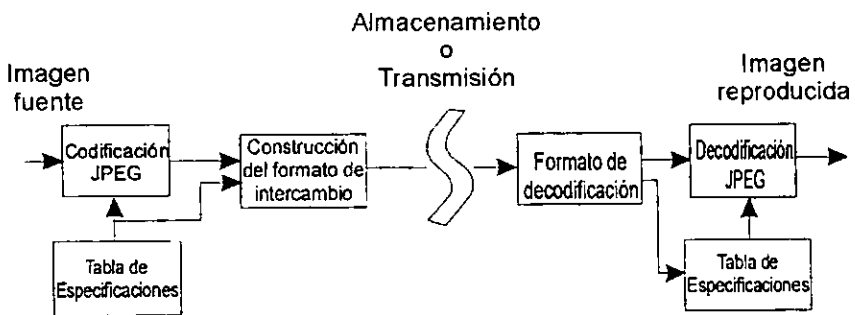


Figura 3.1.1 – Arquitectura de JPEG.

La imagen se codifica bajo el control de una o más tablas de especificaciones. La información ya codificada se envía, y en el otro extremo con un proceso inverso, se recupera la imagen original. Todas las codificaciones con JPEG (excepto en el modo sin pérdidas), se basan en la Transformada Coseno Discreta³.

En la figura 3.1.2 se muestra más detalladamente el proceso que se utiliza en el modo secuencial.

³ TDC (Del inglés Discrete Cosine Transform DCT).

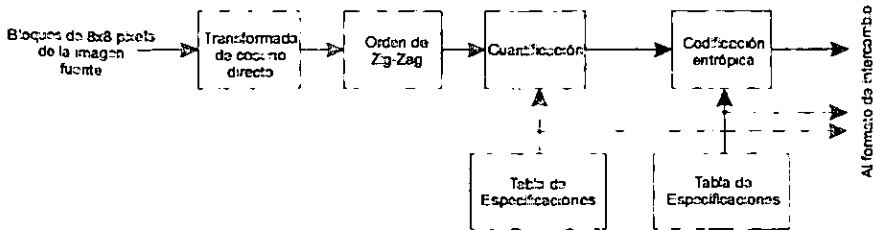


Figura 3.1.2 – JPEG Secuencial.

Cada componente de la imagen original se divide en bloques de 8 por 8 pixeles, a los que se les aplica el proceso TDC. Esto nos da como resultado un coeficiente de DC y 63 componentes de mayor frecuencia, cada uno de los cuales se cuantifica de acuerdo a una tabla de 64 valores de cuantificación.

Esta tabla de cuantificación, es la primera de las tablas de especificaciones del algoritmo, la cual debe ser proporcionada por el usuario y será parte de la cadena de datos que se transmitirá o recibirá. El usuario puede controlar la relación entre el grado de compresión y la calidad de la imagen dependiendo de las especificaciones de ésta tabla. El contenido de la tabla de cuantificación corresponde con el tamaño de los escalones que se utilizarán; el coeficiente DCT simplemente se divide por el tamaño del correspondiente escalón y el resultado se redondea al número de bits especificado. De esta manera, un gran valor de cuantificación resulta en una cuantificación burda y esto significa que valores grandes se redondearán a CERO.

Después de la cuantificación, el coeficiente de DC es codificado diferencialmente contra el coeficiente de DC del bloque de imagen precedente. Los otros 63 coeficientes cuantificados se ubican en un orden de zig-zag y se codifican estadísticamente de acuerdo, ya sea al

método de HUFFMAN o a un método aritmético de codificación que se describe en el estándar. En cualquiera de los casos se requiere de una segunda tabla de especificaciones para definir la codificación estadística. Nuevamente, el usuario puede variar ésta tabla para controlar el rendimiento de la compresión.

El último paso de la codificación es ensamblar los datos comprimidos y las tablas de especificaciones en una cadena de bits unidimensional de acuerdo al formato de intercambio, que se muestra en la figura 3.1.3.

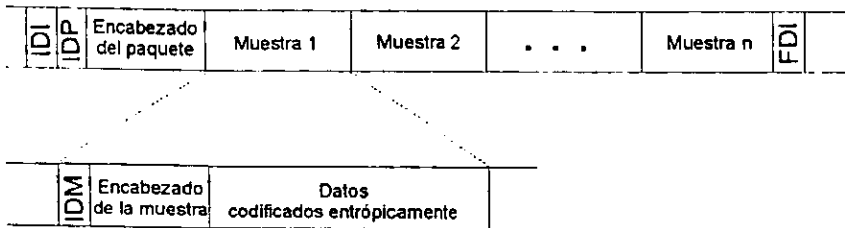


Figura 3.1.3 – Formato típico de intercambio de información en JPEG secuencial.

Los datos de la imagen codificada para un sólo cuadro de imagen se envían en un bloque, que (para formatos TDC) contiene todos los componentes de la imagen de manera intercalada. Las tablas de especificaciones se envían primero de tal manera que estén disponibles para la descompresión de los paquetes de datos que llegarán posteriormente. Los encabezados de los paquetes de información contienen los detalles del formato de la imagen y de su descompresión.

Unas de las grandes ventajas de JPEG es que el grado de pérdida de información se puede variar ajustando los parámetros de compresión. Esto significa que un creador de imágenes puede sacrificar el tamaño del archivo por una buena calidad de imagen de salida. También se

pueden realizar archivos "extremadamente" pequeños si no interesa mucho la poca calidad de la imagen. Este estándar es útil para aplicaciones tales como archivos de índices de imágenes⁴. Por otro lado si la calidad de la imagen a la salida no es la adecuada a los valores preestablecidos de compresión, ésta se puede mejorar hasta que sea satisfactoria, pero teniendo en cuenta que la compresión será mucho menor.

Otro aspecto importante de JPEG es que los decodificadores pueden sacrificar la velocidad de decodificación por la calidad de la imagen, usando aproximaciones más rápidas pero por otro lado más imprecisas que los cálculos requeridos. JPEG comprime simétricamente; lo que significa que la compresión y la descompresión toman la misma cantidad de tiempo. Algunos visualizadores logran un incremento notable de velocidad con estos métodos.

Existen dos razones importantes por las que se usa JPEG: para hacer los tamaños los archivos de imagen más pequeños y para almacenar información de color de 24 bits/pixel en lugar de información de 8 bits/pixel.

El obtener que los archivos de imágenes sean mas pequeños, es un gran logro para la transmisión de imágenes a través de una red y para el almacenamiento de bibliotecas de imágenes. Con JPEG se tiene la posibilidad de comprimir un archivo de color de 2 Mbytes a, aproximadamente 100 Kbytes, con lo que se logra un gran ahorro tanto en medio de almacenamiento como en tiempo de transmisión. JPEG puede, fácilmente, proporcionar una compresión de datos de color de

⁴ Este tipo de aplicaciones se utilizan para la previsualización de imágenes previamente a la selección de la imagen a recuperar.

20:1. Si se comparan los archivos GIF⁵ y JPEG, la relación de tamaños es normalmente más de 4:1.

Si los programas de visualización no soportan directamente JPEG, entonces se tiene que convertir el archivo JPEG en una imagen visible en otro formato. Aún cuando el visualizador sea capaz de ver imágenes JPEG, la decodificación y la visualización de imágenes JPEG toma mayor tiempo que la visualización de una imagen con un formato simple, tal como el GIF. Así pues, el usar JPEG esencialmente implica un sacrificio de tiempo/espacio. Pero esto puede llegar a ser mejor cuando se está hablando de un sistema de transmisión a través del teléfono o a través de una red, ya que el ahorro en tiempo debido a la transferencia de archivos más pequeños puede ser mucho mayor que el tiempo empleado para descomprimir el archivo.

La segunda ventaja fundamental de JPEG es que ésta almacena la información a pleno color: 24 bits/píxel (16 millones de colores). El formato GIF, que es el otro formato ampliamente utilizado, solamente puede almacenar imágenes de 8 bits/píxel (256 o menos colores). Actualmente el formato GIF está siendo reemplazado por el formato JPEG, como lo fueron, hace algunos años desplazados formatos en blanco y negro. Aún más, JPEG es mucho más utilizado que GIF para el intercambio de imágenes entre las personas que tienen una gran diversidad de equipos de despliegue, ya que esto evita la predefinición de cuantos colores se tienen que utilizar. Así pues, el uso de JPEG en lugar del GIF, es mucho más apropiado como formato estándar.

Mucha gente se espanta del término "compresión con pérdidas", pero no existe formato de imágenes digitales que puedan contener toda la

⁵ GIF es un método de compresión de imágenes sin pérdida que utiliza códigos entrópicos.

información que reciben nuestros ojos. La desventaja real de la compresión con pérdidas, es que si se comprime y descomprime una imagen en repetidas ocasiones, en cada ocasión que se hace este procedimiento existe una pequeña pérdida de calidad. Esto puede ser un severo problema sobre todo en algunas aplicaciones, pero afortunadamente para la mayoría de ellas no.

JPEG comprime imágenes bastante bien en realidad, cuando se utiliza con cierto tipo de imágenes, tales como fotografías. Para imágenes a pleno color, la información no comprimida es normalmente de 24 bits/píxel. El método de compresión con menos pérdidas mejor conocido puede comprimir la información aproximadamente con una relación promedio de 2:1. A este tipo de relaciones de compresión si se quisieran guardar 30 segundos de vídeo en un medio de almacenamiento como un disco duro, de requeriría un espacio en disco aproximado de 400 MBytes.

JPEG puede lograr compresión de 10:1 a 20:1 "sin pérdida de información visible", logrando con esto una reducción de almacenamiento efectivo a 1 ó 2 bits/píxel. Actualmente se pueden alcanzar compresiones de 30:1 a 50:1 con algunos pequeños defectos moderados, y cuando se requieren imágenes de muy poca calidad tales como presentaciones preliminares o archivos de índices, se puede lograr, razonablemente bien, una compresión de 100:1. Una imagen comprimida con JPEG a 100:1, ocupa el mismo espacio que una imagen del tamaño de la uña del dedo gordo a pleno color y contiene información mucho más detallada que esa imagen. La desventaja con JPEG (y otros algoritmos de compresión) es que entre más se comprime el vídeo, éste se ve peor. Además de esto, cada vez que se comprime el vídeo con JPEG a menos de una sexta parte de su tamaño original (6:1), entonces se van a ver pequeños cuadros, bordes

no deseados, y colores no correspondientes a los originales en las imágenes.

Las imágenes en escala de grises no se pueden comprimir a factores tan altos. Debido a que el ojo humano es mucho más sensitivo a las variaciones de brillantez que a las variaciones de color, JPEG puede comprimir datos de color con una relación mayor que lo que puede hacer con información de brillantez (escala de grises). Un archivo con JPEG de escala de grises generalmente es aproximadamente de 10% a 25 % más pequeño que un archivo con JPEG a pleno color de una calidad visual similar. Debido a que los datos en escala grises antes de comprimirse son de solamente de 8 bits/píxel o una tercera parte del tamaño de los datos a color, la relación de compresión calculada es menor. El umbral de pérdida visible por compresión es alrededor de 5:1 para imágenes en escala de grises.

El valor exacto al cual un error se convierte en visible, depende de las condiciones de visualización. Entre más pequeño es un píxel independiente, es mucho más difícil de percibir el error; por lo que los errores son mucho más visibles en monitores de computadora (a aproximadamente 70 puntos/pulgada) que en un monitor de alta resolución a color (de 300 puntos/pulgada o más). Así pues, una imagen de mayor resolución puede tolerar mayor compresión. Las relaciones de compresión mencionadas arriba son relaciones típicas para unidades de visualización típicas.

Actualmente existen muchos muy diferentes métodos de compresión todos ellos conocidos como "JPEG". El método comúnmente usado es el JPEG original (o su variante JPEG progresivo). El mismo estándar de ISO, también define un método bastante diferente llamado "JPEG

con menos pérdidas", y existe en proyecto un nuevo estándar con menos pérdidas⁶.

El JPEG con menos pérdidas es un método completamente diferente en el que realmente existen menos pérdidas. Por otro lado, éste no comprime ni siquiera de manera similar que como lo hace el JPEG original; puede comprimir información a pleno color a una relación aproximada 2:1. Un JPEG con menos pérdidas trabaja bien solamente para imágenes con tono continuo. Éste no provee compresión útil para imágenes con una gama de colores.

Como resumen se puede mencionar que si se necesita de tener una imagen con suficiente calidad sin pérdidas, JPEG no es precisamente el método adecuado.

3.2 RECOMENDACIÓN H.261 [U.I.T.C.C.I.T.T., 1988]

CODEC PARA LOS SERVICIOS AUDIOVISUALES A $n \times 384$ bits/s.

El CCITT considerando:

- Que existe una demanda significativa del servicio de videoconferencia.
- Que la transmisión digital a la velocidad de los canales H_0 o sus múltiplos hasta la velocidad primaria puede proporcionar los circuitos necesarios para satisfacer esta demanda.

⁶ Cuando se menciona la palabra menos pérdidas, se refiere a menos pérdidas matemáticamente: un algoritmo de compresión con menos pérdidas, es aquel que garantiza que la salida descomprimida es idéntica bit por bit a la señal de entrada original. Esto es mucho más estricto que "indistinguible visualmente" del original. El JPEG original puede lograr indistinguibilidad visual para la mayoría de las imágenes de tipo fotográfico pero nunca podrá tener realmente menos pérdidas.

- Que la existencia de diferentes jerarquías digitales y diferentes normas de televisión en diferentes partes del mundo complica los problemas relativos a la especificación de las normas de transmisión y codificación en las conexiones internacionales.

El objetivo básico del CCITT es recomendar soluciones únicas para las conexiones internacionales, reconociendo que el objetivo consiste en proporcionar interfuncionamiento entre los códecs a $n \times 384$ kbits/s y los códecs a $m \times 64$ kbits/s definidos en las recomendaciones de la serie H. El interfuncionamiento se hará tomando como base $m \times 64$ kbits/s; los valores de m están estudiándose.

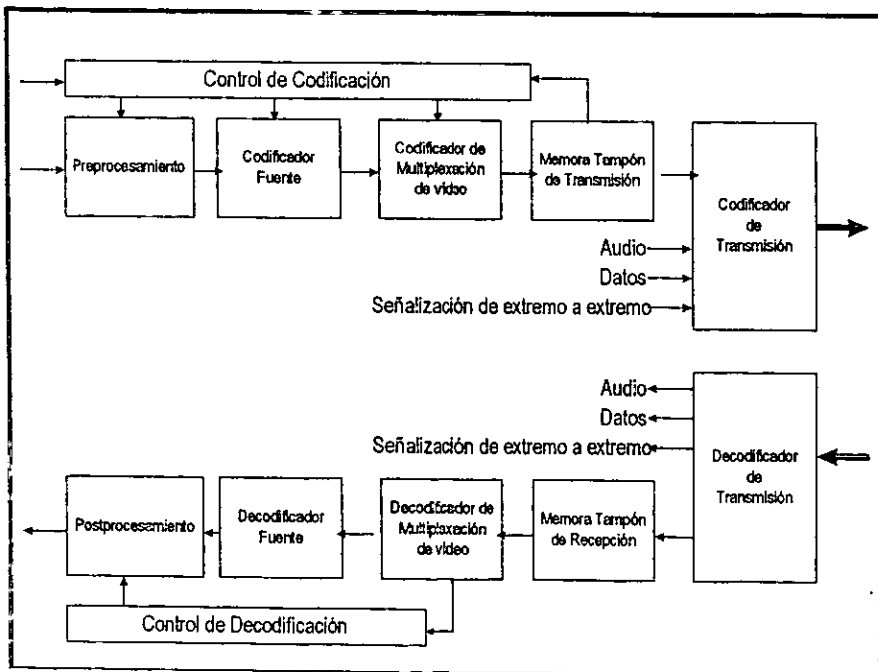


Figura 3.2.1. Diagrama de bloques resumido del códec.

Esta recomendación describe los métodos de codificación y decodificación de los servicios audiovisuales a las velocidades de $n \times 384$ kbits/s, donde n toma los valores de 1 a 5.

En la Figura 3.2.1. se muestra un diagrama de bloques resumido del códec.

Entrada y Salida de Vídeo

Para poder abarcar con una sola recomendación la utilización en las regiones de 625 y 525 líneas entre las mismas, las imágenes se codifican en un formato intermedio común. Las normas de señales de televisión de entrada y salida, que pueden, por ejemplo, ser compuestas o separadas en componentes, analógicas o digitales no son objeto de recomendación, como tampoco lo son los métodos para realizar cualquier conversión necesaria de y hacia el formato de codificación intermedio.

Frecuencia de muestreo

Las frecuencias se muestrean a un múltiplo entero de la velocidad de línea de vídeo. Este reloj de muestreo reloj de red digital son asincronos.

Algoritmo de codificación fuente

Se adopta una combinación de predicción interimágenes para utilizar redundancia temporal y codificación de la transformada de la señal restante para reducir la redundancia espacial. El decodificador tiene capacidad de compensación de movimiento, permitiendo la incorporación facultativa de esta técnica en el codificador.

Codificador Fuente

El codificador fuente opera con imágenes no entrelazadas que ocurren 30 000/1001 (aproximadamente 29,97) veces por segundo. La tolerancia en la frecuencia de imagen es de +/- 50 ppm.

Las imágenes se codifican mediante las componentes de luminancia y las dos de diferencia de color (Y, CR y CB). Estas componentes y los códigos que representan sus valores muestreados se definen en la Recomendación 601 del CCIR.

Negro = 16

Blanco = 235

Diferencia de color nula = 128

Diferencia de color máxima = 16 y 240

Estos valores son nominales y el algoritmo de codificación funciona con los valores de entrada comprendidos entre 0 y 255.

Para la codificación, la estructura de muestreo de luminancia es de 288 líneas por imagen, 352 elementos de imagen por línea en una disposición ortogonal. El muestreo de cada una de las dos componentes de diferencia de color es de 144 líneas, 176 elementos de imagen por línea, ortogonal. Las muestras de diferencia de color se sitúan de manera que sus fronteras de bloque coincidan con las fronteras de bloque de luminancia, como se muestra en la figura 3.2.2. La zona de imagen cubierta por estos números de elementos de imagen y líneas tiene una relación de 4:3 y corresponde a la porción activa de la entrada de vídeo de norma local ⁷

⁷ *Nota:* el número de elementos de imagen por línea es compatible con el muestreo de las porciones activas de las señales de luminancia y diferencia de color de fuentes de 525 ó 625 líneas a 6,75 y 3,375 Mhz respectivamente. Estas frecuencias tienen una relación simple con las de la Recomendación 601 del CCIR.



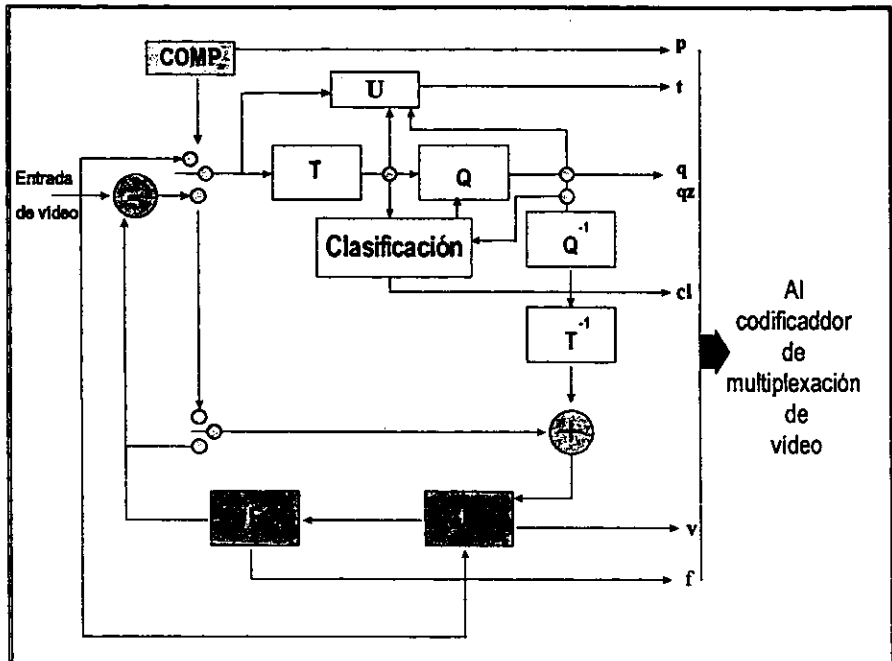
Figura 3.2.2 Posicionamiento de las muestras de luminancia y crominancia.

Algoritmo de codificación de fuente de video

El algoritmo de codificación de video se muestra en forma generalizada en la figura 3.2.3/H.261. Los principales elementos son la predicción, la transformación de bloque, la cuantificación y la clasificación.

El error de predicción (modo INTER) o la imagen de entrada (modo INTRA) se subdivide en 8 elementos de imagen por 8 bloques de línea que se segmentan como transmitidos o no transmitidos.

Los criterios de elección del modo y la transmisión de un bloque no son objeto de recomendación, y pueden variar dinámicamente como parte de la estrategia de control de la velocidad de datos. Los bloques transmitidos se transforman y los coeficientes resultantes se cuantifican y se codifican con longitud variable.



T	Transformación	p	Bandera de INTRA/INTER
Q	Cuantificador	t	Bandera de transmitido/No transmitido
I	Memoria de imagen con movimiento compensado con retardo variable	qz	Indicación de cuantificador
COMP	Comparador de INTRA/INTER	v	Vector de Movimiento
f	activación/desactivación del filtro de bucle	q	Índice de cuantificación para los coeficientes de la transformada
F	Filtro de bucle	cl	Índice de clasificación
u	Umbral		

Figura 3.2.3 Algoritmo de Codificador de vídeo

Predicción

La *predicción* es interimágenes y puede aumentarse por compensación del movimiento y mediante un filtro especial.

Compensación del Movimiento

La *compensación del movimiento* es facultativa en el codificador. El decodificador aceptará un vector para cada bloque de 8 elementos de imagen por 8 líneas. La gama de vectores permitidos está en estudio.

Un valor positivo de la componente horizontal o vertical del vector de movimiento significa que la predicción está formada a partir de elementos de imagen de la imagen anterior, situados espacialmente a la derecha o debajo de los elementos de imagen objeto de predicción.

Los vectores de movimiento están limitados de manera que todos los elementos de imagen por ellos referenciados estén dentro de la zona de imagen codificada.

Filtro de Bucle

El proceso de predicción puede modificarse mediante un filtro especial bidimensional que actúa sobre los elementos de imagen de un bloque predicho.

El filtro es separable en funciones unidimensionales horizontal y vertical. Ambas son no recursivas con coeficientes de $1/4$, $1/2$, $1/4$. En los bordes del bloque, donde uno de los puntos de toma caería fuera del bloque, el elemento de imagen periférico se utiliza para dos puntos de toma. La precisión aritmética se conserva totalmente redondeando a valores enteros de 8 bits en la salida del filtro bidimensional. Los valores de cuya parte fraccionaria es un medio se aproximan al valor superior.

El filtro puede activarse o desactivarse en cada bloque. Está estudiándose el método de señalar esto.

Dispositivo de transformación

Los bloques transmitidos se codifican mediante una transformada de coseno discreto bidimensional separable de 8 x 8. La entrada de la transformada directa y la salida de la transformada inversa tienen 9 bits. Los procedimientos aritméticos para computar las transformadas están en estudio.⁸

Cuantificación

El número de cuantificadores, sus características y asignación están en estudio.

Recorte

Para evitar las amplitudes de distorsión de cuantificación de coeficiente de transformada que causen desbordamiento aritmético en los bucles del codificador y decodificador, se insertan funciones de recorte. Además de las funciones de recorte aplicadas en la transformada inversa, se aplica una función de recorte, tanto en el codificador como en el decodificador, a la imagen reconstruida que se forma sumando la predicción y el error de predicción modificados por el proceso de codificación. Esta función de recorte actúa sobre los valores de elementos de imagen resultantes inferiores a 0 ó superiores a 255, cambiándolos a 0 y 255 respectivamente.

Control de Velocidad de Datos

Las secciones en que pueden variarse parámetros para controlar la velocidad de generación de datos de vídeo codificados incluyen procesamiento antes del codificador fuente, el cuantificador, criterio de significado de bloque y submuestreo temporal. Las proporciones de

⁸ **Nota:** La salida de la transformación directa y la entrada a la transformación inversa serán probablemente de 12 bits

tales medidas en la estrategia de control global no son objeto de recomendación.

Al ser invocado, el submuestreo temporal se realiza descartando imágenes completas. Las figuras interpoladas no se colocan en la memoria de imágenes.

Actualización Forzada:

Esta función se realiza forzando la utilización del modo INTRA del algoritmo de codificación. La estructura y el intervalo de actualización están en estudio.

Históricamente, *H.261* fue desarrollada mucho antes que JPEG. En diciembre de 1984, el Grupo de Estudio XV del CCITT (equipo y sistemas de transmisión) estableció un grupo de "Especialistas en codificación para telefonía visual". El desarrollo de éste estándar de transmisión de vídeo para una baja tasa de transmisión de servicios ISDN ha pasado por varios estados. Al principio la meta fue diseñar un esquema de transmisión para tasas de transmisión de $e(m)$ veces 384) kb/s, donde $e(m)$ estuvo entre 1 y 5. Posteriormente en tasas de $e(n)$ veces 64) kbps donde $e(n)$ de 1 a 5 fueron considerados. Sin embargo, luego fue hecha en 1989 la recomendación final del CCITT *H.26120*.

De hecho, la serie *H* de teleservicios audiovisuales, es un grupo de estándares (o recomendaciones) que consisten en:

H.221 estructura de trama.

H.230 control síncrono de trama.

H.242 comunicación entre terminales audiovisuales.

H.320 sistemas y equipo terminal.

H.261 códec de video.

Los audio códecs (aunque no son tema de la presente tesis) a varias tasas de bits, también fueron especificados en otras recomendaciones del CCITT como la G.725.

Los códecs *JPEG* línea base y H.261 utilizan ambas técnicas *DCT* (*Discrete Cosine Transform*) y *VLC* (*Variable-Length Code*). La principal diferencia entre el esquema de compresión *JPEG* y H.261 es que *JPEG* codifica cada trama individualmente, mientras que H.261 realiza codificación entre tramas. En H.261 la compensación de movimiento es realizada basada en bloques para registrar las diferencias entre tramas, las cuales son entonces codificadas en *DCT*. Aquí, el dato del cuadro en la trama anterior puede ser usado para predecir los bloques de imagen en la trama actual. Como resultado solamente las diferencias, típicamente de pequeña magnitud entre los bloques previamente desplazados y el bloque actual son transmitidos.

Hay algunas características interesantes o consideraciones de diseño en H.261.

- Primero, define esencialmente solo el **decodificador**. Sin embargo, el **codificador**, el cual no es completa y explícitamente especificado por el estándar, se espera sea completamente compatible con el decodificador.
- Segundo, porque H.261 es propuesto para comunicaciones en tiempo real, usa solamente la previa y más cercana trama como predicción para reducir el retardo de codificación.
- Tercero, trata de balancear las complejidades del hardware del codificador y del decodificador, dado que ambos son necesarios para aplicaciones de videoteléfono en tiempo real. Otros esquemas de codificación, como el vector de cuantificación (VQ), pueden tener

un decodificador un poco más simple, pero un codificador muy complejo.

- Cuarto, H.261 es un compromiso entre el funcionamiento del codificador, requerimiento de tiempo real, complejidad de implementación, y robustez del sistema. El movimiento compensado codificado en DCT es un algoritmo maduro, y después de años de estudio, bastante general y robusto que puede manejar varios tipos de cuadros.
- Quinto, las estructuras de codificación finales y parámetros son orientados más hacia aplicaciones de baja tasa de bit. Esta selección es lógica, porque la estructura del codificador y parámetros de codificación es más crítica para el desempeño del códec a muy bajas tasas de bit. A tasas de bit más altas, los valores de parámetros más bajos que el óptimo no afectan mucho el rendimiento del códec.

H.261 especifica un grupo de protocolos que comprime cada ráfaga de bits, y un grupo de operaciones que cada decodificador estándar compatible debe ser capaz de realizar. La actual implementación de hardware del códec y la estructura del codificador puede variar drásticamente de un diseñador a otro.

3.3 NORMA H.263 [UIT-T, 5 de Julio de 1995]

TRANSMISIÓN EN LÍNEA DE SEÑALES NO TELEFÓNICAS. Codificación de Vídeo para la Comunicación a Baja Velocidad Binaria.

Esta recomendación especifica una representación codificada, que se puede usar para comprimir el componente de imagen en movimiento de los servicios audiovisuales a velocidades binarias bajas.

El decodificador tiene capacidad de compensación de movimiento, lo que permite la incorporación opcional de esta técnica en el codificador. En la compensación de movimiento se utiliza la precisión de mitad de pixel, mientras que en la recomendación H.261 se utiliza la precisión de pixel entero y un filtro de bucle. Se emplea la codificación de longitud variable para la transmisión de los símbolos.

La Recomendación *UIT H.263*, es un híbrido de la predicción entre imágenes que utiliza la redundancia temporal y la codificación con transformada de la señal restante para reducir la redundancia espacial. La configuración básica del algoritmo de codificación de fuente de vídeo se basa en la *Recomendación UIT H.261*, a continuación en la figura 3.3.1 aparece el diagrama de bloques resumido del códec.

Para poder abarcar con una sola recomendación la utilización dentro y entre regiones que emplean normas de televisión de 625 y 525 líneas, el codificador de fuente actúa sobre imágenes basadas en un formato intermedio común (*CIF, Common intermediate format*).

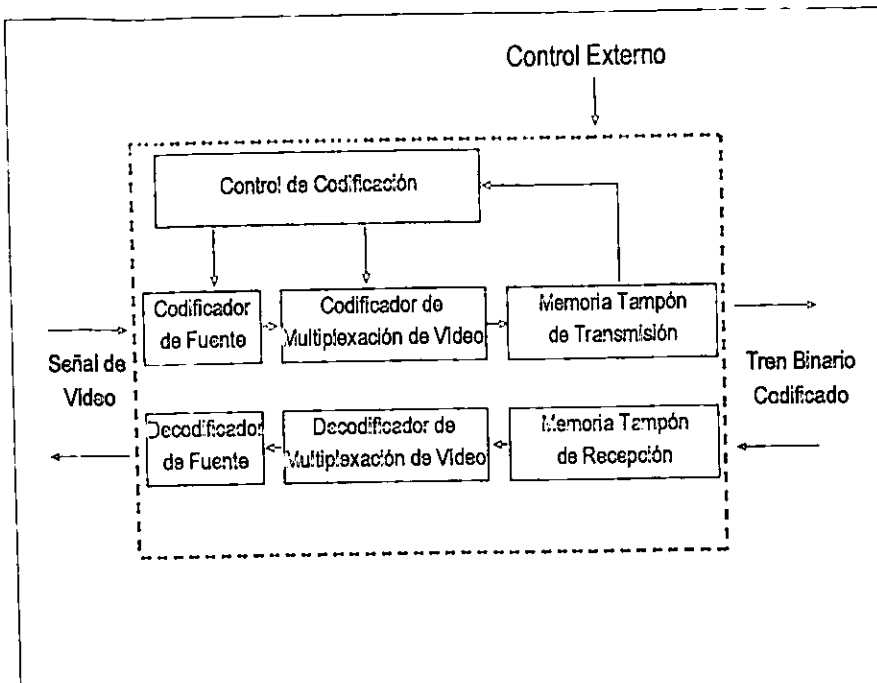


Figura 3.3.1 Diagrama de Bloques resumido del Códec de Vídeo.

El codificador de vídeo proporciona un tren binario digital auto contenido que se puede combinar con otras señales con multifacilidades, (como se define en la norma H.223). Y el codificador de vídeo efectúa el proceso inverso. Las imágenes se muestrean a un múltiplo entero de la frecuencia de línea de vídeo. El reloj de muestreo y el reloj de red digital son asíncronos. Además del algoritmo de codificación de fuente de vídeo básico, se incluyen cuatro opciones de codificación negociables para mejorar el funcionamiento, estas opciones se pueden utilizar juntas o por separado.

Vector de Movimiento sin restricción.

En este modo opcional, los vectores de movimiento están autorizados para apuntar fuera de la imagen. Los píxeles de borde se utilizan como predicción de los píxeles "no existentes". Con este modo se consigue una ganancia significativa cuando hay movimiento a través de los bordes de la imagen, especialmente con los formatos de imagen más pequeños. Además, este modo incluye una extensión de la gama de vectores de movimiento de modo que se puedan utilizar vectores de movimiento mayores, lo que es especialmente útil en el caso de movimiento de la cámara

Codificación Aritmética basada en Sintaxis

En este modo facultativo, se utiliza la codificación aritmética en vez de la codificación de longitud variable. La *relación señal a ruido (SNR)* y las imágenes reconstruidas serán las mismas, pero se producirá un número considerable menor de bits.

Predicción avanzada

En este modo opcional, se utiliza la compensación de movimiento de bloques superpuestos (*OBMC, overlapped block motion compensation*) para la parte de luminancia de las imágenes P. En algunos macrobloques de la imagen, se utilizan cuatro vectores de 8x 8 en vez de un vector de 16x16. El codificador decidirá el tipo de vector que se utilizará. Cuatro vectores utilizan más bits, pero dan una predicción mejor. La aplicación de este modo permite, en general un perfeccionamiento considerable. Más especialmente, se consigue una ganancia subjetiva porque OBMC resulta en menos dispositivos de bloqueo.

Trama PB.

Una trama PB consiste en dos imágenes que se codifican como una unidad. El nombre "PB" proviene del nombre de los tipos de imagen de la recomendación H.262, en la que hay imágenes P e imágenes B. Por consiguiente, una trama PB consiste en una imagen P, que se predice a partir de la anterior imagen P decodificada, y una imagen B, que se predice a partir de la anterior imagen P decodificada y la imagen P que está siendo decodificada en ese momento.

El reloj de transmisión se proporciona por medios externos. La velocidad binaria de vídeo puede ser variable. En esta recomendación no se imponen restricciones a la velocidad binaria de vídeo, estas restricciones las determinará la terminal o la red. El codificador controlará su tren binario de salida para cumplir los requisitos del decodificador ficticio de referencia. Los datos de vídeo se proporcionarán en cada ciclo de reloj válido. Esto se consigue mediante relleno de tipo de macrobloque y patrón de grupo modificado (*MCBCP, macroblock type and codec block pattern*) (véase las figuras 3.3.11 y 3.3.12) o, cuando se utiliza corrección de errores hacia adelante, también mediante tramas de relleno de corrección de errores hacia adelante.

El tratamiento de los errores se suministra por medios externos (recomendación H.223). Si no se proporciona por medios externos (por ejemplo, Recomendación H.221), se puede utilizar la alineación de trama y la corrección de errores opcional.

El decodificador puede enviar una instrucción para codificar uno o más GOB de la imagen próxima en modo INTRA con los parámetros de codificación adecuados para evitar el desbordamiento de la memoria tampón. Y también puede enviar una instrucción para que se

transmitan únicamente los encabezados GOB no vacíos. Estas señales se transmiten por medios externos por ejemplo, Recomendación H.245.

Codificador Fuente

El *codificador de fuente* trabaja con imágenes no entrelazadas que se presentan 30,000/1,001 (aproximadamente 29,97) veces por segundo. La tolerancia de frecuencia de imagen es de ± 50 ppm.

Las imágenes se codifican como una componente de luminancia y dos componentes de diferencia de color (Y, CR, CB). Estas componentes y los códigos que representan sus valores muestreados son los que define la recomendación 601 del CCIR.

Negro = 16

Blanco = 235

Diferencia de Color Nula = 128

Diferencia de Color Máxima = 16 y 240

Estos valores son nominales y el algoritmo de codificación funciona con valores de entrada comprendidos entre 1 y 254. Hay cinco formatos de imagen normalizados: sub-QCIF, QCIF, CIF, 4 CIF, y 16 CIF. Para cada uno de estos formatos de imagen, la estructura de muestreo de luminancia dx píxeles por línea y dy líneas por imagen en una disposición ortogonal. El muestreo de cada una de las dos componentes de diferencia de color se efectúa a $dx/2$ píxeles por línea y $dy/2$ líneas por imagen ortogonal. Los valores de dx , dy , $dx/2$ y $dy/2$ se ofrecen en la Tabla 3.3.1 para cada uno de los formatos de imagen.

En cada uno de los formatos de imagen, las muestras de diferencia de color se colocan de forma tal que sus límites de bloque coincidan con los límites de bloque de luminancia, como se muestra en la figura

3.3.2. La relación de aspecto de pixel en la misma para cada uno de estos formatos de imagen, es la misma que se define para QCIF y CIF en la recomendación H.261: $(4/3) \cdot (288/352)$. El área de la imagen cubierta por todos los formatos de imagen, salvo el sub-QCIF, tiene una relación de aspecto 4:3.

Formato de Imagen	Número de Pixeles de Luminancia dx	Número de líneas de Luminancia Dy	Número de Pixeles de Crominancia dx/2	Número de líneas de Crominancia dy/2
sub-QCIF	128	96	64	48
QCIF	176	144	88	72
CIF	352	288	176	144
4 CIF	704	576	352	288
16 CIF	1408	1152	704	576

Tabla 3.3.1 número de pixeles por línea y números de líneas de cada formato de imagen de H.263.

Todos los decodificadores deberán poder funcionar con sub-QCIF. Los medios externos, como la recomendación H.245 indican los formatos que puede manipular el codificador. Para un panorama completo de los posibles formatos de imagen y los algoritmos de codificación de vídeo, por ejemplo la recomendación H.324.

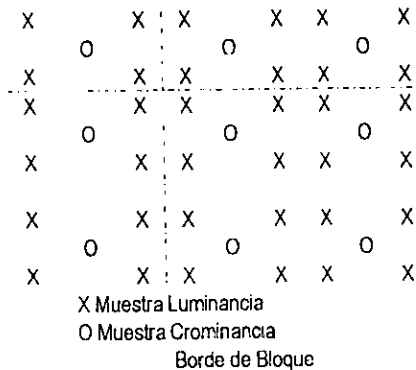
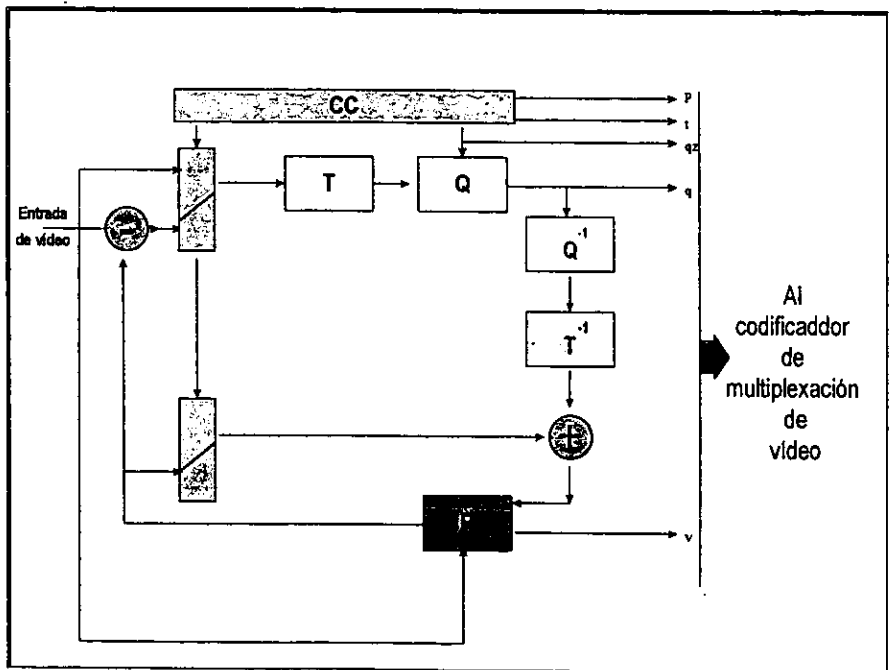


Figura 3.3.2 Posición de las muestras de luminancia y crominancia.

Colocando un número mínimo de imágenes no transmitidas, entre las imágenes transmitidas se proporcionarán los medios para restringir el periodo máximo de transmisión de imagen de los codificadores.

Algoritmo de codificación de Fuente de Video.

El codificador de fuente se muestra de forma generalizada en la figura 3.3.3. Los principales elementos son la predicción, la transformación de bloques y la cuantificación.



T	Transformación	p	Bandera de INTRA/INTER
Q	Cuantificación	t	Bandera de transmitido/No transmitido
P	Memoria de imagen con retardo variable y compensación de movimiento	qz	Indicación de cuantificador
CC	Control de Codificación	v	Vector de Movimiento
		q	Índice de cuantificación para los coeficientes de la transformada

Figura 3.3.3 Codificador de fuente.

Grupos de Bloques y Macrobloques.

Cada imagen se divide en *grupos de bloques* (GOB, *Group of Blocks*), Un Grupo de bloques comprende $k^2 \cdot 16$ líneas, según el formato de imagen $k=1$, para sub-QCIF, QCIF y CIF, $k=2$, para 4CIF $k=4$, para 16CIF. El número de GOB por imagen es de 6 para sub-QCIF, 9 para QCIF y 18 para CIF, 4 CIF y 16 CIF.

Cada GOB se divide en *macrobloques*. Un macrobloque se relaciona con 16 píxeles por 16 líneas de Y, y con 8 píxeles 8 líneas correspondientes espacialmente de C_B y C_R . Además, un macrobloque consiste en cuatro bloques de luminancia y los dos bloques de diferencia de color correspondientes espacialmente, como se muestra en la figura 3.3.4. Cada bloque de luminancia o de crominancia se relaciona con 8 píxeles por 8 líneas de Y, C_B o C_R .

La numeración de macrobloques se efectúa mediante un barrido horizontal de las filas de macrobloque de izquierda a derecha, iniciando por la línea de macrobloque superior y terminando por la línea de macrobloque inferior. Los datos de macrobloque se transmiten por macrobloque en orden de macrobloque creciente.

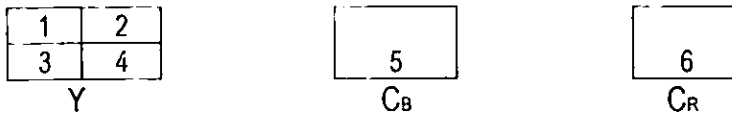


Figura 3.3.4 Disposición de los bloques en un macrobloque.

Los criterios de elección del modo y la transmisión de un bloque no son objeto de recomendación y pueden variar dinámicamente como parte de la estrategia de control de la codificación. Los bloques transmitidos

se transforman y los coeficientes resultantes se cuantifican y se codifican en entropía.

Predicción

La predicción se efectúa entre imágenes y se puede aumentar por compensación de movimiento. El modo de codificación en el que se aplica la predicción se denomina INTER, en modo de codificación se denomina INTRA cuando no se aplica la predicción. El modo de codificación INTRA se puede señalar en el nivel de la imagen (INTRA para imágenes I o INTER para imágenes P) o el nivel de macrobloque en las imágenes P. En el modo de trama PB opcional, las imágenes B están siempre codificadas en modo INTER. Las imágenes B se predicen parcialmente de forma bidireccional.

Compensación de Movimiento

El codificador aceptará un vector por macrobloque o, si se utiliza el modo de predicción avanzada de H.263, aceptará uno o cuatro vectores por macrobloque para la adaptación de vectores de movimiento destinado a la predicción de macrobloque B.

Un valor positivo del componente horizontal o vertical del vector significa que la predicción se forma a partir de los píxeles de la imagen referenciada que están espacialmente a la derecha o por debajo de los píxeles que se predicen.

Los vectores de movimiento están restringidos de modo que todos los píxeles referenciados por ellos se encuentren dentro de la zona de imagen codificada, salvo cuando se utiliza el modo de vector de movimiento sin restricción y/o el modo de predicción avanzada.

Quantificación.

El número de cuantificadores es 1 para el primer coeficiente de bloque INTRA y 31 para todos los demás coeficientes, excepto el primero de los bloques INTRA. No se definen los niveles de decisión. El primer coeficiente de los bloques INTRA es, nominalmente, el valor de la transformada en continua (cc) cuantificado uniformemente con un paso de valor 8. Cada uno de los restantes 31 cuantificadores utilizan niveles de reconstrucción de espaciado igual con una zona muerta central alrededor de cero y un paso de valor par entre 2 y 62.

Control de la cuantificación

Se pueden cambiar varios parámetros para controlar la velocidad de generación de datos de vídeo codificado. Se trata del proceso efectuado antes del proceso del codificador de fuente, el cuantificador, el criterio de significación de bloque y el submuestreo temporal. Las proporciones de estas medidas en la estrategia de control global no son objeto de esta recomendación. Al ser invocado, el muestreo temporal se ejecuta mediante el descarte de imágenes completas.

Un decodificador puede indicar su preferencia por un compromiso específico entre la resolución espacial y la temporal de la señal de vídeo. El codificador señalará su compromiso por defecto al comienzo de la llamada e indicará si es capaz de responder a las solicitudes del decodificador de cambiar este compromiso. Estas señales se transmiten por medios externos, como se puede observar en la recomendación H.245.

Actualización Forzada

Esta función se realiza forzando la utilización del modo INTRA del algoritmo de codificación. El esquema de actualización no se define.

Para controlar la acumulación de errores por desajuste de la transformada inversa, cada macrobloque se codificará en modo INTRA al menos una vez cada 132 veces cuando se transmiten los coeficientes de este macrobloque en imágenes P.

Alineación en byte de los códigos de comienzo

La *Alineación en Byte de los Códigos de Comienzo* se consigue insertando una palabra de código de menos de 8 bits cero antes del código del comienzo, de manera que el primer bit del código del comienzo es el primer bit más significativo de un byte. Por consiguiente un código de comienzo esta alineado en byte si la posición de su bit más significativo es un múltiplo de 8 bits del primer bit del tren binario de video. Es obligatorio que los códigos de comienzo de todos los tipos de imagen estén alineados en byte, y los códigos de comienzo de GOB y EOS no están sujetos a esta obligación.

Sintaxis y Semántica

El múltiplex de video tiene una estructura jerárquica de cuatro capas, de arriba hacia abajo, esas capas son:

- Imagen
- Grupo de Bloques
- Macrobloque
- Bloque

A continuación en la figura 3.3.5 se muestra el diagrama de sintaxis, y a menos que se especifique otra cosa, el bit más significativo se transmite primero. Es el bit 1 situado en el extremo izquierdo de las tablas de bits de esta recomendación. A menos que se especifique otra cosa, todos los bits no utilizados o de reserva se fijan a 1. Los bits de reserva no se utilizarán hasta que la *UIT* no especifique sus funciones.

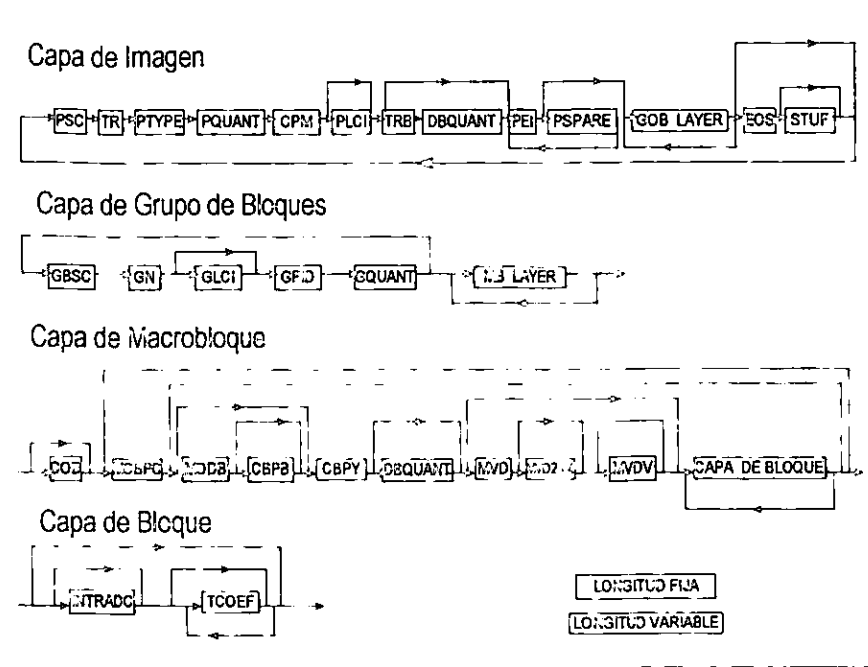


Figura 3.3.5 Diagrama de Sintaxis del tren binario de Vídeo.

Capa de Imagen

Los datos de cada imagen consisten en un encabezamiento de imagen, seguidos de datos de grupo de datos, y éstos a su vez seguidos de un código de fin de secuencia y bits de relleno. La estructura aparece En la figura 3.3.6. PSBI solo está presente si así lo indica CPM. TRB y BDQUANT, solo están presentes si PTYPE indica "trama PB". Las combinaciones de PSPARE y PEI pueden no estar presentes. EOS puede no estar presente, mientras que STUF puede estar presente solo si EOS esta presente, los encabezamientos de imagen de las imágenes descartadas no se transmiten.

PSC	TR	PType	BOQUANT	CMP	PSBI	TRB	OBOQUANT	PEI	PSPARE	PEI	Grupo de Bloques	STUF	EOS	PSTUF
-----	----	-------	---------	-----	------	-----	----------	-----	--------	-----	------------------	------	-----	-------

Figura 3.3.6 Estructura de la capa de Imagen

Capa de Grupos de Bloques

Los datos de cada grupo de bloques (GOB) consiste en un encabezamiento GOB, seguido de datos de macrobloque. La estructura se muestra en la figura 3.3.7. Cada GOB contiene una o más filas de macrobloques

GBSC	GN	GLCI	GFID	GQUANT	Datos de Macrobloque
------	----	------	------	--------	----------------------

Figura 3.3.7 Estructura de la capa GOB.

Capa de Macrobloque

Los datos de cada macrobloque consisten en un encabezamiento de macrobloque, seguidos de los datos de bloque. La estructura aparece en la figura 3.3.8.

COD	MCBPC	MODB	CBPY	DQUANT	MVD	MVD ₂	MVD ₃	MVD ₄	MVDB	Datos de Bloque
-----	-------	------	------	--------	-----	------------------	------------------	------------------	------	-----------------

Figura 3.3.8 Estructura de la capa de Imagen.

Capa de Bloque

Si no se utiliza el modo de tramas PB, un macrobloque comprende cuatro bloques de luminancia y uno de cada uno de los bloques de diferencia de color. La estructura de la capa de bloque se muestra en la figura 3.3.9. INTRADC esta presente para cada bloque del macrobloque si MCBPC indica MB tipo 3 ó 4 como se puede observar en las tablas 3.3.2 y 3.3.3 TCOEF está presente si es indicado por MCBPC Y CBPY.

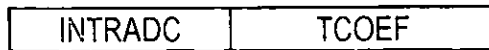


Figura 3.3.9 Estructura de la capa de Bloque.

Índice	Tipo de macrobloque	CBPC (56)	Número de Bits	Código
0	3	00	1	1
1	3	01	3	001
2	3	10	3	010
3	3	11	3	011
4	4	00	4	0001
5	4	01	6	0000 01
6	4	10	6	0000 10
7	4	11	6	0000 11
8	relleno	----	9	0000 0000 1

Tabla 3.3.2 Tabla de Códigos de longitud variable (VCL) para MCBPC (imágenes INTRA).

Índice	Tipo de Macrobloque	CBPC (56)	Número de Bits	Código
0	0	00	1	1
1	0	01	4	0011
2	0	10	4	0010
3	0	11	6	0001 01
4	1	00	3	011
5	1	01	7	0000 111
6	1	10	7	0000 110
7	1	11	9	0000 0010 1
8	2	00	3	010
9	2	01	7	0000 101
10	2	10	7	0000 100
11	2	11	8	0000 0101
12	3	00	5	0001 1
13	3	01	8	0000 0100
14	3	10	8	0000 0011
15	3	11	7	0000 011
16	4	00	6	0001 00
17	4	01	9	0000 0010 0
18	4	10	9	0000 0001 1
19	4	11	9	0000 0001 0
20	relleno	----	9	0000 0000 1

Tabla 3.3.3 De Códigos de Longitud Variable (VCL) para MCBPC (imágenes INTER).

3.4. MPEG.

Imágenes en movimiento.

Muchas imágenes digitales contienen un alto grado de redundancia, que mediante una técnica eficiente de compresión puede reducir significativamente la cantidad de información necesaria para ser almacenada o transmitida. Esta redundancia puede ser encontrada entre píxeles simples, entre líneas, entre tramas, cuando la escena es estacionaria o con movimiento suave.

La Organización Internacional para Estandarización (ISO) ha intentado hacer un esfuerzo por desarrollar un estándar para almacenamiento en medios digitales de video y su audio asociado, donde el concepto de medio de almacenamiento digital incluye dispositivos de almacenamiento convencionales CD-ROM, DAT, unidades de la cinta, winchester, dispositivos ópticos escribibles, así como canales de telecomunicación como ISDN's y redes de área local.

Este esfuerzo es conocido con el nombre de Grupo de Expertos de Imágenes en Movimiento MPEG (Moving Picture Experts Group) y es parte de la ISO-IEC/ JTC1/ SC2/ WG11. Las actividades de la MPEG empezaron en 1988 y cubren además de la compresión de video, la compresión del audio asociado. MPEG-Video se dirige a la compresión de señales de video alrededor de 1.5 Mbits/s, MPEG-Audio se dirige a la compresión de una señal digital de audio con tasa de 64, 128 y 192 Kbits/s por canal. Nosotros nos enfocaremos al MPEG-Video. La premisa de MPEG es que una señal de video y el audio asociado pueden ser comprimidos a una tasa de bit de 1.5 Mbits/s con una calidad aceptable.

Aplicaciones de la compresión de video en medios de almacenamiento digital.

o Medios de almacenamiento digital.

Muchos medios de almacenamiento y canales de telecomunicación son perfectamente adaptados a las técnicas de compresión de video en las tasas de 1 a 1.5 Mbits/s. El CD-ROM es un medio muy importante de almacenamiento de una gran capacidad y bajo costo. La cinta de audio digital (DAT) es también apropiado para la compresión de video; la grababilidad o registro del medio es una ventaja, pero tiene inconvenientes cuando es requerido el acceso

aleatorio. El disco de computadora tipo Winchester proporciona el máximo de flexibilidad (grababilidad o registro y acceso aleatorio) pero tiene un alto costo y limitada portabilidad. El disco óptico escribible combina las ventajas de los otros medios (almacenamiento o registro, acceso aleatorio, portabilidad y bajo costo).

- **Aplicaciones Asimétricas.**

Las aplicaciones asimétricas requieren el uso frecuente de procesos de descompresión, aún cuando el proceso de compresión es realizado una sola vez, como por ejemplo la publicidad electrónica, educación, entretenimiento (películas), juegos de vídeo y videotexto.

- **Aplicaciones simétricas.**

Las aplicaciones simétricas requieren esencialmente igual uso de procesos de compresión y descompresión. En las aplicaciones simétricas tenemos al videoteléfono, la videoconferencia, el vídeo mail, y la producción de publicidad electrónica.

- **Características del algoritmo de compresión de vídeo.**

Las siguientes características han sido identificadas como importantes para conocer la necesidad de las aplicaciones de MPEG.

- **Acceso aleatorio.**

El acceso aleatorio es una característica esencial para el vídeo en un medio de almacenamiento si el medio es o no, un medio de acceso aleatorio tal como un CD (Compact Disc) o un disco magnético. O un medio secuencial como una cinta magnética. El acceso aleatorio implica la existencia de puntos de acceso, como por

ejemplo segmentos de información codificados solamente con referencia a ellos mismos. Se debe alcanzar un tiempo de acceso aleatorio de aproximadamente $\frac{1}{2}$ segundo sin una significativa degradación de la calidad.

- ***Búsqueda rápida hacia adelante/hacia atrás.***

Dependiendo del medio de almacenamiento, debe ser posible buscar el flujo de bits comprimidos (posiblemente con la ayuda de una estructura de directorio de una aplicación específica) y usando los puntos de acceso apropiados, desplegando la imagen seleccionada para obtener un efecto rápido hacia adelante o rápido hacia atrás. Esta característica es esencialmente una forma más exigente de accesibilidad aleatoria.

- ***Sincronización audiovisual.***

La señal visual debe ser sincronizada al audio asociado. Un mecanismo debe proporcionar la permanente resincronización, las dos señales de audio y vídeo deben ser derivadas de relojes un poco distintos. Esta característica es proporcionada por el grupo de sistemas de MPEG cuyas tareas son el definir las herramientas de sincronización para una mejor integración de múltiples señales de vídeo y audio.

- ***Robustez contra los errores.***

Muchos medios de almacenamiento digital y canales de comunicación no están libre de errores, y mientras se espera que un apropiado esquema de codificación de canal deba ser usado para muchas aplicaciones, el esquema de codificación fuente debe ser robustecido para los errores remanentes sin corregir.

- **Retardo codificación/decodificación.**

Como se mencionó anteriormente, aplicaciones tales como videotelefonía necesitan mantener en el sistema un retardo total por debajo de 150 ms, para mantener la conversación "cara a cara" natural de la aplicación. El algoritmo debe llevarse a cabo a través de rangos de retardo aceptables, y el retardo debe ser considerado como un parámetro.

- **Editabilidad.**

Ya que toda imagen no puede ser comprimida independientemente (como por ejemplo imágenes fijas), es deseable poder estructurar unidades de edición de corta duración y codificadas sólo con referencia a ellas, de manera que se obtenga un nivel aceptable de editabilidad en forma comprimida.

- **Flexibilidad de formato.**

El paradigma de la computadora de "vídeo en una ventana" supone una gran flexibilidad de formatos en términos de tamaño de trama (ancho, alto) y velocidad de trama.

- **Costo de manufactura.**

Todas las propuestas de soluciones algorítmicas fueron evaluadas para verificar que un decodificador se puede implementar en un número pequeño de chips, con la tecnología de 1990. El algoritmo propuesto también tenía que encontrar la restricción del proceso de codificación que debería ser realizado en tiempo real.

MPEG1 tiene como objetivo la compresión de vídeo de 320 x 240, de movimiento total y optimizado para aplicaciones de almacenamiento en CD-ROM, multimedia interactiva y es codificado aproximadamente a una tasa de bits de 1.5 Mbit/s.

MPEG2 provee aplicaciones con requerimientos de alta calidad (resoluciones altas) de video tal como transmisión de TV, comunicación de video a través de redes ISDN usando ATM. MPEG2 soporta formatos de video entrelazados y un número de características para HDTV. Esta permite ambas secuencias entrelazadas y progresivas. El decodificador MPEG2 debe poder decodificar también MPEG1.

MPEG3 fue pensada para aplicaciones de HDTV (televisión de alta definición) y ofrecía codificación a una tasa de bits entre 20 y 40 Mbit/s. Se descubrió más tarde que con algunas mejoras, la sintaxis de MPEG2 funcionaba muy bien para HDTV. De manera que HDTV es ahora parte de la norma MPEG2.

MPEG4 tiene como destino aplicaciones a muy baja tasa de bits como en el videoteléfono. La tasa de bits de codificación está entre 4,800 y 6,4000 bit/seg. Las actividades de MPEG4 se instalaron en Septiembre de 1993. MPEG4 está actualmente en fase de identificación de aplicaciones.

3.4.1. MPEG1

La dificultad del desafío en el diseño del algoritmo MPEG es el siguiente: Por una parte, los requerimientos de calidad demandan una alta compresión no alcanzada solamente con la codificación dentro de las tramas (intraframe); y por otra parte los requerimientos de acceso aleatorio se satisfacen mejor con codificación dentro de las tramas. El algoritmo puede satisfacer los requerimientos en la medida que alcance la alta compresión asociada con la codificación entre las tramas (interframe), la cual no cumple con el compromiso de acceso

aleatorio que se demanda. Esto requiere un balance entre codificación dentro y entre las tramas, y reducción de la redundancia temporal entre la recurrente y la no recurrente. Para dar respuesta a éste desafío, los miembros de la MPEG han acudido al uso de dos técnicas de codificación entre tramas: Interpolada y predictiva.

El algoritmo de compresión de vídeo cuenta con dos técnicas básicas: compensación de movimiento para la reducción de la redundancia temporal y la compresión basada en DCT para la reducción de la redundancia espacial. La técnica de compensación de movimiento es aplicada para ambas; a la causal (codificación predictiva) y la no causal (codificación interpolada). La señal restante (el error de predicción) es comprimida con reducción de la redundancia espacial (DCT). La información de movimiento esta basada en bloques de 16x16 y es transmitida junto con la información espacial. La información de movimiento es comprimida usando códigos de longitud variable para alcanzar la máxima eficiencia.

Las actividades de MPEG iniciaron en 1988 para aplicaciones de CD-ROM a una tasa de 1.5 Mb/s. La entrada recomendada del tamaño del cuadro en la norma MPEG es de 360x240 pixeles para luminancia y 180x120 para crominancia, la velocidad de tramas es de 29.97 tramas por segundo para la compatibilidad con el sistema NTSC. El esquema de codificación de MPEG1 es similar al H.261, la principal diferencia entre los dos es que MPEG1 permite compensación de movimiento bidireccional. Tres tipos de tramas en una secuencia de vídeo son codificadas en tres diferentes algoritmos en MPEG. (Le Gall, 1991).

1. Imagen intra (trama 'I').

Estas imágenes son codificadas usando técnicas basadas en DCT similares a la JPEG, y también son codificadas por sí mismas sin tomar en cuenta a las otras. Cada imagen es descompuesta en bloques de 8x8 píxeles, cada uno de éstos son codificados sin tomar en cuenta a los otros bloques y son enviados directamente al proceso de transformación por bloque. Las tramas 'I' proporcionan puntos de acceso para el acceso aleatorio, y dan baja relación de compresión.

2. Imagen predictiva (trama 'P').

Estas imágenes son codificadas con referencia a la trama pasada ('I' o 'P'), es decir la actual trama es codificada con referencia a la trama previa ('I' o 'P'), y en general son usadas como una referencia para la futura imagen a predecir.

3. Imagen bidireccional o imagen interpolada (tramas 'B').

Estas son imágenes codificadas usando dos tramas de referencia de una trama pasada y una trama futura (la cual puede ser una trama 'P' o una trama 'I'). Una trama 'B' no puede ser usada como imagen de predicción de referencia.

En todos los casos cuando una imagen es codificada con respecto a una referencia, la técnica de compensación de movimiento es usada para proporcionar la eficiencia de la codificación. La relación entre los tres tipos de tramas se ilustra en la figura 3.4.1.1. La organización de las tramas en MPEG es flexible y dependerá de la aplicación específica, y de parámetros tales como accesibilidad aleatoria y retardo de codificación.

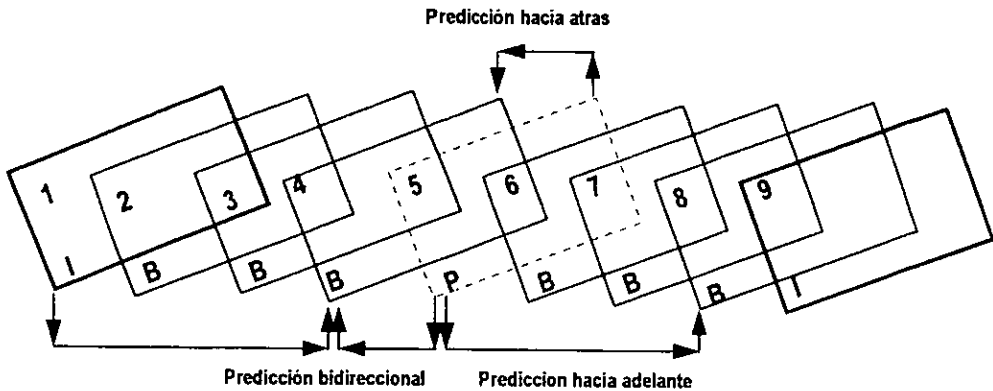


Figura 3.4.1.1 Tipos de tramas en la norma MPEG.

Note que en la figura 6, las primeras tres tramas B (2, 3 y 4) son codificadas bidireccionalmente usando la trama pasada 'I' (trama 1), y la trama futura 'P' (trama 5). Por consiguiente, el orden de la decodificación tendrá una diferencia del orden de la codificación. La trama 'P' (5) debe ser decodificada antes que las tramas 'B' (2, 3 y 4) y la trama 'I' (9) antes que las tramas 'B' (6, 7 y 8). Si la secuencia es transmitida a través de una red, el orden actual de transmisión será: {1, 5, 2, 3, 4, 9, 6, 7, 8}

La aplicación MPEG determina la secuencia de las tramas 'I', 'P' y 'B'. Esto es necesario para un rápido acceso aleatorio, la mejor resolución se alcanzará codificando la secuencia entera de las tramas 'I'. Por lo tanto, la mas alta razón de compresión puede ser alcanzada por la incorporación de un largo numero de tramas 'B'. [Borko furht, 1995]

Compensación de movimiento.

○ *Predicción.*

De las técnicas que explotan la redundancia temporal de la señal de vídeo, la más ampliamente usada es la predicción de compensación de movimiento. Esta es la base para muchos algoritmos de compresión para la telefonía visual como se especifica en la norma H.261. Predicción de compensación de movimiento asume que "localmente" la trama presente puede ser modelada como una translación de la imagen a algún tiempo previo.

Localmente significa que la amplitud y dirección de los desplazamientos no necesariamente deben ser los mismos en todas las partes del cuadro. La información de movimiento es parte de la información necesaria para recuperar la imagen y debe ser codificada apropiadamente.

○ *Interpolación...*

La interpolación de compensación de movimiento es una característica importante de MPEG. Esta técnica ayuda a satisfacer algunos de las aplicaciones que dependen de los requerimientos, ya que proporciona acceso aleatorio y reduce el efecto de errores y al mismo tiempo contribuye significativamente a la calidad de la imagen.

En la dimensión temporal, la interpolación de compensación de movimiento (también llamada predicción bidireccional en la terminología de MPEG) es una técnica multiresolución: Una subseñal con baja resolución temporal (típicamente 1/2 ó 1/3 de la velocidad de tramas) es codificada, y la resolución total de la señal es obtenida por interpolación de la señal de más baja resolución adicionando los términos de corrección. La señal a reconstruir por

interpolación es obtenida por la adición de términos de corrección de una combinación de una referencia pasada y una futura. La compresión obtenida por codificación interpolativa es muy alta.

- **Representación del movimiento:**

La información de movimiento consiste de un vector de predicción hacia adelante y hacia atrás, y dos vectores para predicción bidireccional. La información de movimiento asociada con cada bloque de 16x16 es codificada diferencialmente, con respecto a la información de movimiento presente en el bloque previo adyacente. El rango del vector de movimiento diferencial puede ser seleccionado cuadro por cuadro; para igualar la resolución espacial, temporal y la naturaleza del movimiento en una secuencia en particular. La información de movimiento diferencial es además codificada por medio de códigos de longitud variable para proporcionar gran eficiencia.

- **Estimación de movimiento.**

La estimación de movimiento cubre un conjunto de técnicas usadas para extraer la información de movimiento de una secuencia de vídeo. La sintaxis de MPEG especifica cómo representar la información de movimiento: uno o dos vectores de movimiento por 16x16 sub-bloques de la imagen dependiendo del tipo de compensación de movimiento: predicción hacia adelante, hacia atrás o un promedio. El MPEG no especifica como se calculan los vectores. Debido a la representación de movimiento basado en bloques, sin embargo se usan técnicas de block matching. (Le Gall, 1991)

El proceso de codificación para las tramas 'P' y 'B' incluye la estimación de movimiento, que es usado para extraer la información de una secuencia de vídeo, el cual encuentra el mejor block

matching en la trama de referencia. Las tramas 'P' usan siempre predicción hacia delante; las tramas 'B' usan predicción bidireccional, también llamado interpolación de compensación de movimiento, es decir un bloque en la trama presente (trama 'B') puede predecirse a partir de otro bloque de la trama de referencia pasada, o de la trama futura de referencia o por el promedio de los dos bloques (interpolación).

◦ **Transformación de bloque.**

En transformación de bloque, los bloques INTRA codificados así como los errores de predicción son procesados por una función DCT de dos dimensiones.

◦ **Cuantificación.**

La cuantificación de los coeficientes DCT es una operación importante. Debido a la combinación de cuantificación y la codificación de longitud variable, contribuyen a más compresión. El propósito de este paso es llevar a cabo más compresión para la representación de los coeficientes de DCT con la precisión necesaria para lograr la calidad requerida.

◦ **Códigos de longitud variable.**

Aquí se hace compresión extra para asignarles una codificación de palabras más corta para eventos frecuentes, y una codificación de palabras más larga para eventos menos frecuentes. La codificación de Huffman es usualmente usada para la implementación de este paso.

Capas de la estructura de sintaxis del flujo de bits

La meta de esta estructura en capas es separar las entidades en el flujo de bits para distinguirlos lógicamente, y facilitar el proceso de

decodificación. La sintaxis de MPEG vídeo de un flujo de bits contiene seis capas (ver tabla 3.4.1.1); cada capa soporta una función definida: cualquiera de las funciones de procesamiento de la señal (DCT, Compensación de Movimiento) o una función lógica (resincronización, puntos de acceso aleatorio). [Le Gall, 1991]

Seis capas de la sintaxis del flujo de bit de MPEG	
Capa de secuencia de imágenes	(Unidad de acceso aleatorio: contexto)
Capa de grupo de imágenes	(Acceso aleatorio: Codif. de vídeo)
Capa de cuadro	(Unidad de codificación primaria)
Capa de slice	(Unidad de sincronización)
Capa de macrobloque	(Unidad de compensación de movimiento)
Capa de bloque	(unidad de DCT)

Tabla 3.4.1.1 Sintaxis MPEG de vídeo de un flujo de bits.

- **Secuencia de vídeo.**

Esta es la estructura más alta de la codificación del flujo de bits. Se puede considerar como una unidad de acceso aleatorio. Una secuencia de vídeo consiste de grupos de cuadros o imágenes (GOPs).

- **Grupo de cuadros o tramas (GOPs).**

Esto es opcional en MPEG2. Aquí se tienen tres posibles tipos de cuadros o tramas: trama 'I', trama 'P' y trama 'B'. Los grupos de tramas ('I', 'P' y trama 'B') pueden tener una estructura y una longitud arbitraria. El primer cuadro en la codificación del flujo de bits debe ser una trama 'I'.

◦ **Cuadro o trama.**

Este correspondería a un simple cuadro en la secuencia de vídeo. Para el cuadro de campo en secuencias entrelazadas, el cuadro entrelazado se representa por dos cuadros separados en el flujo de codificación. Este puede ser codificado en el mismo orden que viene de la salida del decodificador.

◦ **Slice.**

Este corresponde a un grupo de macrobloques. El número actual de macrobloques dentro de un slice no está sujeto a una estandarización.

El slice no necesita cubrir el cuadro completamente. Es un requisito que si el cuadro fue usado para las siguientes predicciones, entonces las predicciones deben pertenecer únicamente a esas regiones de la imagen que fue incluida en los slices.

◦ **Macrobloque.**

La unidad fundamental del vídeo es el macrobloque. Un macrobloque es un grupo de 16x16 píxeles de vídeo. Un macrobloque contiene una sección de la componente de luminancia y la correspondiente componente de crominancia. Consiste entonces de 16x16 píxeles de luminancia, y 2 bloques de 8x8 píxeles de crominancia. Los 16x16 píxeles de luminancia son divididos en cuatro bloques, de modo que un macrobloque consiste de seis bloques de 8x8 componentes de píxel. MPEG soporta tres formatos. El formato 4:2:0 consiste de un macrobloque de 6 bloques (4Y, 1Cb, 1Cr), el formato 4:2:2 que es un macrobloque consistente de 8 bloques. (4Y, 2Cb, 2Cr), y el formato 4:4:4 que es un macrobloque consistente en 12 bloques (4Y, 4Cb, 4Cr).

• **Bloque.**

Un bloque es un término usado para describir cualquier bloque de 8x8 componentes de pixeles.

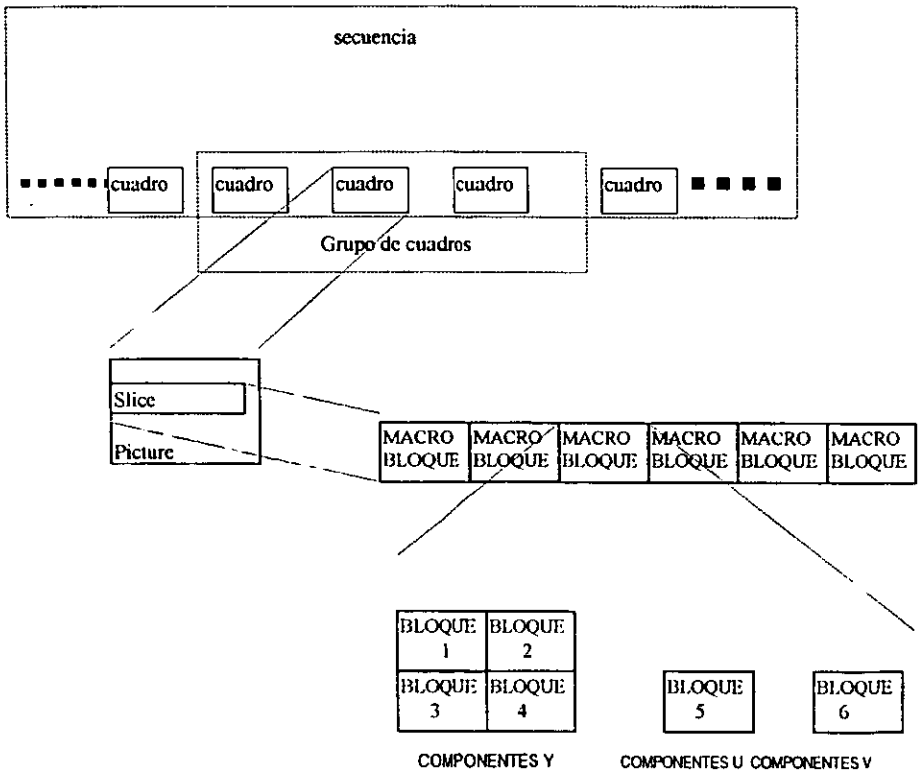


Figura. 3.4.1.2. Jerarquía de datos de video.

El video es codificado jerárquicamente. Una secuencia esta hecha de grupos de cuadros. Un grupo de cuadros esta formado de cuadros individuales. Cada cuadro corresponde a una simple trama de video en movimiento. Un cuadro está dentro de un slice. Cada slice en turno esta compuesto de macrobloques. Los macrobloques están

compuestos de seis bloques de 8x8 componentes de pixel. La codificación de los datos de vídeo ocurre bloque por bloque contenidos en un macrobloque, y es codificado en el orden enumerado en la figura 3.4.1.2. [Motorola,1995]

MPEG soporta secuencias de vídeo entrelazadas y progresivas. La salida del proceso de decodificación, para secuencias entrelazadas, consiste de una serie de campos que son separados en el tiempo por un período de campo. Los dos campos de una trama pueden ser codificados independientemente (cuadro-campo) o pueden ser codificados juntos como una trama (trama de imagen).

El MPEG1 solamente especifica la sintaxis del flujo de bits y el proceso de decodificación. A una tasa de bit dada, son posibles diferentes esquemas de codificación. Se ha demostrado que a 1.5 Mb/s, la calidad del vídeo es comparable a la de una VCR. Además se espera mejorar la calidad gracias a la flexibilidad de la codificación. [Borko,1995]

3.4.2. MPEG2

La norma MPEG1 puede proporcionar un rango amplio de tasa de transmisión, aún cuando la norma fue originalmente concebida para funcionar por debajo de 1.5 Mb/s para aplicaciones de CD-ROM. Sin embargo, ésta cualidad y flexibilidad no son consideradas ideales para aplicaciones a tasas de bits más altas. Para encontrar el objetivo de alta calidad y altas tasas de bits, fue formulada la norma MPEG2.

Los algoritmos de codificación que conforman el MPEG1 y MPEG2 son similares, esto es, esquemas de codificación entre tramas de compensación de movimiento usando DCT y ambos predicción hacia

adelante y hacia atrás. Como en el MPEG1 se usan tres tipos de codificación de tramas; tramas 'I', tramas 'P' y tramas 'B'. El procesamiento de estos tres tipos de tramas son similares a los de MPEG. No obstante, la norma MPEG2 provee más características que la MPEG1, tales características incluyen manipulación de vídeo entrelazado, escalabilidad, compatibilidad, y otras opciones para muy alta resolución de codificación de vídeo. Una de las principales diferencias entre los esquemas de codificación de las normas MPEG1 y MPEG2 es que la norma MPEG2 es capaz de manipular secuencias de vídeo entrelazado adaptable con cualquier trama o modo de campo, y la MPEG1 solo puede tratar con modo fijo. En un vídeo entrelazado, el codificador MPEG1 solo puede tratar con cada campo o cada trama (para unir o fusionar los dos campos adyacentes) como una unidad total o integral, en cambio el codificador MPEG2 puede, en adición, explotar la correlación entre los dos campos en una trama y seleccionar el modo óptimo. La norma MPEG2 proporciona más opciones para los esquemas de codificación, incluyendo escoger búsqueda en zig-zag ó sesgada, tabla de cuantificación lineal o no lineal, y la resolución de DCT dentro de los macrobloques. La escalabilidad para satisfacer la necesidad de diferente resolución de imágenes, proporciona servicios de grado múltiple, y múltiple ancho de banda. El estándar MPEG2 proporciona cinco tipos de herramienta de escalabilidad: Partición de datos, escalabilidad de ruido SNR, escalabilidad espacial y escalabilidad temporal, los detalles pueden ser encontrados en "*Generic coding of moving pictures and associated audio*". [96]

MPEG2 intenta ser una norma genérica para muchos requerimientos y muchos tipos de aplicaciones, ISO/MPEG ha decidido asociar a la

[96] ISO/IEC, "Generic coding of moving pictures and associated audio," Committee Draft 13818-2, Nov. 1993

norma MPEG2 con perfiles y niveles. Un perfil usa un subconjunto de elementos sintácticos disponibles para soportar un número de características técnicas y requerimientos de funcionalidad para un grupo de aplicaciones similares, los niveles son definidos por el rango de parámetros tales como tamaño del cuadro o imagen, velocidad de tramas, tasa de bits, tasa de pixel, tamaño de buffer, etc.

Es posible obtener una calidad de video NTSC con una tasa de 4 a 6 Mb/s usando la norma de MPEG2. [Chen,1995]

- **Perfiles y niveles**

Los perfiles y niveles proporcionan los medios para definir los subconjuntos de la sintaxis y semántica, y por lo tanto la capacidad del decodificador para decodificar un cierto flujo. Un perfil es un subconjunto definido de la sintaxis de flujo de bits que es definida por MPEG2. Un nivel es un conjunto definido de parámetros restringidos impuestos en el flujo de bits.

Se definen cinco perfiles:

1. Perfil principal: sin escalabilidad y máxima calidad
2. Perfil simple: igual que el perfil principal pero sin tramas interpoladas y en orden para ser guardadas por una RAM.
3. Perfil escalable SNR: Una mejora en escalabilidad de relación señal-ruido
4. Perfil escalable espacial: donde también es soportada la escalabilidad en la resolución de cuadro espacial.
5. Perfil alto: soporta 4: 2: 2 y total escalabilidad.

Junto con cuatro niveles:

1. Nivel bajo, similar a CIF de la H.261 o SIF¹ de MPEG1.
2. Nivel principal, corresponde a la televisión convencional.
3. Nivel alto 1440, corresponde aproximadamente a HDTV con 1440 muestras por línea.
4. Nivel alto, corresponde aproximadamente a la HDTV con 1920 muestras por línea.

Nivel alto		X			X
Nivel alto-1440		X		X	X
Nivel principal	X	X	X		X
Nivel Bajo		X	X		
	Perfil Simple	Perfil principal	Perfil escalable SNR	Perfil escalable Espacial	Perfil alto

Tabla 3.4.2.1 Combinación de perfiles y niveles

El objetivo del perfil principal es soportar las aplicaciones más comunes que probablemente serán introducidos primero. El perfil simple permite abatir memoria y complejidad del hardware a expensas de menos características técnicas y disminución de la calidad. Los últimos tres perfiles proporcionan un entorno para escalabilidad y alta resolución de crominancia.

¹ *SIF(Source Input Format) es un término general para describir que la entrada ha sido restringida a las limitaciones de la tasa de datos de MPEG y típicamente consiste de 360x240 pixeles de componente de luminancia y 180x120 pixeles componentes de crominancia Cr y 180x120 pixeles componente de crominancia Cb

Comparando con H.261, da como resultado que la calidad de MPEG1 y MPEG2 es mejor debido al uso de predicción bidireccional. Estos estándares, junto con H.261 ayudarán a estimular aplicaciones sofisticadas de multimedia y educación.

Debido a la compresión temporal hacia adelante y hacia atrás usada por MPEG, una mejor compresión y mejor calidad puede ser producida. Como MPEG no limita la resolución del cuadro, datos con resolución alta pueden aun ser comprimidos usando MPEG: El proceso de fuera de orden, el cual ocurre en ambos lados de codificación y decodificación puede introducir retrasos. Esto indeseable en el videoteléfono y la videoconferencia. (Chiariglione,1995)

3.4.3. RECOMENDACIÓN MPEG4

APLICACIONES MULTIMEDIA.

La MPEG4¹⁰ tiene como objetivo aplicaciones que incluyen comunicación audiovisual en tiempo real, multimedia y percepción remota. Esta también tiene identificadas categorías de requerimientos incluyendo material de entrada, calidad, formato de vídeo, formato de audio, tasa de bit, retardo, complejidad, seguridad, operación interactiva, interoperabilidad en red, anotación de otros datos y extensibilidad. Un algoritmo de compresión simple no permite cubrir al amplio rango de objetivos, especificaciones y requerimientos.

¹⁰Nota: Es importante mencionar que hasta la fecha (Mayo de 1998), no se ha logrado la estandarización de la MPEG4, por lo cual lo aquí expuesto sufrirá rápidamente un gran número de modificaciones.

El foco común del grupo MPEG4 es el desarrollo de *Modelos de Verificación (VM)*. Un modelo de verificación VM es una plataforma común con una definición precisa de algoritmos de codificación y decodificación que pueden presentarse como herramientas dirigidas a funciones específicas. Hasta octubre de 1994, el grupo de vídeo MPEG4 tiene enfocado su esfuerzo sobre un simple modelo de verificación VM que ha tenido gradual evolución desde la versión 1.0 hasta la versión 4.0, durante éste proceso se ha incrementado el número de funciones deseadas, satisfaciendo el objetivo principal de escalabilidad temporal, escalabilidad espacial y eficiencia de compresión.

3.4.3.1. FUNCIONALIDADES.

Como el estándar MPEG4 intenta cubrir un amplio rango de aplicaciones multimedia, debe soportar ciertas funciones tales como seguridad, bajo retardo, sincronización, interconexión, etc.

Algunas de estas aplicaciones las tiene ya, o existen aplicadas a un número de estándares comunes. El estándar MPEG4 se usará comúnmente o mejorará soluciones a fin de aplicarse a soluciones bajo este alcance.

El comité MPEG4 está buscando soluciones para soportar 8 funciones principales (aún en discusión), estas sin embargo no están soportadas por los estándares existentes. Dichas funciones se dividen dentro de 3 clases mayores, basadas en los requerimientos soportados por ellas.

- **Contenido básico interactivo:**

Esta clase incluye cuatro funciones centrales sobre requerimientos para aplicaciones mezcladas de alguna manera interactiva entre el usuario y los datos, es decir, el contenido básico de datos

multimedia y acceso a sus herramientas, manipulación del contenido básico y edición del flujo de bits, codificación híbrida, natural y sintética, y mejora temporal de acceso aleatorio. Las aplicaciones beneficiadas por estas funciones incluyen reconstrucción de datos desde librerías en línea, compras interactivas desde la casa, producción y edición de películas.

- **Compresión:**

Esta clase se compone de dos funciones; mejorar la eficiencia de codificación y decodificación de múltiples flujos de datos comunes. Esencialmente dirigido a aplicaciones que requieren un eficiente almacenamiento y transmisión de información audiovisual, así como su eficiente sincronización. Las aplicaciones beneficiadas con estas funciones son: búsqueda de información sobre *Internet* y *realidad virtual*.

- **Acceso universal:**

Las dos funciones restantes son: robustez en ambientes propensos a errores y escalabilidad del contenido básico. Estas funciones permiten a MPEG4 decodificar datos accesibles sobre un amplio rango de aplicaciones multimedia y con varias cualidades en términos de resolución temporal y espacial para objetos específicos, que podría ser decodificado por varios decodificadores con diferentes complejidades. Las aplicaciones beneficiadas por estas funciones son comunicaciones inalámbricas, búsqueda en bases de datos y acceso a diferentes niveles de contenido, resoluciones y cualidades.

EL MODELO DE VERIFICACIÓN (VM)¹

Una importante fase en el proceso de estandarización de la MPEG4 es el desarrollo del modelo de verificación VM. Un modelo de verificación VM es una codificación completamente definida, así como el algoritmo de decodificación. Está compuesto de herramientas y algoritmos que sirven como referencia para evaluar el desempeño de otras herramientas y algoritmos.

ESTRUCTURA GENERAL DEL MODELO DE VERIFICACIÓN (VM).

MPEG4 considera que una escena está compuesta de *objetos de vídeo* (VO) con propiedades intrínsecas tales como forma, movimientos y textura. Creyendo que tal representación del contenido básico es clave para la interactividad con objetos de una gran variedad, y representar objetos sintéticos de 2D y 3D generados por computadora, que evoluciona en el tiempo. Como un tercer ejemplo, uno podría considerar un VO como una estructura representando un objeto extraído desde una escena automáticamente.

El proceso para representar un VO no necesita estandarización. Generalmente el siguiente método es usado en MPEG4 para definir los VOs. Para todas las pruebas de secuencia usadas en MPEG4, el VO es definido considerando una secuencia completa de vídeo como un simple VO, o proveyendo estructuras de objetos segmentados considerando cada una como un VO diferente.

La figura 3.4.3.1. muestra un diagrama de bloques del codificador MPEG4.

¹ En adelante tomaremos VM para referirnos al Modelo de Verificación de MPEG4.

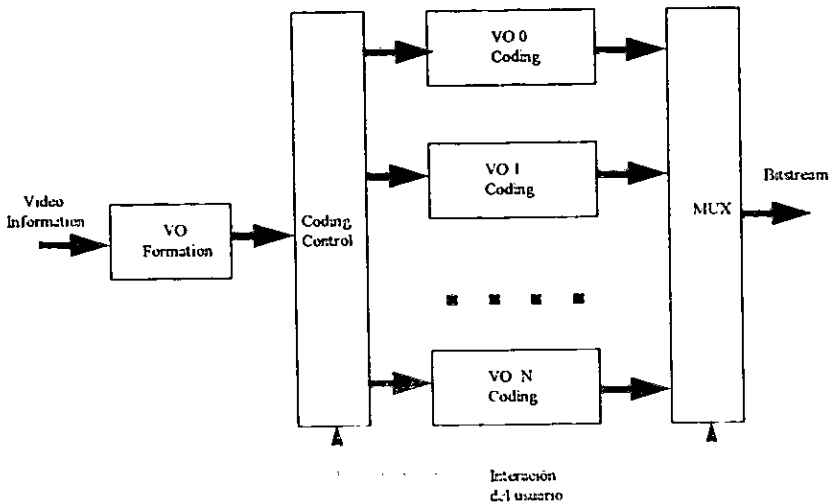


Figura 3.4.3.1 Estructura general de un codificador MPEG-4 en el video VVi.

La característica más importante de éste codificador es la representación intrínseca basada en VO cuando definimos una escena visual. En realidad, un usuario, o un algoritmo inteligente puede seleccionar a los diferentes VOs codificados, componiendo una fuente de datos con diferentes parámetros y diferentes métodos de codificación.

Otra característica importante del VVi es que la velocidad del frame no es explícitamente definida por este acceso, esto significa que el codificador y el decodificador pueden funcionar a diferentes velocidades de frame que no necesiten ser constantes a través de la secuencia de vídeo.

La interactividad entre el usuario y el codificador o el decodificador puede conducir a diferentes caminos. El usuario puede decidir el nivel

de interactividad en el codificador, distribuir el control en ambos, la tasa de bit disponible entre los diferentes VOs, o influenciar el multiplexaje al cambiar parámetros tales como la composición de script en el codificador.

La figura 3.4.3.2 muestra un diagrama a bloques del decodificador en el VM.

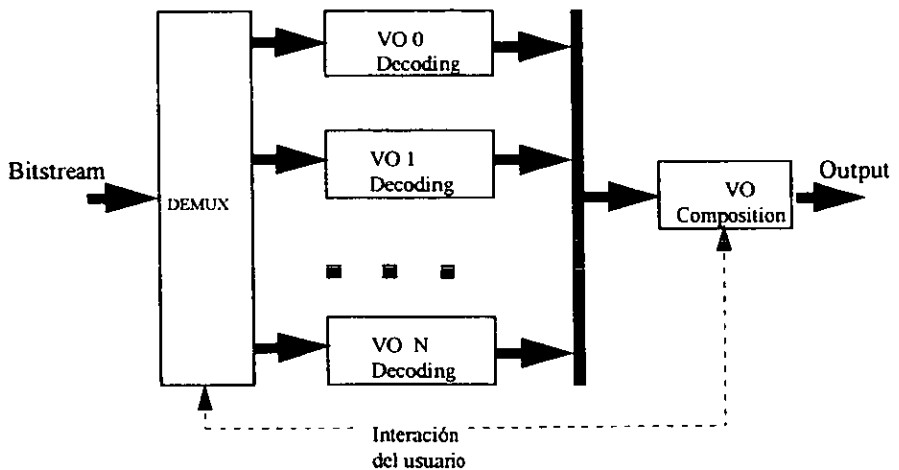


Figura 3.4.3.2 Estructura general de un decodificador MPEG4 en el video VM.

La estructura de decodificador es similar al codificador pero inversa, excepto por el bloque de composición al final. El método de composición (mezclado) de los diferentes VOs depende de la aplicación y el método de multiplexaje usado en el sistema. EL VM 4.0 usa una solución recursiva en que se mezclan secuencialmente dos planos al mismo tiempo.

ESTRUCTURA DE DATOS DE UN MODELO DE VERIFICACIÓN (VM).

El VM común usa las siguientes clases de sintaxis al permitir una fácil y eficiente representación de funcionalidades.

- **Sesión de Vídeo (VS)**
Una secuencia completa de vídeo puede formarse de varios VS.
- **Objetos de Vídeo (VO):**
Es una clase que define objetos específicos en una escena.
- **Capa de Objetos de Vídeo (VOL)**
Es una clase que realiza la resolución espacial y temporal de un VO.
- **Planos de Objetos de Vídeo (VOP)**
Es una ocurrencia de un VO a un tiempo dado, dos diferentes VOP's pueden pertenecer al mismo objeto en dos diferentes tiempos y no necesariamente a diferentes objetos.

En suma, una Sesión de Vídeo (VS) es una colección de uno o más Objetos de Vídeo (VO), cada uno puede consistir de una o más Capas de Objetos de Vídeo (VOL), y cada una consiste de una secuencia de fotografías instantáneas en forma de Planos de Objetos de Vídeo (VOP's). Un VO puede ser de forma arbitraria (rectangular en un caso especial). Para una simple codificación únicamente un VOL existe por VO.

Comúnmente, debido a falta de datos de 3D/4D en las pruebas de datos de MPEG4, además de la falta de tecnología adecuada para su representación, manipulación y codificación, el Modelo de Verificación (VM) de la MPEG4 es basado en objetos de vídeo planos y por lo tanto es limitada a la manipulación de objetos en 2D.

3.4.3.3. HERRAMIENTAS OCUPADAS COMÚNMENTE EN EL MODELO DE VERIFICACIÓN.

FORMA DE CODIFICACIÓN

Las dos formas de codificar información son consideradas como características inherentes de un objeto de video. Y éstas son conocidas como:

Forma de codificación de información binaria

Es un medio de etiquetar información que define, qué porción de pixeles del objeto corresponden al objeto de video en un tiempo dado. La forma de información binaria es comúnmente representada como una matriz del mismo tamaño, como el de la caja ligada a un VOP. Todos los elementos de la matriz pueden tomar uno de dos posibles valores, dependiendo si el pixel se encuentra dentro o fuera del objeto de video.

Forma de codificación de escala de gris

Tiene una estructura similar a la forma binaria con la diferencia que todos los pixeles (elementos de la matriz) pueden tomar un rango de valores (usualmente 0 a 255), representando el grado de transparencia del pixel. Los valores que toma la escala de gris son similares a la noción del plano alfa usado en gráficos de computadora; donde un valor de 0 corresponde a un pixel completamente transparente y el valor de 255 corresponde a un pixel completamente opaco.

ESTIMACIÓN Y COMPENSACIÓN DE MOVIMIENTO¹.

La estimación y compensación de movimiento tiene que ser introducida en el Modelo de Verificación en orden para explotar la redundancia temporal existente en la secuencia de video. El principio de la

¹ Ver punto 2.2.2 de este texto .

reducción de redundancia temporal en el VM de la MPEG4 es similar al ocupado por otros estándares MPEG1 y MPEG2.

Codificación de Textura:

La técnica de codificación de textura usada en un VM es similar en varios aspectos los usados en otros estándares MPEG1 y MPEG2. Debido a ésta simplicidad, al buen funcionamiento y también al orden incorporado en un modo completo en las funcionalidades del VM, así como a la robustez de errores, la codificación de textura del VM es seleccionada como una técnica de bloques básica, donde se toma un especial cuidado en el orden de extender los bloques básicos aproximándolos al manejo arbitrariamente del VOP.

Codificación Escalable.

Satisfacer la capacidad de escalabilidad espacial y temporal son dos funcionalidades requeridas por un número de aplicaciones. Un flujo de bits es escalable si al menos un subconjunto del flujo de bits es suficiente para generar una presentación segura.

Robustez de Errores.

Es una característica muy importante del VM, en ambientes adversos así como en redes inalámbricas. La robustez se logra insertando resincronización de marcas de campo en el flujo de bit, con aproximadamente 3 espacios constantes. La resincronización de las marcas de campo debe ser insertada por el codificador antes del primer macrobloque, y también si el número de bits de salida desde la última marca de resincronización exceden un valor predeterminado.

El presente escrito, es una breve introducción a MPEG4. Es importante notar que en poco tiempo nuevas versiones de MPEG4 aparecerán y algunas de las herramientas descritas aquí serán diferentes.

Resultados concretos muestran que MPEG4 provee nuevas funcionalidades así como contenido básico interactivo. Únicamente el futuro dirá el impacto que el estándar MPEG4 tendrá en aplicaciones emergentes de multimedia.

Glosario NORMA H.263

CBPY	Patrón de bloque codificado para luminancia (codec block pattern for luminance)
COD	Indicación de macrobloque codificado (codec macroblok indication)
CPM	Multipunto de presencia continua (continuos presence multipoint)
CPPB	Patrón de bloque codificado para bloque B (codec block pattern for B blocks)
DBQUANT	Información de cuantificación para imágenes B (quantization information for B-pictures)
DQUANT	Código de dos bits que define un cambio de QUANT (two bit code to define change in QUANT)
EOS	Fin de secuencia (end of sequence)
GBSC	Código de comienzo de grupo de bloques (group of bloks start code)
GFD	Indicador de trama de GOB (GOB trame ID)
GLCI	Indicador de canal lógico de GOB (GOB logical channel indicator)
GN	Número de grupo (group number)
GOB	Grupo de bloques
GQUANT	Información de cuantificación para GOB (GOB quantizer information)
INTER	Modo de codificación cuando se aplica la predicción
INTRA	Modo de codificación cuando no se aplica la predicción
INTRADC	Coefficiente cc de bloque INTRA (DC coefficient for INTRA blocks)
MB LAYER	Capa de macrobloque (macroblock layer)
MCBCP	Tipo de macrobloque y patrón de grupo codificado (macroblock type and coded block pattern)
MODB	Modo macrobloques para bloques B (macroblock mode for B-blocks)
MVDB	Datos de vector de movimiento para macro bloque B (motion vector data)
OBMC	Compensación de movimiento por bloque superpuestos (overlapped block motion compensation)
PEI	Información de inserción suplementaria de imagen (picture extra insertion information)
PLCI	Indicador de canal lógico de imagen (picture logical channel indicator)

PQUANT	Información de cuantificador de imagen (picture quantifier information)
PSC	Código de comienzo de imagen (picture start code)
PTYPE	Información de tipo de imagen (picture type information)
STUF	Información de relleno (stuffing information)
TCOEF	Coefficiente de la transformada (transform coefficient)
TR	Referencia temporal (Temporal Reference)
TRB	Referencia temporal para tramas B (temporal reference for B-frames)

Bibliografía.

1. Arch C. Luther
Principles of Digital Audio and Video
Artech house, Inc., 1997
2. <http://www.cis.ohio-state.edu/hypertext/faq/usenet/jpeg-faq/part1/faq.html>
3. <http://www.execpc.com/~cwhite/squeeze.htm>
4. U.I.T/C.C.I.T.T. H.263. Libro Azul. Tomo III, Fascículo 3.6
Transmisión en línea de señales no telefónicas. Transmisión de señales radiofónicas y de televisión.
Recomendaciones de las series H y J. IX Asamblea Plenaria. Melbourne, 14-25 de Noviembre 1988, Ginebra 189, p. 120-124.
5. UIT-T, Sector de Normalización de las Telecomunicaciones de la UIT.
Proyecto H.263.
Transmisión en línea de señales no telefónicas. Codificación de vídeo para la comunicación a baja velocidad binaria.
UIT-T 5 de Julio de 1995, p. 1-60

6. D.J. Le Gall
MPEG: a video compression standard for multimedia applications
Comm. ACM, vol. 34, no 4 pp.46-57. Apr. 1991
7. Furht, Borko
A survey of multimedia techniques and standards part 2: Video
Compression, Real -Time Imaging
1995) 1:pag. 319 - 337
8. Motorola
multimedia device data : Q1/95, DL158/D, REV 0.:
1995PAG 4-43 and 4-44.
9. Proceedings of the IEEE
"The Development of an Integrated Audiovisual Coding Standard: MPEG"
L. Chiariglione
February 1995, vol.83, No. 2, pp 151-158
10. Proceedings of the IEEE
Video in Phone Lines: Technology and Applications
D.T. Chen, and T.R. Hsing, February 1995, vol.83, No. 2, pp 175-194
11. Swiss Federal Institute of Technology
MPEG4 Video Verification Model: A video encoding/decoding algorithm
based
on content representation; Touradj
http://drogo.cselt.stet.it/ufv/leonardo/icjfiles/mpeg-4_si/paper5.htm

PARTE II

**LOS SISTEMAS DE
VÍDEO DIGITAL**

CAPÍTULO 1

TELEVISIÓN DE ALTA DEFINICIÓN

La televisión ha obtenido enormes avances, desde su colorido y la calidad de la imagen, hasta las transmisiones vía satélite y los sistemas por cable, sólo por mencionar algunos aspectos. Pero hoy en día se está en vísperas de un nuevo avance: la alta definición.

Resulta claro para cualquier observador que la calidad de la imagen de la televisión actual, aún en las mejores condiciones, es bastante mala. Si se compara la definición de una fotografía tomada con un cámara de 35 mm, con la de la misma escena tomada con una cámara de televisión veremos que el detalle de la fotografía es muy superior; es decir, contiene más elementos de imagen (píxeles) que la primera. Existe entonces una evidente necesidad de desarrollar tecnologías de televisión de mayor calidad.[De la herran,1989]

La televisión de hoy depende de tres diferentes sistemas usados en todo el mundo (NTSC, PAL y SECAM). Consecuentemente para intercambio internacional de vídeo, los sistemas de conversión son requeridos.

Por ésta razón, dado que igual que antes el intercambio internacional de video sigue siendo tan importante, ha habido un gran interés entre las naciones para que todos los esfuerzos que se hagan busquen una unificación global en los sistemas de televisión de próximas generaciones, cuando se dio a conocer el comienzo de investigaciones en alta definición.

Los primeros pasos en la investigación de un sistema de televisión de alta definición comenzaron en 1970, y fue la NHK (Nippon Hoso Kyokai), Corporación de la Televisión Japonesa, la que inició los primeros estudios y experiencias, y desde 1980 los japoneses comenzaron a impulsar la televisión de alta definición. El objetivo al principio planteado por los japoneses consistía en crear una norma de producción de calidad equivalente a un film de 35 mm. Para ello proponían doblar el número de líneas (1,050) y mantener la frecuencia de campo en 60 Hz. Simultáneamente, desarrollaron equipo de producción en esta norma que incluía cámaras, magnetoscopio, telecine, mezclador, etc. En Junio de 1989, se inicio en Japón, una hora diaria de transmisión experimental de alta definición, usando la transmisión directa del satélite japonés Bs-2, y extendiéndose desde 1991 a ocho horas de transmisión diaria usando un canal del satélite Bs-3b entonces lanzado y asignado exclusivamente para el sistema de HDTV.

Fue hasta 1980 cuando los europeos empezaron a interesarse por la televisión de alta definición, porque los sistemas PAL y SECAM eran superiores en calidad al sistema NTSC y además porque en esa década Japón y USA presentaron en el CCIR (Comité Consultivo Internacional de Radiocomunicaciones) una norma mundial para la producción en televisión de alta definición pero esta era más

compatible con el NTSC 525 líneas/ 59.94 campos que con los sistemas europeos.

En 1986 en Yugoslavia, hubo una asamblea plenaria donde se realizó una revisión, pero no hubo acuerdo para formar una recomendación para hacer el sistema de 1,125/60 un estándar global unificado, por la oposición levantada por los países europeos, quienes propusieron un sistema de 1250 líneas y 50 campos como un estudio estándar para alta definición.

En 1989 y 1990, se logró un acuerdo general en 23 artículos incluyendo relación de aspectos, colorimetría y forma de la señal de sincronización, sin embargo no se llegó a un acuerdo sobre el número de líneas de exploración ni frecuencia de campo haciéndolo imposible para integrar el estándar internacional por este punto de vista. Se decidió entonces que este estudio debería continuarse. [De la Fuente, 1994

1.1. Definición y Estándar para HDTV.

La televisión de Alta Definición es un sistema televisivo en el cual el televidente, no alcanza a distinguir las líneas o rayas horizontales que tan molestas nos parecen en la televisión común y corriente, éste efecto sólo se consigue con una pantalla que tenga más de 1,000 líneas de barrido electrónico, se ha demostrado que con más de 1,000 líneas activas, la resolución se compara con la de las películas de 35 mm, considerada como referencia de alta calidad.

Este receptor de Alta Definición posee por lo menos el doble de la definición de una televisión actual y su contenido informativo equivale en ambos casos, a unos 900 a 1000 Mbits/s. Así como es diferente el

receptor, también lo son los transmisores, además la televisión de Alta Definición emplea sonido estereofónico de alta calidad comparable a la del disco compacto y excelente capacidad en la reproducción del color. [De la Fuente, 1994]

En cuanto a la estructura de la exploración, la progresiva (1:1) requiere más ancho de banda que la entrelazada (2:1), por lo que es apropiado el sistema de exploración entrelazado, por la economía del ancho de banda, la posibilidad de una gran calidad de la imagen y la facilidad de la conversión de la televisión de alta definición al sistema existente.

También se contempla el aspecto ergonómico de la misma y se ha llegado a la conclusión de que la visión horizontal en el ser humano es más amplia que la vertical, y pruebas subjetivas han demostrado que los telespectadores prefieren una relación de aspecto mayor, cuando crece el tamaño de la pantalla, lo que ha inducido a los constructores de este nuevo sistema de televisión a fabricar pantallas mucho más largas en lo horizontal, un formato de pantalla ancha mejora el realismo de la escena, porque ésta "inunda" al ojo. El sistema NHK fue inicialmente diseñado para una relación de (5:3), ésta relación fue estudiada posteriormente por SMPTE de los Estados Unidos y se decidió adoptar una relación de (16:9), atendiendo a la importancia de tener una base común con el formato del cine, ésta relación, ahora ha sido adoptada también por muchos otros sistemas propuestos, incluyendo el de NHK, ésta relación de aspecto de 16:9 de alta definición cubre casi toda la pantalla de imagen amplia, excepto del cinemascope. [Benson and Whitaker, 1992]

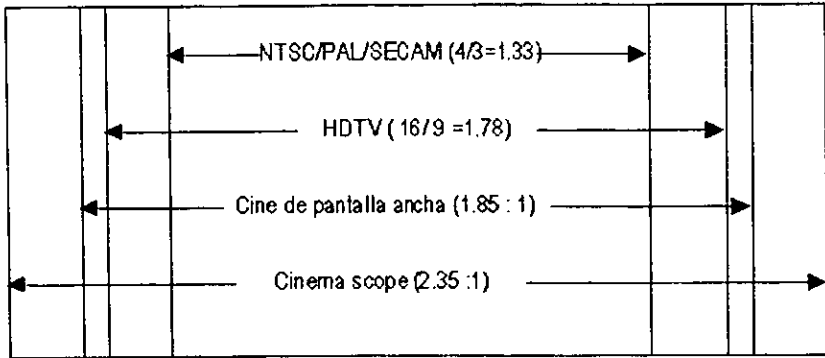


Figura 1.1.1 Comparación de la relación de aspecto entre el cine y la televisión

El ancho de banda nominal de video para el estándar de estudio de alta definición quedó determinado en 30 Mhz.

En el inicio del desarrollo del cine se encontró que el movimiento podría ser presentado suavemente en alguna frecuencia de cuadros más rápida que 15 por segundo, la experiencia en la producción para salas de cine, mostró que movimientos muy rápidos, podrían no ser fielmente mostrados en 16 cuadros por segundo, así se cambio a 24 por segundo, éste permanece como estándar en el mundo. La frecuencia adoptada por el servicio de televisión europea fue de 25 Hz, en acuerdo con la planta de 50 Hz, de esa región, y éste estándar ha persistido en los sistemas PAL y SECAM. La frecuencia de cuadro para el sistema NTSC fue inicialmente seleccionada en 30 por segundo, después cuando fue introducido el servicio de color NTSC, la frecuencia de cuadro de 30 HZ del servicio monocromático NTSC, fue cambiada a 29.97 Hz. El hecho de que la frecuencia de campo y cuadro de PAL y SECAM (25 y 50 Hz) y de NTSC (29.97 y 59.94 Hz) sean diferentes ha afectado el diseño del servicio de alta definición. [De la Fuente, 1994]

1.2. Algunos tipos de sistemas híbridos.

El sistema japonés es analógico aún cuando internamente sus sistemas son digitales, la parte de entrada y salida de la señal es analógica pero todo el proceso de compresión de la señal se realiza con un algoritmo de computadora lo que lo hace digital. Esta compresión de señal es básica para poder manejar anchos de banda de transmisión no demasiado grandes, dada la enorme cantidad de información visual detallada que la alta definición requiere transmitir. Japón es el único país donde la televisión de Alta Definición es toda una realidad. En este país el sistema Muse ya tiene varios años trabajando. Una señal de Alta definición requiere aproximadamente 30 Mhz de ancho de banda de transmisión. Con la compresión de la señal se ha conseguido que el ancho de banda disminuya a 8.1 Mhz. Japón ha querido imponer su sistema al mundo, en especial en el mercado comercial de los Estados Unidos. Pero ha habido un gran rechazo de estos últimos y de Europa, lo que los ha llevado a su vez a desarrollar sus propios sistemas. En especial se argumenta que el sistema analógico japonés contradice la tendencia mundial tecnológica en la que todo se vuelve digital. Europa tiene su propio sistema, HD-Mac. Por el momento los europeos están esperando a ver lo que pasa en Estados Unidos. Éstos, por su parte, han visto cómo diversas compañías, universidades y grupos de trabajo han intentado imponer su propio sistema, luego de muchas batallas, se formó al final un grupo de trabajo que a cada sistema le encontró ventajas y desventajas. De hecho se demostró que todos tienen problemas, unos en la transmisión, otros en la concepción de modulación, etc. Eso hace muy difícil la creación de una norma de fabricación. Entonces se concibió una unión de todos esos sistemas a la cual le llamaron "La Gran Alianza" y la constituyen principalmente NBC, Philips, el centro David Sarnoff (antes laboratorios RCA), la Thomson, el Massachusetts

Institute of Technology (MIT), AT&T, General Instrument y Zenith. De esto debe salir una norma, pero seguramente el futuro de la HDTV será digital.[De la Fuente, 1994]

Los Estados Unidos Proyectan ahora pasar directamente a la tecnología digital en televisión , salto que los pondrá en vanguardia de las soluciones adoptadas por Europa y Japón para la HDTV, resulta, pues muy probable que sea digital el sistema elegido para la nueva generación de TV en los Estados Unidos. Ello responde a la motivación esencial de mejorarse la calidad de las imágenes recibidas en un grado considerable frente al actual sistema NTSC, pero al mismo tiempo, los nuevos canales de TV digitales han de acomodarse en la misma anchura de banda de 6 Mhz que los existentes, sin interferirlos en lo absoluto. Este último requisito que parecía disuasorio, demostró ser viable gracias a las técnicas de compresión de señal digital, y fue General Instrument en San Diego, quien primero proclamó la posibilidad de utilizar canales de 6 Mhz para transmitir imágenes digitalizadas, a razón de 1,050 líneas por cuadro (alta definición). Se pudo apreciar que la transmisión digital ofrece una recepción igualmente nítida a todos los espectadores situados en su zona de cobertura, mientras que la recepción de señales analógicas sufre una degradación paulatina al acercarse al contorno de dicha zona, esta y otras ventajas de calidad se unen a la posibilidad de nuevos enfoques operativos como por ejemplo que los propios espectadores manipulen las imágenes utilizando programas interactivos facilitados por la entidad emisora. Las transmisiones de televisión, radio, la telefonía móvil, e incluso la convencional se aprestan a asumir su papel en la era digital que se avecina. [Corcoran, 1992]

1.3. Sistema MUSE

Para establecer un estudio estándar de alta definición, la NHK realizó varias pruebas psico-físicas por más de veinte años. Basados en los resultados de las pruebas, los parámetros de la señal tales como el número de líneas (1,125), relación de aspectos (16: 9) y frecuencia de campo (60 Hz) fueron seleccionados. Los valores de éstos parámetros fueron adoptados por BTA (technology Association) de Japón como estándar BTA-001, y por SMPTE (Society of Motion Picture and Television Engineers) en los Estados Unidos como estándar SMPTE 240. La opción de 1,125 líneas fue justificada por el estándar existente esto es, el punto medio entre 525 y 625 líneas es 575, el doble de este valor es 1,150 líneas, este número par de líneas no podría producir entrelazado de línea alternado, el número non más próximo, teniendo un factor común con 525 y 625 seleccionado por la NHK es de 1,125.

NHK ha desarrollado una versión para reducir el ancho de banda a solo 8 Mhz. La reducción del ancho de banda utiliza una técnica llamada multiple sub-Nyquist sampling encoding (MUSE).

MUSE: Sistema de codificación con muestreo múltiple ("Multiple Sampling Encoding"). Norma elaborada en Japón durante los años setenta para radiodifusión por satélite de TV de alta definición. Tiene las siguientes características:

Líneas exploradas por trama	1,125
Número de líneas activas por trama	1,035
Relación de aspecto	16:9
Estructura de exploración	2:1
Frecuencia de campo	60 Hz
Frecuencia de línea	33,750 Hz
Ancho de banda de luminancia ('Y')	20 Mhz a 22 Mhz

Ancho de banda de crominancia (señal diferencia de color) 7 Mhz

El sistema MUSE es en esencia un sistema de transmisión de televisión analógica que usa técnicas de muestreo de elementos individuales de una imagen entrelazada, seleccionando los fragmentos de estos para formar la señal a transmitir. El ancho de banda total de la señal HDTV es muestreado con una compresión de cuatro usando un patrón de muestreo dimensional. En el receptor un buffer de tramas es usado para almacenar las muestras que llegan en cuatro campos sucesivos para la reconstrucción de la imagen de alta definición.

Para imágenes fijas, dos tramas sucesivas, cuatro campos, son combinadas y después los elementos de imagen faltantes son obtenidos por interpolación en el receptor.

En el sistema MUSE las partes de movimiento de la imagen son submuestreadas a una fracción del muestreo principal ya que la resolución del sistema visual humano es menor para objetos en movimiento que para objetos fijos.

La figura 1.3.1 muestra gráficamente la versión presente del sistema MUSE. Comenzando con una señal HDTV del tipo Hi-Vision, la cual submuestrea cada tercer elemento de imagen, quedando una señal resultante que puede ser transmitida a través de un canal satelital que puede ser manejado en una banda base de 8 Mhz. Al mismo tiempo, la información del movimiento de la imagen es enviada del transmisor al receptor en el período del borrado vertical de la imagen. Esta información del vector de movimiento es usada para hacer posible la reconstrucción sofisticada de la imagen HDTV en el receptor.

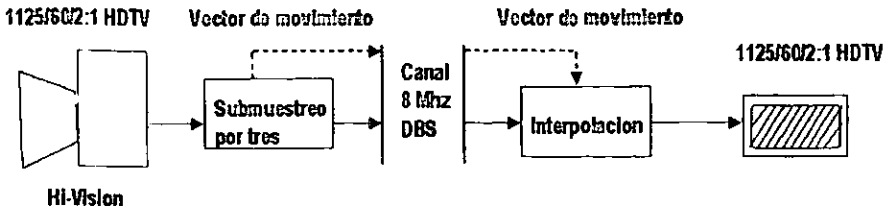


Figura 1.3.1 La esencia del sistema

Además la señal de audio digital y la señal del vector de movimiento son transportadas en el intervalo del borrado vertical.

El hardware o aparatos para la codificación y decodificación son muy complejos y los circuitos convertidores para recepción emplean complejos circuitos integrados. [Slater, 1991]

1.4. Sistema HD-MAC

MAC: Componentes Analógicas Multiplexadas ("Multiplexed Analogue Components"),

Es una familia de normas desarrolladas en la Comunidad Europea durante los años ochenta para introducir la radiodifusión de TV por satélite y la mejora de la calidad. Se basa en multiplexar en el tiempo las componentes analógicas de la imagen (luminancia y diferencia de color) con señales digitalizadas de sonido y de datos, la resolución es de 180,000 píxeles, un 50% más que PAL/SECAM que es de aproximadamente 120,000 píxeles. Forman parte de ella la D2-MAC (625 líneas) y la HD-MAC (1250 líneas), ambas a 50 campos o tramas por segundo. [Corcoran, 1992]

Las naciones de la Comunidad Económica Europea han establecido conjuntamente una organización conocida como EUREKA, y han desarrollado un sistema europeo capaz de competir con el NHK de Japón. El proyecto es uno de la familia MAC del sistema de transmisión directa por satélite (DBS), este es conocido como HD-MAC (alta definición MAC). Uno de los objetivos de EUREKA, EU-95 es establecer normas para el sistema europeo de HDTV compatible (la compatibilidad significa que la señal HD-MAC pueda enviarse por los canales de transmisión MAC, y que el receptor pueda captarla en su aparato MAC sin una degradación apreciable, además de que la transmisión de imagen fuese analógica) con los sistemas de transmisión directa por satélite (DBS) convencionales D-MAC y D2-MAC en uso. Las organizaciones principales involucradas en tal proyecto son Philips (Holanda), Thomson (France), y Siemens (Alemania del oeste).

El punto de partida común con el sistema adoptado por NHK en Japón, consistió en admitir unas características esenciales para todo sistema HDTV, es decir:

- a) Duplicar la resolución (tanto vertical como horizontal) de la imagen, en Europa, se pasaría de 625 a 1250 líneas por cuadro.
- b) Mejorar la resolución temporal (mayor nitidez de imágenes en movimiento).
- c) Separar la señal de vídeo en varias componentes analógicas, una de brillo o luminancia, y dos de diferencia de color (crominancia), que luego se multiplexan (preferiblemente por división en el tiempo) para ser transmitidas por vía radioeléctrica terrestre o a través de satélite.
- d) Producir las imágenes en un formato extendido, más agradable al usuario, se ha adoptado la relación de aspecto 16:9.

e) Transmisión de señales digitales de sonido en alta fidelidad, así como canales de datos con información relativas a los proceso de transmisión y recepción.

Exploración.

El sistema HD-MAC emplea el doble de líneas exploradas del sistema MAC de 625 líneas con una relación de aspecto de 16:9, entrelazado, 1250/50/2:1/16:9. Las bandas base de la luminancia y crominancia son respectivamente, 21 Mhz y 10.5 Mhz, para una relación de compresión de 3:2 y 3:1 respectivamente. Una frecuencia de muestreo común es de 20.5 Mhz y esta impone un ancho de banda máximo de 10.125 Mhz. Esta es la señal de frecuencia modulada hacia el satélite, usando el ancho de banda de 27 Mhz disponible en los satélites de Europa. La información adicional (sonido, datos, DATV) es transportada por multiplexaje en los intervalos de borrado vertical y horizontal. [Benson and Fink, 1991]

HD-MAC es un sistema híbrido analógico-digital. La parte analógica es la señal visual, la cual está basada en conceptos MAC; la parte digital contiene canales de sonido, teletexto, datos de acceso condicional, y datos de televisión asistida digitalmente DATV (digitally assisted television). Los datos de 20.25 Mbit/seg de DATV, son transportados en el intervalo del borrado vertical. La señal visual analógica tiene un ancho de banda de 10.125 Mhz, comparada con 8.4 Mhz del MAC convencional. Medidas especiales como pre-énfasis no lineal han sido adoptadas para mejorar la relación señal a ruido S/N. [Benson and Whitaker, 1992]

Señal Asistida Digitalmente.

Un elemento esencial del sistema HD-MAC es la señal DATV televisión asistida digitalmente (digitally assisted television) transmitida durante el intervalo del borrado vertical de $1/50 \text{ seg.} = 20,000 \text{ ms.}$ Este tiempo

permite una máxima tasa de transmisión de 1 Mbit/s. Una función principal de la señal DATV es el custodiar o mantener el switcheo de las ramificaciones (branch-switching), mostradas más adelante, en preciso sincronismo en el transmisor y receptor.

Reducción de ancho de banda.

Para transmitir los 21 Mhz de banda base de luminancia compatible con los receptores D-MAC y D2-MAC, se requiere una reducción de 4 veces de ancho de banda. Esto es cumplido con los procesos de codificación y decodificación mostrados en la figura 1.4.1.

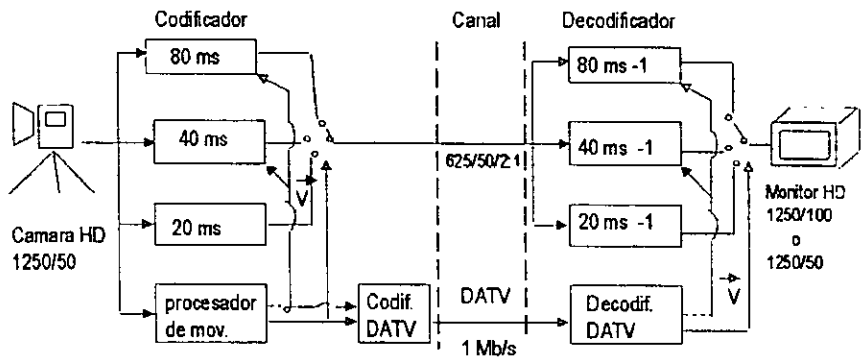


Figura 1.4.1 Elementos y trayectoria de transmisión del sistema HD-MAC EUREKA

El codificador tiene bifurcaciones (*branches*) para tres grados de movimiento:

- Una bifurcación de 80 ms (4 campos) para áreas de movimiento lento o estáticas de la escena.
- Una bifurcación de 40 ms (2 campos) para áreas de movimiento.

- Una bifurcación de 20 ms (1 campo) para movimientos rápidos y cambios de escena bruscos.

Estas ramificaciones son conmutadas para el canal de transmisión por un circuito procesador de movimiento. Las señales conmutadas son transmitidas después para el receptor por el canal DATV, donde la bifurcación en uso en un tiempo particular, después de la decodificación, es conectada al receptor para el procesamiento y despliegue en el régimen 1250/50/2:1/16:9 de la cámara o equipo de telecine en el transmisor. La señal de crominancia, cerca de 10.5 Mhz de banda base, es transmitida después igualmente con codificación de tres ramificaciones pero sin compensación de movimiento.

La codificación en la ramificación de los 80 ms se extiende a través de 4 campos, de aquí que, el ancho de banda de luminancia para áreas estacionarias es reducido de 21 Mhz a 5.25 Mhz. Pero las ramificaciones de 40 ms y 20 ms se extienden solamente a través de dos campos y un campo respectivamente, una reducción adicional del ancho de banda es requerido, esta reducción es realizada por algunos procesos, incluyendo:

- Exploración "quincunx" o campos alternados. Exploración quincunx es la exploración de sucesivos elementos de imagen alternados de dos líneas adyacentes. Este proceso produce un entrelazado "sintético".
- Línea "shuffling" con muestras intercaladas de alta definición de manera que dos líneas dentro de un campo son transmitidas como una línea de paquete/MAC. Para el cual los receptores D-MAC y D2-MAC responden compatiblemente a este proceso. Los datos requeridos para realizar la operación inversa en el receptor son transmitidos a través del canal DATV. [Bensond and Fink, 1991]

La Gran Alianza.

El 31 de mayo de 1990 General Instrument Corp. sometió la primera propuesta para un sistema totalmente digital. Para Diciembre de 1990 se anuncio la entrada digital de Zenith seguida rápidamente por AT&T, y después por el Massachusetts Institute of Thechnology (MIT). Así había 4 serios competidores para la propuesta de televisión digital, así como una modificación del sistema MUSE y también propuestas para la Televisión de Definición Extendida EDTV (Extended Definition Television). Durante los siguientes años, éstos sistemas fueron estudiados.

En febrero 1993, la FCC toma la importante decisión para una tecnología totalmente digital, pero no podía decidir entre los 4 competidores. Por consiguiente después de tal confusión, se recomendó formar una "***Gran Alianza***". Esta tomaría las mejores características de los cuatro sistemas y desarrollaría una norma HDTV. La mayor parte del resto de 1993 se dedico a establecer las características de esta nueva norma.

Durante 1994 se construyó el sistema y en 1995 se instaló para probarlo y desde 1996, se prueba comercialmente en algunos países (USA, México).

El sistema HDTV de la Gran Alianza

En la Gran Alianza (GA), se llegó a un acuerdo general de que la nueva norma debería ser totalmente digital. Los objetivos para éste sistema fueron:

- Alta calidad digital de sonido e imagen en la HDTV.
- Un sistema que podría coexistir con la transmisión analógica de TV, sin interferir con ella.
- El costo efectivo de equipo debería ser accesible para los consumidores, productores y todo usuario al momento de introducción y duración de la norma.
- Interoperabilidad con otros medios de transmisión y aplicaciones.
- El potencial para una norma mundial.

El sistema de la GA define cuatro capas o niveles. Las capas simplemente separan la norma en cuatro bloques. Esto no solamente hace más fácil entender la norma, sino que proporciona medios para interoperabilidad con otras normas que tienen capas equivalentes. La figura 1.4.2 muestra las cuatro capas de la norma de la GA.

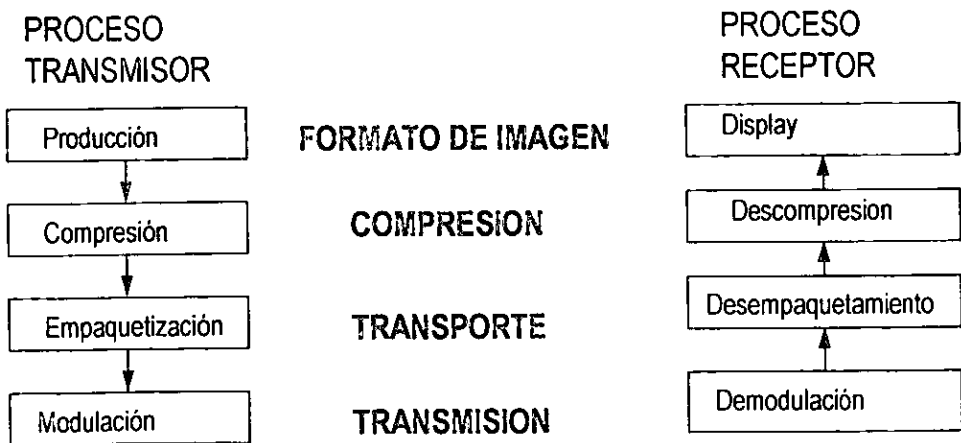


Figura 1.4.2.- Capas de la norma de transmisión de la GA HDTV.

1 Capa de imagen. El sistema proporciona múltiples formatos y velocidad de tramas, las cuales pueden ser decodificadas y presentadas por un receptor HDTV con la norma de GA. Esto permite que servicios diferentes tengan normas de exploración diferentes. Teóricamente, varios formatos pueden ser transmitidos por el sistema, pero ya en la práctica, en el receptor se limita el número de formatos que se pueden aceptar. Por lo tanto la GA limita el formato de exploración escogiendo las variaciones de solamente dos números de líneas activas: 720 y 1,080 en ambos se puede escoger la velocidad de tramas, en el caso del formato de 1,080 líneas la exploración progresiva o entrelazada puede ser usada. En el formato de 720 líneas solamente se puede usar la exploración progresiva. La tabla 1.4.1 muestra los formatos a escoger.

Líneas activas	Píxeles por líneas activas	Tasa (Hz)	Entrelazado	Relación de aspecto
720	1280	59.94/60	No	16/9
	1280	23.98/24	NO	16/9
	1280	29.97/30	NO	16/9
1080	1920	59.94/60	Si	16/9
	1929	23.98/24	No	16/9
	1920	29.97/30	No	16/9

Tabla 1.4.1 Formatos de exploración de la norma de transmisión de la GA

2 Capa de compresión. La compresión de video es posible porque la redundancia es inherente en muchas imágenes en movimiento, la compresión de video del sistema de la GA está basado en la norma

ISO-MPEG2. El sistema de audio usa compresión de audio digital Dolby AC-3, este incluiría un total sonido surround. La velocidad de datos de vídeo para formatos de imagen HDTV es aproximadamente 18.9 Mb/s y para formatos de imagen de definición estándar es de 3 a 5 Mb/s.

3 Capa transporte. El sistema de la GA usa transporte de empaquetamiento basado en la estructura de paquete de MPEG2. La capa de transporte recibe flujos de bits separados de audio y vídeo y los multiplexa por medio del empaquetamiento. Se pueden manejar cualquier número de flujo; múltiple flujo de vídeo, audio, u otro tipo de datos que pueden ser enviados por el mismo canal, limitado solamente por la capacidad de tasa de datos del sistema. El empaquetamiento de la GA es un bloque de longitud de 188 bytes, cada paquete contiene un encabezamiento de 4 bytes de datos de control.

4 Capa de transmisión. Se usa modulación de banda lateral o doble banda vestigial (VSB) de símbolos multinivel. Esta es llamada modulación 8-VSB o 16-VSB, donde éste número, indica el número de niveles de amplitud del símbolo transmitido. El sistema de la GA 8-VSB actualmente transmite 2 bits de datos por cada simbolo. El sistema de radiodifusión terrestre usa el formato 8-VSB que tiene mejor capacidad para protección de error, mientras que el sistema de TV por cable usa el formato 16-VSB que ofrece velocidad de datos más alta, pero también requiere la alta relación señal a ruido (SNR) que el cable puede proporcionar. Finalmente una pequeña portadora piloto es agregada (más que la total supresión de la portadora como es usual en VSB). Esta portadora piloto se utiliza para minimizar la interferencia con el servicio existente NTSC.

El sistema de la Gran Alianza es diseñado pensando en el futuro, la computadora y aplicaciones de multimedia. *El uso de MPEG2 permitirá que la HDTV interactúe directamente con aplicaciones de multimedia de la computadora. Por ejemplo, podría ser grabada HDTV en una computadora de multimedia, y aplicaciones de CD-ROM podrían ser presentados en un sistema HDTV.* [Inglis, 1996].

Bibliografía.

1. José de la Herrán
Información Científica y tecnológica Vol.11 No. 156
Sept. 1989
2. De la Fuente Alberto
Información Científica y tecnológica Vol.16 No. 22
Mayo 1994
3. Blair Benson K., Whitaker Jerry
Featuring Hdtv System
Television Engineering Hand Book, Pp 24.1-24.25
4. Corcoran, Elizabeth
Investigación y Ciencia
Abril 1992 No. 187
Mc Graw Hill 1992
5. K. Blair Benson, Donald G. Fink
Hdtv Advanced Television For The 1990s
Intertext Publications
Mc Graw Hill Publishing Company, Inc 1991, Pp 9.36-9.41

6. Slater, Jim
Modern Television Systems To Hdtv And Beyond
Pitman Publishing 1991. Pp 146-160

7. Andrew F. Inglis, Arch C. Luther
Video Engineering
Second Edition
Mc Graw Hill 1996. Pp 157-178

CAPÍTULO 2

2.2 VIDEOCONFERENCIA

Una combinación de bajos precios, líneas más rápidas, mejores productos y una mayor competencia ha hecho a la *videoconferencia* más atractiva que nunca. Existen cuatro tipos principales de videoconferencia: sistemas de sala, unidades de escritorio (desktop PC -Personal Computer-), PC vídeo basado en LAN y PC videoconferencia basada en Internet.

Una vez que el sistema es instalado apropiadamente, la videoconferencia puede entregar beneficios de aplicaciones como telemedicina, aprendizaje a distancia, y colaboración por compartición de archivos de PC, sin mencionar los ahorros obtenidos con la reducción de viajes y gastos de hospedaje.

La implementación del vídeo no siempre es fácil. Antes debe decidirse la compra de uno o más sistemas, el tipo de sistema que se desea, que protocolos y estándares se necesitan, el costo del ancho de banda requerido, y cuántas localidades necesitan estar simultáneamente puenteadas (vídeo multipunto). Además de las facilidades que podrían desearse opcionalmente, como videocámaras (para videograbar la sesión), un pizarrón en blanco (whiteboard) para aplicaciones de colaboración, y

una cámara de documentos para cambiar a ver documentos en papel. [NIH Telecommunications Branch, Sep. 1997]

La videoconferencia soporta dos vías de audio y vídeo comunicación. Esto significa que dos o más personas en diferentes localidades se pueden ver y escuchar una a la otra al mismo tiempo. No es lo mismo que estar allí, pero la videoconferencia es frecuentemente más conveniente y menos costosa que un viaje, especialmente cuando se trata de un grupo de personas.

Equipamiento Básico.

Un sistema de videoconferencia debe tener equipo audio-visual (monitor, cámara, micrófono y bocinas) así como un medio de transmisión de información entre los sitios. Como puede imaginarse, una conexión en banda ancha vía satélite con equipamiento de calidad de estudio produciría una excelente conexión de vídeo en completo movimiento, pero el costo del equipamiento y transmisión es grande. Los recientes avances tecnológicos en computación y Telecomunicaciones han despertado gran interés en sistemas de compresión de vídeo, los cuáles hoy en día transmiten información vía Internet o la red telefónica, reduciendo con mucho los costos de la videoconferencia.

El hardware que permite realizar la videoconferencia es llamado códec (contracción de codificador-decodificador). El códec toma la señal analógica de vídeo y la codifica (digitaliza y comprime). El códec también la decodifica (descomprime y des-digitaliza) a la recepción de la transmisión, y en tal proceso la consecuente desventaja es un "salto" en la imagen y un retardo en el audio de 0.5 a 2 segundos. Aunque el vídeo comprimido no tiene calidad de radiodifusión, es el más adecuado para muchas aplicaciones de videoconferencia. [Pacific Bell, Agosto 1996]

Los sistemas de videoconferencia desktop se comunican de diferentes

formas, las cuales son descritas a continuación.

Conferencia POTS.

POTS (Plain Old Telephone Service) es el servicio telefónico básico que proporciona acceso a la Public Switched Telephone Network (PSTN) -red telefónica pública conmutada-. Este servicio es ampliamente disponible, pero tiene un muy bajo ancho de banda (las velocidades típicas de los módems son 9.6 kbps, 14.4 kbps o 28.8 kbps). Hay muy pocos productos de videoconferencia desktop que tratan de operar a esas velocidades. Funcionan actualmente sobre el H.324, un estándar de interoperatividad para productos de videoconferencia operando a una tasa de 28.8 kbps.

Conferencia Conmutada 56.

Conmutada 56 es un servicio digital que entrega 56 kbps de ancho de banda. Es muy similar a ISDN excepto que sus canales son de un menor ancho de banda. Conmutada 56 es una tecnología anterior, y está siendo reemplazada por la ISDN. Típicamente los productos de videoconferencia desktop que operan sobre la conmutada 56 pueden también operar sobre la ISDN.

Conferencia ISDN.

ISDN (Integrated Services Digital Network) es un servicio digital. Existen dos tasas de acceso definidas para ISDN, la Basic Rate Interface (BRI) -interfaz de tasa básica- y la Primary Rate Interface (PRI) -interfaz de tasa primaria-, la BRI proporciona dos canales de datos a 64 kbps (canales B) y un canal de señalización de 16 kbps (canal D). Hay muchos productos de videoconferencia desktop en el mercado que utilizan ISDN BRI. Sin embargo, existen problemas con el acceso a ISDN porque todavía no se encuentra disponible en todos los lugares. Aún cuando sea disponible, puede todavía resultar no ser muy trivial el configurar al sistema correctamente. La interfaz de acceso primario PRI proporciona 23 ó 30

canales B de 64 kbps, y un canal de D a 64 kbps. La ISDN PRI es costosa y por consiguiente no realmente aplicable para videoconferencia en escritorio.

Porque los canales ISDN ofrecen 64 kbps de ancho de banda, los estándares y algoritmos de conversión han sido diseñados alrededor de éste número, 64 kbps se ha convertido en algo así como un número mágico para videoconferencia. [Bell Atlantic Corporation, Mayo 1996]

Los sistemas avanzados de videoconferencia pueden estar configurados de varias diferentes formas. Las dos principales configuraciones son los sistemas roll-about y los sistemas basados en salas.

Sistemas ROLL-ABOUT

Los *sistemas roll-about* son sistemas integrados móviles de videoconferencia. Éstos sistemas son compactos y usualmente integrados en un gabinete rodante. Generalmente, éstos sistemas proporcionan el servicio de videoconferencia para grupos de tamaños pequeños y medianos. Los componentes estándar de un sistema roll-about incluyen un sistema de control, un códec, al menos una cámara, dos monitores, un sistema de audio, micrófonos y amplificadores. Durante una videoconferencia, los participantes utilizan el sistema de control para manipular los ángulos de las cámaras, los niveles de sonido y otras aplicaciones como el pizarrón blanco o la transferencia de archivos. Los sistemas son frecuentemente configurados con una segunda cámara que es usada para mostrar mapas, cartas y documentos. Debido a lo compacto y su movilidad, los sistemas roll-about son la versión más popular de configuración de videoconferencia actualmente en uso.

Sistemas de Videoconferencia en Sala.

Los *sistemas en sala* incluyen el mismo equipo que los sistemas roll-about; sin embargo éstos están permanentemente instalados dentro de una sala de juntas importante como puede ser una sala de consejo o una de videoconferencia. Los sistemas basados en sala pueden ser configurados para atender grupos de personas de mayor tamaño que los sistemas roll-about. Aún cuando los sistemas para sala poseen los mismos componentes estándar de los sistemas roll-about, pueden también integrar componentes adicionales como cámaras extras. Generalmente éste tipo de sistemas son instalados de manera que puedan estar ocultos a la vista cuando no están en uso. Las señales de una videoconferencia son generalmente transportadas sobre las líneas telefónicas, utilizando comúnmente velocidades de 128 a 384 kbps, o de 1 a 3 líneas ISDN BRI. [Advanced Video Conferencing]

Típicamente utilizan componentes de alta calidad audiovisual, sofisticados códecs, y destacando costosos aparatos de interfaz con objeto de crear un ambiente apropiado para una sala llena de participantes. (Pacific Bell, Agosto 1996]

Ofrecen más flexibilidad, y una mejor calidad en audio y video que los sistemas desktop (usualmente a causa del incremento del ancho de banda en las líneas). Además, soportan fuentes de video mezcladas para usos como presentaciones, documentos y video cintas. También proporcionan entradas múltiples de audio. Los precios inician en aproximadamente \$8,000 sin embargo un sistema típico de monitor dual cuesta aproximadamente \$ 50,000. [NIH Telecommunications Branch, Sep. 1997]

Sistemas de Videoconferencia DESKTOP (Escritorio).

Utiliza una computadora personal y un hardware y software especial para codificar y decodificar la señal. Este tipo de sistemas utiliza componentes más económicos y apropiados para uso individual o pequeños grupos. El

sistema desktop frecuentemente incluye la facilidad de compartición de documentos, la cual permite a los participantes ver y editar documentos en la computadora mientras que ellos se puedan ver y escuchar unos a otros. La compartición de documentos y el relativamente bajo costo de los sistemas desktop hacen de ésta una herramienta ideal de comunicación, colaboración, y aprendizaje. [Pacific Bell, Agosto 1996]

Estos sistemas cuestan entre \$ 1,200 y \$ 2,000, pero tienen una pobre calidad de video. El sistema desktop puede no convenir para aplicaciones como telemedicina, donde el color, el tamaño del monitor, y las tasas de trama son importantes, pero funciona como ya se mencionó anteriormente para aplicaciones de enseñanza remota y compartición de aplicaciones de PC. Ahora bien, por la demanda que el video tiene del disco duro de la PC, es recomendable tener un disco duro de la mayor capacidad posible. [NIH Telecommunications Branch, Sep. 1997]

Videokonferencia en LAN (Local Area Network).

Las LAN están comúnmente dentro de un campus y en compañías para interconectar computadoras. En la capa física, las LAN usualmente consisten de segmentos Ethernet de 10 Mbps, o en Token Ring de 4 ó 16 Mbps. La diferencia entre Ethernet y Token Ring está en como los usuarios ganan el acceso al medio de transmisión. En Ethernet es Carrier Sense Multiple Access With Collision Detección (CSMA/CD) -acceso múltiple por detección de portadora con detección de colisión-, red en la que los usuarios transmiten los datos y escuchan para detectar las colisiones. Si ocurren colisiones, los usuarios esperan durante cierto rango de tiempo antes de intentar transmitir nuevamente. Token Ring es una red donde un token ó estafeta pasa alrededor de los usuarios y dichos usuarios deben ganar el acceso al token antes de comenzar a transmitir. Los productos de videokonferencia desktop soportan una variedad de protocolos de red como TCP/IP, Novell IPX/SPX, NetBIOS, y Appletalk. [Bell Atlantic Corporation, 1996]

Alentados por los bajos precios del equipo y el crecimiento del ancho de banda en la red, los proveedores de sistemas de videoconferencia ofrecen ahora sistemas que llevan videoconferencias full-dúplex a través de redes LAN (Local Area Network) y WANs (Wide Area Networks). Comúnmente, los proveedores no satisfacen el estándar de la ITU (International Telecommunication Union) H.323, que permitirá la interoperabilidad entre unos productos y otros, y permite a los múltiples usuarios participar en una videoconferencia (multipunto). Ésta junto con el pequeño tamaño de la imagen de vídeo, son las principales deficiencias de una videoconferencia basada en LAN.

La mayor parte de las soluciones de videoconferencia desktop (escritorio) fueron originalmente diseñadas para ISDN. Con los recursos suficientes la mayor parte de las compañías han puesto en desarrollo extensas infraestructuras LAN y WAN, lo cual ha hecho que los proveedores de servicios de videoconferencia también proporcionen vídeo servicios desktop sobre la LAN. Un sistema de videoconferencia basado en LAN es básicamente plug-and-play (conectar y utilizar). No son requeridos componentes de servidor, y los usuarios pueden inicializar los sistemas ellos mismos. La imagen en una videoconferencia basada en LAN es quarter CIF (Common Intermediate Format), o QCIF, alrededor de un octavo de una pantalla completa; una imagen de videoconferencia desktop ISDN es usualmente full CIF (FCIF), proporcionando una mejor imagen, y una pantalla de imagen mucho más grande.

El administrador debe observar cuidadosamente el ancho de banda en la red. Un solo circuito de videoconferencia full-dúplex con sonido y vídeo continuos entre el transmisor y el receptor pueden utilizar hasta un décimo del ancho de banda de una red LAN Ethernet de 10 Mbps. Y procesando toda la información de vídeo puede colocar una carga pesada a los CPU's (Central Processing Unit) de las PC's en ambos extremos. Tal carga puede impedir a la computadora completar otras tareas. En muchos casos,

el ancho de banda usado puede ser reducido con solamente un deterioro menor en la calidad de la imagen. [NIH Telecommunications Branch, Junio 1997]

Conferencia en INTERNET.

Las LANs proporcionan conectividad entre una comunidad local. El *Internet* conecta LANs a otras LANs. El protocolo desarrollado para interconectar varias redes es llamado Internet Protocol (IP) -protocolo de internet-. Dos protocolos de transporte fueron desarrollados con IP, a saber TCP y UDP. TCP (Transmission Control Protocol) -protocolo de control de transmisiones- proporciona un servicio confiable extremo a extremo porque usa un recuperador de errores y reordenador. UDP (User Datagram Protocol) -protocolo de diagrama de usuario- es un servicio no muy confiable, sin recuperación de errores.

Las aplicaciones de videoconferencia que operan sobre la Internet principalmente utilizan UDP para transmisiones de audio y vídeo. TCP no es práctico por su mecanismo de recuperación de errores. Si fueron reenviados paquetes perdidos, éstos llegarán muy tarde estando fuera ya de cualquier uso. TCP es usado para algunas aplicaciones de videoconferencia y para otros datos que no son muy sensibles al tiempo como los datos en pizarrón blanco y datos de aplicaciones compartidas. [Bell Atlantic Corporation, 1996]

Conferencia BACKBONE MULTICAST (MBone).

La *conferencia Multicast Backbone ó Mbone* -conferencia múltiple en red primaria- ha sido llamada red virtual porque es dividida en partes de la Internet. Utilizando la Mbone, es posible transmitir audio, vídeo y otros datos en tiempo real a múltiples destinos a través de la red global Internet. Para entender como funciona la Mbone, definimos primero unicast y multicast. Unicast es una transmisión de datos punto a punto. Para lograr una o muchas transmisiones deben enviarse copias separadas de datos por el originador ó fuente a cada destino. Multicast posibilita una manera

más efectiva de entregar la misma información de datos a múltiples destinos.

Multicast ha sido implementado en redes locales como Ethernet y Fiber Distributed Data Interface (FDDI) -interfaz de datos distribuidos por fibra- durante algún tiempo. Hasta que en 1989, Steve Deering especificó una extensión al Protocolo de Internet para soportar multicasting conocido también como RFC (Request For Comment) 1112. Permitiendo ésta extensión al concepto de multicast expandirse para incluir así a la Internet entera. Básicamente la Mbone está compuesta de "islas" soportando IP multicast, las cuales son conectadas por "túneles" de enlaces punto a punto.

Con IP multicast, los datos son transmitidos a un "host group" (grupo huésped). Grupos que son especificados por una clase D IP address en el rango de 224.0.0.0 a 239.255.255.255. Los hosts que desean recibir los datos se enlazan al grupo de host. Es entonces responsabilidad de los routers (ruteadores) multicast el entregar eficientemente los datos enviados por los originadores a todos los receptores que están enlazados al grupo. El Internet Group Management Protocol (IGMP) -grupo de administración de protocolo de Internet-, es utilizado por los routers multicast para determinar qué grupos están activos en una subred en particular. Hay varios protocolos de ruteo que los routers multicast pueden utilizar para construir árboles de entrega eficientes de los originadores a los receptores, incluyendo DVMRP (Distance Vector Multicast Routing Protocol) -vector de distancia de protocolo de ruteo múltiple-, MOSPF (Multicast Open Shortest Path First) - ruta primaria abierta múltiple más corta- y PIM (Protocol Independent Multicast) -protocolo múltiple independiente-. El problema es que esos protocolos de ruteo no han sido disponibles tradicionalmente en las producciones de routers. El trabajo al respecto de éste problema es "tuneleo". Un router multicast encapsula el paquete múltiple dentro de un paquete regular IP y fija el destino a otro router multicast. Al intervenir

routers y subredes, los datos aparecen como un paquete unicast normal.

Los retos de transmitir audio y video sobre la Internet están enfocados al desarrollo de un nuevo protocolo de transporte propuesto por el Audio/Video Transport working group of the IETF (IETF-AVT working group) -grupo de trabajo de Transporte de Audio y Video de la IETF-. El RTP (Real-Time Transport Protocol) - protocolo de transporte en tiempo real- proporciona el soporte para secuenciamiento, temporización, y calidad de servicio reportado para enlaces punto a punto o multipunto. Las herramientas Mbone comúnmente utilizadas implementan alguna versión del RTP haciendo algunas de ellas disponibles comercialmente como son Communique! (InSoft), InPerson (Silicon Graphics), y ShowMe (Sun Microsystems). [Bell Atlantic Corporation, 1996]

Por otra parte; cuando el backbone está altamente congestionado, las ráfagas de video y audio podrían salirse de sincronía. Una forma de controlar el ancho de banda en la LAN es usar un esquema de interconexión de redes llamado IP multicast (dirección Internet múltiple). Éste es especialmente útil donde varios usuarios estarán observando la misma ráfaga. En lugar de enviar los mismos paquetes cinco veces a cinco diferentes usuarios, todos pueden compartir un solo paquete en secuencia. Para que IP multicast funcione, cada router en la LAN debe estar actualizado para soportar Ese protocolo. Siendo otra forma de administrar el ancho de banda por medio del administrador de la red, con objeto de administrar cuidadosamente el ancho de banda y solamente admitir ciertas ráfagas de video a un tiempo. [NIH Telecommunications Branch, Sep. 1997]

ATMs y Videoconferencia

Asynchronous Transfer Mode (ATM)- modo de transferencia asincrono-es la última tecnología entrante en videoconferencia para escritorio. Terminando con el hecho de que hasta últimamente las líneas ISDN a cada PC fueron necesarias para realizar la videoconferencia en escritorio,

pero actualmente con la tecnología ATM, esto no se requiere. Ahora se puede realizar por medio de un gateway (o compuerta) ATM. Esta compuerta incluye un software especial, una tarjeta de códec, y un software para que funcione.

Uno de tales productos producido por First Virtual solamente requiere una computadora 486/66 con 8 Mb en disco duro y 5 Mb de espacio libre, operando sobre alguna versión de MS-Windows. Este sistema tiene capacidad para video CIF (Common Intermediate Format), el cual es considerado de alta calidad CIF (352 x 288 pixeles) con resolución de 30 tramas por segundo. Algunos otros programas admiten solamente QCIF (176 x 144 pixeles), el cual es todavía bastante bueno. QCIF se parece en mucho a un video en VHS. Aun cuando ATM es más caro, es todavía más efectivo en costo por el ancho de banda extra que es posible alcanzar con el ATM comparándolo sobre las líneas separadas ISDN. Además, las empresas no tienen que comprar nuevas PC's para aplicar tal software, ahorrando dinero con esto. La ISDN puede dar 128 kbps mientras que la nueva tecnología ATM permite 384 Kbps. Este tipo de transferencia es posible todavía con el tráfico regular sobre la red LAN. El cual puede ser manejado mediante el uso de Multimedia Conference Servers (MCS) - servidores de conferencia multimedia-. Los cuales son usados para conectar muchas conferencias a un mismo tiempo por la misma conexión. Permitiendo esto a diferentes sistemas de videoconferencia trabajar en conjunto. Cada MCS debe soportar completamente el estándar H.320 para ser un verdadero MCS, de otra manera podría no lograrse la comunicación entre los diferentes sistemas. Algunas veces un MCS puede soportar algo más que el estándar, lo cual le permitiría comunicarse también con otros sistemas. [ATM's and Video Conferencing]

Ahora bien, como un medio de comunicación de video interactivo, en el ámbito educativo por ejemplo, la videoconferencia ofrece las siguientes ventajas:

- Establece una comunicación visual entre los participantes.

El maestro puede ver y escuchar a estudiantes remotos, a la vez que puede usar la conversación y el lenguaje corporal para enfatizar la comunicación. La frecuente interacción incrementa el entendimiento y fomenta la instrucción personalizada. Las estrategias de enseñanza interactiva como discusión y cuestionamiento pueden también ayudar a acoplar y motivar a los estudiantes haciéndolos participantes.

- Posibilita la conexión con recursos externos.

Expertos remotos, pueden auxiliar validando lo entendido, proporcionando retroalimentación, e introduciendo ejemplos prácticos. Esta conexión real del mundo puede proporcionar una gran motivación, especialmente si los estudiantes participan y el experto interactúa con ellos en un nivel apropiado.

- Soporta el uso de diversos medios.

Las gráficas y fotos en color se ven bien y pueden ayudar a transmitir un concepto difícil o simplificar instrucciones. Los sistemas basados en salas incluyen una cámara de documentos que permite una transmisión de imágenes de alta calidad. Esta característica puede ser usada para mostrar también objetos, así como fotos y gráficas, y algunas veces se incluyen "pizarrones" -simples pantallas de texto con unas pocas anotaciones-.

El compartir documentos facilita la colaboración y retroalimentación.

Algunas aplicaciones permiten compartir aplicaciones, permitiendo a los usuarios en cada sitio ver y editar un documento. Esta facilidad de compartir fomenta la colaboración y retroalimentación en tiempo real. [Pacific Bell, Agosto 1996]

Trabajando con Limitaciones Tecnológicas.

En general, el vídeo comprimido debe transmitir información por un canal o "conducto" más pequeño que el de la radiodifusión televisada. La cámara y el micrófono llevan más información de la que el canal puede manejar, de modo que la información de vídeo y audio debe ser procesada por un componente del equipo llamada códec (coder-decoder) antes de que ésta pueda ser transmitida. Las señales de entrada son decodificadas por el códec antes de ser enviadas al monitor y las bocinas. De modo que el desarrollo de todo éste proceso toma su precio en la calidad de imagen y sonido resultantes, observándose usualmente las siguientes características:

- "Fantasmas" en el video o "imágenes temblorosas".

El códec es la ruta de compensación para flujos rápidos de información. Una forma en que el códec compacta la información es reduciendo la tasa de la trama (número de imágenes de vídeo por segundo), y en la que el efecto resultante al hacer movimientos rápidos es que aparece como una "sacudida" de la imagen. El códec también puede bajar la resolución para comprimir la información, lo cual puede resultar en una imagen "borrosa" o "fraccionada". De manera que para reducir esos efectos en la imagen, reduce la cantidad de flujo de información visual. Esto es; evita los movimientos rápidos, emplea un fondo sencillo, y coloca un fondo pastel detrás de los participantes para reducir la información visual extraña.

- Retardos de audio.

Podrían llegar a presentarse, pues toma alrededor de un segundo comprimir la información, enviarla, y descomprimirla en el otro extremo. Los usuarios principiantes en videoconferencia usualmente experimentan algunos inconvenientes al seguir las conversaciones debido a este tiempo de retardo. Dado que no hay manera de evitar el retardo, como usuario hay que aprender a terminar las ideas en una sola exposición con

una conclusión obvia. Y a su vez, los oyentes deberán evitar interrumpir y emplear indicaciones visuales (como inclinar la cabeza) en lugar de usar afirmaciones verbales.

- "Recorte" de audio o eco.

Puede presentarse si el sistema de audio no está apropiadamente configurado. Aunque si se tiene experiencia en problemas de audio, se puede resetear el cancelador de eco (si se tiene uno) y reducir el ruido de fondo. También se puede checar la documentación del equipo y checar las guías de volumen y posicionamiento del micrófono. Por otra parte, dependiendo del equipo, se pueden utilizar diademas y bocinas externas que pueden también ayudar a mejorar la calidad del audio. [Pacific Bell, Mayo 1996]

Un sistema de videoconferencia debe satisfacer algunos estándares. El estándar H.323 de ITU (International Telecommunications Union) permite interoperar entre sí a los diferentes equipos de videoconferencia. Entre otras cosas, H.323 especifica la codificación y decodificación de las señales de video y audio. Mientras que T.120 es un estándar que soporta la colaboración de datos, con beneficios que incluyen la interoperabilidad entre múltiples plataformas, transporte de red, y velocidad independiente. T.120 soporta conferencia en pizarrón blanco, pantalla compartida, anotaciones, e intercambio de imágenes. De manera que resulta importante el seleccionar un producto que soporte ambos estándares H.323 y T.120. [NIH Telecommunications Branch, Sep. 1997]

A continuación, se presenta un **resumen de los estándares de videoconferencia de la ITU**: [Network Computing]

H.320 Una serie de estándares que definen la *interoperabilidad de los sistemas sobre la ISDN*. Estos estándares definen las reglas para el establecimiento de las comunicaciones, trama, sincronización, y canales

de multiplexaje inverso ISDN. H.320 también incluye los siguientes estándares de audio y video codificación:

- H.261** Un estándar de videocompresión para anchos de banda múltiplos de 64 Kbps.
- G.711** Audio compresión básica de 48 Kbps a 64 Kbps. -técnica de modulación por códigos de pulsos como la usada en la telefonía regular.
- G.722** Audio de alta calidad en el mismo ancho de banda utilizando audio procesadores más sofisticados.
- G.728** Audio compresión a más baja tasa de bits. -16 Kbps.-

H.323 Una serie de estándares que definen la *interoperabilidad de la videoconferencia sobre redes de paquetes conmutados* con calidad de servicio -Quality of Service- (QoS) no garantizado. Estos estándares definen el empaquetamiento y sincronización del medio sobre las redes de paquetes, no sobre la QoS y el puerto (gateway) para la interoperabilidad de los sistemas H.320 y H.323. Los siguientes estándares de audio y video codificación son parte de la H.323:

- H.261** El mismo estándar de video compresión también usado en H.320 y H.323.
- H.263** Una alternativa de estándar de compresión para comunicaciones a baja tasa de bits.
- G.711, G.712, G.728** Estándares de audio compresión usados también en H.320.
- G.723** Audio compresión para 5.3 Kbps y 6.5 Kbps.

H.324 Una serie de estándares que definen la *interoperabilidad de la videoconferencia sobre las POTS* que establece el multiplexaje a bajo ancho de banda/protocolos de control. H.324 incluye los siguientes estándares:

H.263 Aunque H.261 está incluida en el estándar, H.263 es comunmente usada dado que puede comprimir video a muy bajas tasas de transmisión de bits.

G.723 H.324 incorpora solamente audio compresión a baja tasa de bit.

V.80 Estándar de interface de aplicación para sistemas desarrollando H.324 que convierte las ráfagas de datos sincronos a modem asincronos, posibilitando ajustes en la tasa durante una llamada, y notificando al usuario sobre las pérdidas de paquetes.

T.120 Una serie de estándares de *videoconferencia para colaboración de datos* que están incluidas en los tres estándares H.32x descritos anteriormente. T.120 incluye:

T.124 Estándar de control de conferencia.

T.126 Estándar para la compartición de imágenes.

T.127 Estándar para la transferencia de archivos binarios.

Por último; se presentan una serie de tablas (cortesía de *Network Computing, Junio 1997*); en la cual se presentan algunos de los principales productos de videoconferencia en desktop que existen actualmente en el mercado. Y en la cual se consideran algunas de las características más importantes a considerar en éstos equipos como son: descripción del producto, los estándares que soporta, precio, etc.

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
Hacer (Taiwan)	Acer Axxess 100	\$ 1,200 (PCI board, ISDN phone, camera, mike)	Desktop videoconferencing system over ISDN	Pentium PC 8 MB RAM 20 MB disk	Win 95	BRI ISDN	H.320 T.120	Bundled with Microsoft's NetMeeting
Aeontech International (Taiwan)	AEON AvShare, version 2.00	\$ 495 (Software, DSV-D modem, digital color camera)	Low-cost software-based audio/videoconferencing Kit over Internet, LAN/WAN or DSV-D modem	486/33 PC, 4MB RAM, 4MB disk, camera uses parallel port, no special board	Win 3.1, Win NT	LAN/WAN, DSV-D analog 28.8 Kbps modem	H.320 T.120	Whiteboard, image sharing
Alpha Systems Lab	Mega Conference	\$ 1,195	Video/audio tools over POTS	Intel PC	Win 3.x, 95 & NT	Analog POTS	Proprietary	Whiteboard, file transfer
Applied Communication Concepts	Interact	\$ 5,995 (Codec, camera, speaker phone)	Windows desktop videoconferencing system over, ISDN, high-quality person/document camera	i386SX, 4MB RAM	Win 3.x,	BRI ISDN	H.320	Whiteboard, shared, clipboards, file transfer, OLE
Boca Research	Boca VideoPhone	\$ 399	H.324-compliant videoconferencing user supplies NTSC/PAL mike, camera, speakers	Pentium PCI Bus master slot 12 MB RAM	Win 95	Video-ready V.34 modem	H.324	N/A
CINECOM Corp	CineVideo/ Direct Version 1.09	\$ 149 (QuickCam mike, software)	Desktop videoconferencing over 28.8 Kbps modems or TCP/IP, Works with QuickCam and any MSVIDEO compliant device	Intel PC	Win 3.x, 95 & NT	28.8 Kbps modem, TCP/IP	CINECOM Proprietary audio and video	Works with Intel ProShare
Compression Labs	CLI 1000	\$ 1,795	Video/audio/collaboration tools over ISDN, runs 15 fps in full CIF	Intel PC	Win 3.1 & 95	1-BRI ISDN, 128 Kbps operation	H.320	Shared whiteboard, file transfer (opt. Feat.)
	CLI 2000	\$ 2,295	Video/audio/collaboration tools over ISDN, runs 22 fps in full CIF	Intel PC	Win 3.1 & 95	3-BRI ISDN, 384 Kbps operation	H.320	Shared whiteboard, file transfer (opt. Feat.)
Core Tech Software	PictureView	Call for pricing	Audio/video conferencing over Ethernet, Token-Ring, ATM, ISDN, Switched 56	SUN	Sun-SPARC 1,1+	LAN/WAN, ISDN, ATM	H.320	Whiteboard, Share X
C-Phone Corp	C-Phone Videoconferencing System	\$ 1,995 (desktop equipment)	Windows desktop videoconferencing system over, LAN ISDN, Fractional T1 or Switched 56 digital lines	Intel PC	Win 3.1 & 95	Broadband modulation on LAN, Gateway to WAN	Proprietary on LAN, H.320 on WAN	High-speed file transfer, document, Camera (optional) & video playback
Connectix Corp	Connectix VideoPhone	\$ 99 (software) \$ 199 (with QuickCam)	Software-only audio/video and whiteboard conferencing bundled with Connectix QuickCam	Pentium PC Quadra or Power Mac	Win 3.x, 95 & NT, MacOS	TCP/IP EtherTalk	H.261 video, True-Speech audio	Whiteboard

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
Corel Corp	Corel VIDEO, Versión 1.0	\$ 499 per seat (excluding camera)	VideoPBX High res. full-motion video to desktops in enterprise campus setting	Macintosh Intel PC	Win 3.x, 95 & NT, MacOS OS/2	Unshielded twisted pair or coax in-house	Baseband modulation in-house, H 320 Via gateway to PSTN	Video broadcasts shared screen & document, file transfer, chat
Creative Software Technologies	Collaborator System 4000	Call for pricing	Audio/video/data conferencing PC, upgrade Kit	i486/33, ISA bus 8 MB RAM 5 MB disk	Win 3.1, WFW 3.11	ISDN, LAN/WAN	H 320	App Sharing, file, chat, transfer, shared whiteboard
CyberTron	VuFone, Version 1.1	\$ 3,995 (codec card, 1-BRI ISDN interface camera software)	Audio/video/data over ISDN, IsoEthernet, Switched 56	i486/33 or higher, ISA Bus	Win 3.x, 95 & NT	1 - or - 3 BRI ISDN, IsoEthernet, Switched 56	H 320	TalkShow, FarSite, ProShare, NetMeeting
Datapoint Corp	MINX Networked Video System	\$ 1,750 (camera, speakers, mike, video, overlay)	LAN and WAN Videoconferencing system high res. full-motion video to desktops in enterprise campus setting	Intel PC	Win 3.x, 95 & NT, MacOS OS/2 for windows	Unshielded twisted pair or coax in-house	Proprietary modulation in-house, H 320 via gateway to PSTN	Share Whiteboard Interactive documents, document camera, shared video resources
DynaVision	DynaVision	\$ 200 (1 Unit) 330 for a set (1 or 2 cameras, software, cables)	Audio/Video system designed to operate over standard analog telephone lines or various network situations	Macintosh	MacOs	UDP/IP, IP Multicast	Proprietary protocol for point to point modem connections	N/A
The Electronic Studio, London	ES + F2F, (Electronic Studio's Face 2 face), version 1.0	\$ 995 (video), \$ 995 (text & image exchange) \$ 1,495.00 (both)	Video/tools over ISDN, analog, Ethernet, audio requires ISDN or analog phone line	Macintosh	Apple Communications Toolbox	BRI ISDN, Analog phone, Ethernet	Proprietary	Text and image exchange
ELSA	ELSAVision Version 1.30	Call for pricing (PCI board, color camera, headset, software)	single slot PCI board solution for videoconferencing, standard telephone, ISDN board, MPEG decoder, two inputs for analog cameras, speakerphone	Intel PC, PCI slot	Win 3.1, 95 & NT	ISDN, IPX/PP, TCP/IP	H 320	Intel ProShare, Premier
EyeTel Communications	Insight 100 version 3.5	\$ 129 (software), \$ 369 (add video card)	Software based videophone program for POTS or cellular phone connections	i486/80, PCI or ISA slot 8 MB RAM, 6 MB disk	Win 3.x	28.8 Kbps modem	Proprietary	Phone box, chat box

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
	Insight 300 version 3.5	\$5,495 (codec, camera, mike, speakers)	Video/audio/tools over Switched 56, ISDN, T1, Ethernet or Token Ring	1486/50, 3 ISA slots, 8 MB RAM, 10 MB disk	Win 3.x, TCP/IP stack	Switched 56, ISDN, T1, Ethernet	H.320	Chat, App Sharing, whiteboard, file transfer, clipboard
Fiber & Wireless	Mediafone Version 7.1q	\$ 124 (software only)	Video/audio conferencing and collaboration over analog phone, ISDN, Internet or LAN/WAN	1486/33 PC, 4 MB RAM, 12 MB disk, 256 KB cache	Windows	BRI ISDN, LAN/WAN, 14.4 Kbps or higher modem	H.261 video G.723 audio	Whiteboard App Sharing
Incite Division of Intecom	Incite Conversational Multimedia Network	\$ 3,450 (codec, camera, speakerphone, NIU)	Videophone and video distribution over IsoEthernet using Zydracron codes	1486/66, 32 MB RAM, 70 MB disk	Win 3.11, NT 3.5.1	IsoEthernet 64 Kbps multiples	H.320	Whiteboard online image editing
InSoft	Communique!	Call for pricing (platform dependent)	Integrates real-time digital video technology with fully interactive, real-time collaborative tools	SUN, HP, IBM 6000, Digital, SGI, Intel PC	Unix, Win 3.1, 3.11, 95 & NT	Packet, Cell, PSTN	H.320, CellB, Motion JPEG, DVE, Indeo	App Sharing, shared whiteboard, chat, audio, text & graphics
Intel Corp	ProShare Video System 200, version 2.0	\$ 1,499 (2 ISA Boards, camera, earpiece/mike)	Audio/videoconferencing over LAN/WAN/ISDN	1486/66, 16 MB RAM, 17 MB disk, Pentium 133 recommended	Win 3.x & 95	BRI ISDN, TCP/IP, Novell IPX	H.320	Whiteboard App sharing
	Intel Video Phone	≈ \$ 200 (bundled w/ht OEM PCs)	Audio/videoconferencing over POTS	Pentium 133, 16 MB RAM	Win 95	v.80 modem	H.324	N/A
	Intel Internet Video Phone	Free download from Intel	Audio/videoconferencing over The Internet	Pentium 133, 16 MB RAM	Win 95	TCP/IP	H.323	N/A
Intellect Visual Communications	Intellect LANscape	\$ 2,995	PC-LAN based videoconferencing on single board	Intel PC, Unix	Win 3.x, 95 & WFW	TCP/IP	H.320 Motion JPEG	T120 Collaborative software
Intelligence at Large	BeingThere, Version 2.0	\$ 299 (standard), \$ 599 (PRO), \$ 149 (starter kit)	Video/audio/tools over LAN, ISDN and other switched digital telephone lines	Mac 68040 Power Mac, Quick Time video	MacOs 7.1 or later	9.6 BRI ISDN, POTS; 28.8 Kbps modem (9.6 Kbps collaboration only)	Proprietary	Whiteboard drag & drop file transfer, clipboard transfer
InVision Systems Corp	InVision, Version 3.0	\$ 595 (software only)	Video/audio/tools over Ethernet, Token Ring, FDDI, Frame Relay, ATM, ISDN, V.32 or faster modem	1486/33, 8 MB RAM, 3 MB disk	Win 3.x	ISDN, FDDI, Frame Relay, ATM, TCP/IP	Proprietary	Vision Graphics document-sharing software

Sistemas de Video Digital
Capítulo 2

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE	
Microsoft Corp	NetMeeting, Version 2.0b4	Download from www.microsoft.com/msdownload	Video/audio/data collaboration tools over Internet, modem or IP over ISDN, can receive video without special hardware	Macintosh 90 MHz Pentium, 24 MB RAM	MacOs, Win 95	LAN/WAN, IP over ISDN, 28.8 Kbps modem	H 323	App Sharing, whiteboard, chat, file transfer	
Peregrine Systems	Ntv	Call for pricing	Video/audio/tools over Ethernet, Token Ring	i386/25, 4 MB RAM	Win 3.x, WFW	LAN/WAN	Proprietary	App Sharing,	
PictureTel Corp	PictureTel Live 200 Version 1.1	\$ 1,495 (one board, camera, headset or speakers, mike, software)	Single board solution for video and audio communications, plus interactive information sharing	Intel PC, PCI Bus 16 MB RAM 20 MB disk	Win95	BRI ISDN	H 320	Shared Whiteboard & clipboard, app Sharing, drag & drop file transfer, messaging, remote control	
	PictureTel Live 100 Version 1.6	\$ 4,995 (codec & comm Board, camera, headset or speakers, mike, software)	Two board solution for video and audio communications plus interactive information sharing networks at rates from 56 to 384 Kbps	i386/225 PC 2 ISA slots 8 MB RAM 20 MB disk	Win 3.x, & 95	Switched 56 ISDN, V.35/RS449	H 320 PT 724 audio	Whiteboard, shared clipboard, app Sharing, file transfer, messaging, remote control	
	PictureTel Live 50 Version 1.6	\$ 2,495 (one board, camera, headset or speakers, mike, software)	Single board solution for video and audio communications plus interactive information sharing networks at rates from 56 to 128 Kbps	i386/225 PC ISA slot, 8 MB RAM 20 MB disk	Win 3.x, & 95	Switched 56, BRI ISDN, V.35/RS449	H 320 PT 724 audio	Whiteboard, shared clipboard, app Sharing, file transfer, messaging, remote control	
	PictureTel Live LAN Version 3.0	\$ 1,495 (codec card, camera, speakers, mike, software)	LAN Ethernet H 323 compliant videoconferencing product with integrated T.120-compliant multipoint data collaboration tools	Pentium PC, PCI Bus 16 MB RAM 25 MB disk	Win 95	WinSock 1.1-compliant stack	TCP/IP	H 323	Whiteboard, shared clipboard, app Sharing, file transfer, messaging T.120 multipoint
	RSI Systems	Eris Visual Communications System, version 1.8	\$ 3,995	Standalone codec that can use PC via SCSI bus	Intel PC, Macintosh running System 7.0 Sun Workstations	Windows, MacOs System 7.0, SunOs	BRI ISDN	H 320	Cross-platform document sharing & file transfer

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
SAT-SAGEM	Meet-Me, Version 2 02	\$ 1,850 (NuBus codec, handset, camera)	Video/audio collaboration tools over ISO Ethernet or ISDN	Mac, (PCI & NuBus), Mac AV	MacOs	LAN/WAN, ISDN (NT1 is bundled)	H.320	Whiteboard high-speed file transfer
	Meet-Me, Light 1 0	\$ 295 software only	Software-only videoconferencing solution Macintosh	100 Mhz Mac with AV capability	MacOs	LAN/WAN	H.320	N/A
Shark Multimedia	SeeQuest	\$ 399 (Connectix Color QuickCam, Baby Tiger modem)	Real-Time full-duplex audio videoconferencing Kit	1486/66 PC, 16 MB RAM, CD ROM	Win 3 x & 95	Baby Tiger 33.6 Kbps DSVD modem	Proprietary video, V 70 DSVD modem	fax, image capture, voicemail, internet browser, TalkShow Whiteboard, QuickPICT
Siemens Nixdorf	jointX, version 2 2	Call for pricing	Multipoint desktop conferencing system	Sun SGI, HP, IBM AIX	SunOs Solaris, HP-UX, AIX	ISDN, JPEG/UDP, CellB/UDP	H.320	X11 app. Sharing, shared group filestore, chatbox & shared Whiteboard
Silicon Graphics	InPerson version 2 0	\$ 495	Video/audio/tools over ISDN, T1, Ethernet, FDDI	Intel PC, SUN, HP, any SGI platform	Win 3 11, 95 & NT IRIX 5.3	UDP/IP, Whiteboard TCP/IP	H.320, SGI's HDCC	SGI MindShare collaborative environment
Sun Microsystems	ShowMe version 2 01	\$ 3,270 (Sun Video board and camera)	Video/audio tools over the Internet	Sun SPARC station	Solaris 2 3 or later	UDP/IP, TCP/IP, Multicast, RTP	Proprietary	Whiteboard, X11 R4/R5 app sharing
Teles Communications Corp	TELES, Vision-85	\$ 599 (Color Camera, headphones/mike)	Single PCI Board solution including terminal adapter and frame grabber for two-way audio and video	Pentium 133, PCI slot, 16 MB RAM 30 MB disk	Win 3 1x, 95 & NT	BRI ISDN	H.320	App. Sharing Remote Control
VCON	VCON Cruiser Armada 150	\$ 1,895	Video/audio/data collaboration tools over IsoEthernet, ISDN, ATM	Pentium PC, 16 MB RAM 2 MB VRAM	Win 3 x, & 95	BRI ISDN, ATM, LAN/WAN	H.320 T 120	FarSite, NetMeeting
VSI	Omega MVP	\$ 2,895	Desktop videoconferencing system	i486/66, 16 MB RAM, 4 MB disk	Win 3 x, & 95	BRI ISDN	H.320	FarSite, ProShare, TalkShow LiveShare Plus
VTEL Corp	VTEL Personal Collaborator Version 1 0	\$ 2,495 (codec, camera, speaker, mike)	Single board videoconferencing solution for win95 based PCs	Intel PC, ISA or PCI slot	Win95	BRI ISDN	H.320	VTEL ObjectShare Intel ProShare

Sistemas de Video Digital
Capítulo 2

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOLS SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
VIC Hi-Tech Corp	Global Phone Version 1.5	\$ 149	Software for simultaneous voice, data and video over DSVD modem	LapTop PC, with PC Card video capture, V 2.0, 1485 PC	Win 3.x, 95 & NT	DSVD Modem	H 320	Whiteboard file transfer, 2-way chat, app Sharing in GP 2.0, 2 way chat boxes
	Global Phone Version 2.0	\$ 595	Hardware codec for simultaneous voice, data and video over ISDN or TCP/IP	486DX 25+ PC, 8 MB RAM	Win 3.x, 95 & NT	ISDN, TCP/IP	H 320	Whiteboard file transfer, 2-way chat, app Sharing
	Global Phone Lite Version 2.0	\$ 595	H 324 compliant visual communications over POTS	Pentium with 28.8 Kbps modem, preferably synchronous	Win 3.x & 95	POTS	H 324	Full version will support T 120 app & whiteboard sharing
VideoLive	VDCPhone Version 2.0	\$ 79 (buy one, get one free)	Software for video audioconferencing over Internet or direct modem connection	Pentium 75, 16 MB RAM	Win95	LANWAN, 28.8 Kbps modem	Proprietary VDO wavelets	NetMeeting
VideoCam	VideoCam VCI-10	\$ 2,950	OEM Video codec supporting up to 384 Kbps	i486/33, 1 ISA slot	Win 3.1 & 95	API to IP or LANWAN	H 320, T 120	T 120, T 123, software
Visual One	InterVideo	\$ 5,999	Full motion, full screen videoconferencing system over LAN or ISDN WAN	Intel PC, Unix Systems	Windows, Unix	ISDN, TCP/IP, IPX/SPX	H 320	Info Share Plus Data Collaboration software
Vivo Software	Vivo 320	\$ 1,495	Videoboard & tools over ISDN	i486/66 PC, 8 MB RAM, VESA or PCI display adapter, 2 ISA slots	Win 3.1	BRI ISDN	H 320	Image Prescription markup, FarSite 2.0 collaboration included
White Pine Software	Enhanced CU-SeeMe, v2.1 Windows v 2.0 Mac	\$ 69 (no Internet, bundled with some OEMs)	Person-to-person or group conferencing over the Internet using a variety of platforms	Intel PC, Macintosh	Win 3.1, 95 & NT, MacOS	TCP/IP, UDP/IP, IP Multicast	White Pine Color software codec	White-Pine Board Chat Window
Winov	Vidum AV	\$ 199	Video and audio capture board for videoconferencing and multimedia authoring, uses Enhanced CU-SeeMe software	Pentium 90, 16 MB of RAM	Win 3.1, 95 & NT 4.0	TCP/IP, UDP/IP, IP Multicast	H 323	Enhanced CU-SeeMe White Pine Board
	VidumConf Pro	\$ 399	Internet videoconferencing and multimedia authoring kit includes Vidum AV, VidumCam and Enhanced CU-SeeMe	Pentium 90, 16 MB of RAM	Win 3.1x, 95 & NT 4.0	TCP/IP, UDP/IP, IP Multicast	H 323	Enhanced CU-SeeMe White Pine Board

COMPANY	PRODUCT NAME/MODEL	LIST PRICE	SYSTEM DESCRIPTION	PLATFORM SUPPORTED	OPERATING SYSTEM SUPPORTED	TRANSMISSION PROTOCOL SUPPORTED	STANDARD SUPPORTED	COLLABORATION TOOLS AVAILABLE
Wintronix	XIX Internet Communications Suite, Version 1.2	\$ 69 (Downloaded from the Internet)	Software-only multipoint videoconferencing solution for the Internet	Pentium 75, 16 MB RAM, 10 MB disk	Win95 w/ Video for Windows	TCP/IP, UDP/IP, IP Multicast	Proprietary	Whiteboard chat, audio/video recording
Zydacron	INTERVu	Call for pricing (Zydacron is a manufacturer for OEMs)	Computer-independent video/audio and collaboration over IsoEthernet (802.9) ISDN, Switched 56, or V.35/RS366	Intel 386 or higher	Win 3.x, 95 & NT	Switched 56, ISDN, IsoEthernet	H.320	Runs Talkshow, Vis-a-Vis, Person to person, FarSite, Terminal, Carbon Copy, ProShare, etc

Bibliografía.

1. NIH Telecommunications Branch
Video Conferencing Overview
Sept. 1997
[wysiwyg://13/http://tcb.od.nih.gov/vidover.html](http://tcb.od.nih.gov/vidover.html)
2. Pacific Bell
A Brief Description of Videoconferencing
Agosto 1996
<http://www.kn.pacbell.com/wired/vidconf/description.html>
3. NIH Telecommunications Branch
LAN-Based video conferencing Overview
Junio 1997
[wysiwyg://23/http://tcb.od.nih.gov/lanvideo.html](http://tcb.od.nih.gov/lanvideo.html)

4. Pacific Bell
Compressed Video
Mayo 1996
<http://www.kn.pacbell.com/wired/vidconf/compressedVid.html>

5. Bell Atlantic Corporation
DVC Technology
1996
<http://www.brookwood.com/desktech.htm>

6. ATMs and Video Conferencing
<http://www.ee.mtu.edu/courses/ee465/groupk/atm.html>

7. Advanced Video Conferencing
<http://yerkes.mit.edu/NARCT/Technodistribution/advancedvideoconf.html>

8. Network Computing
Designing a Videoconferencing solution
<http://pubsys.cmp.com/nc/net design/video4.html>

9. Network Computing
Junio de 1997
<http://www.Network Computing.Com>

CAPÍTULO 3

COMUNICACIÓN MULTIMEDIA VÍA REDES DIGITALES DE SERVICIOS (N-ISDN).

En el presente capítulo se hablará de las características de los sistemas y equipos terminales, estructura de trama, técnicas de multiplexaje, protocolos para inicialización y desconexión; intercambio de información dentro de banda y administración del canal, en Banda Estrecha. En Banda Estrecha las velocidades binarias están comprendidas entre los 64 Kbps y 1920 Kbps, esta capacidad de canal puede establecerse en forma de un solo canal B/Ho/H11/H12¹ ó múltiples canales B/Ho en la ISDN (Integrated Services Digital Network; por sus siglas en inglés).

¹Canales B:64 Kbps, H₀:384 Kbps, H₁₁:1536 Kbps, H₁₂:1920 kbps.

La ISDN es un servicio de datos digitales de alta velocidad prestados típicamente por una compañía telefónica. En la ISDN se puede transmitir grandes cantidades de datos, voz y señales de vídeo sobre una sola línea telefónica, a altas velocidades, y a menor costo que cualquier otro servicio digital o módem análogo.

Aplicaciones típicas de ISDN:

- Acceso a Internet.
- Teleconmutación.
- Enrutamiento de oficinas remotas.
- Conferencias de vídeo.
- Pantalla compartida y colaboración de PC a PC.
- Voz y comunicación de datos simultáneos.

Hay tres clases de servicio en la ISDN: Tasa de interfase básica ("Basic Rate Interface") (BRI), Tasa de Interfase Principal ("Primary Rate Interface") (PRI), y Multirate. Los servicios PRI y Multirate son típicamente usados por oficinas grandes, para soportar solicitudes de gran ancho de banda, tales como sistemas telefónicos PBX. Estos sistemas son más costosos y por lo tanto, menos disponibles, en términos generales que el servicio BRI (Ver figura 3.1).

A continuación se muestra la configuración básica de la ISDN, mostrando las interfases¹ que la ISDN ocupa.

¹ Ver Glosario de Términos.

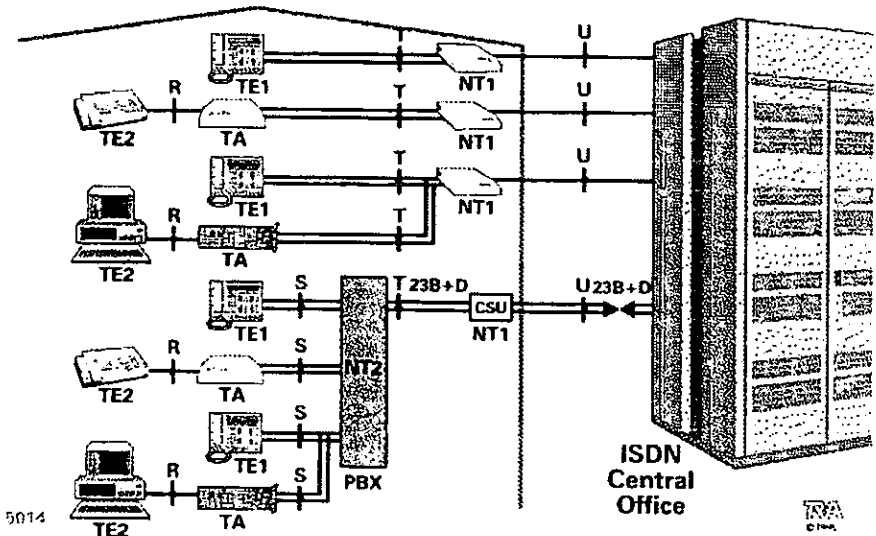


Fig.3.1. ISDN Configuración de referencia

Servicio BRI

Una sola línea BRI puede adaptar hasta 128 Kbps sobre dos canales portadores virtuales (B) de 64 Kbps. Cada canal B puede simultáneamente transmitir datos y señales de fax, junto con voz digitalizada y tráfico de video. Un solo canal (D) de 16 Kbps es asignado para soportar funciones del sistema, tales como señalización para iniciar una llamada; este servicio BRI es conocido también por línea 2B+D (ver figura 3.2).

Lo que distingue a la ISDN es que cada canal B es un circuito de comunicación por separado. Eso significa que solamente una línea ISDN puede soportar comunicaciones simultáneas de dos vías para

dos dispositivos, tales como una computadora y un teléfono, o una computadora y una cámara de video para teleconferencias.

A continuación se muestra el servicio básico BRI.

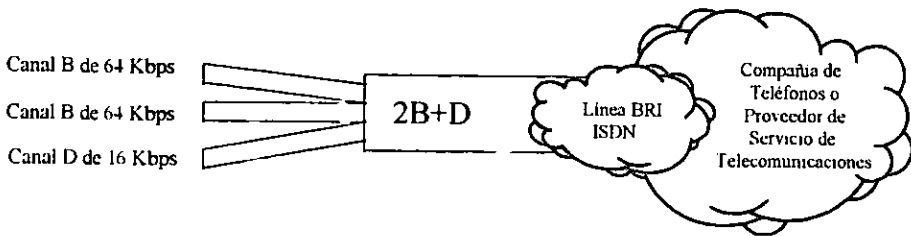


Fig.3.2 Servicio BRI 6 (2B+D).

Adicionalmente, si se necesitan enviar más datos de los que puede manejar un canal B de 64 Kbps, ISDN también soporta estándar **BONDING**, para **multiplexado invertido**. Este enlaza dos canales B a un solo circuito lógico que puede soportar velocidades de transmisión de hasta 128 Kbps. Se pueden usar multiplexores ISDN, con lo que se podrán fusionar múltiples líneas ISDN para velocidades máximas de transmisión de datos hasta de 512 Kbps. El servicio más común y menos costoso es BRI. Su aplicación más usual es de combinación de voz y datos sobre una sola línea entre lugares a pequeña escala.

CONTEXTO RDI EN MEXICO

TELMEX, promovió la solución de una red superpuesta que crecería sobre la red de cobre existente para efectuar los enlaces de comunicación en las organizaciones más importantes de México. Esta red pasó a ser la Red Digital Integrada (RDI), red que creció de los

33,600 enlaces en 1990, a más de 400,000 accesos en 1995 en base a la fibra óptica para canales E1 (2.048 Mbps).

La infraestructura RDI, sólo es una parte de la Red Digital de Servicios Integrados (Integrated Services Digital Networks-ISDN), que por un par de cobre entrega servicios de comunicación a 128 Kbps, en un contacto (roseta) estándar. RDI empezó siendo muy caro, en un principio se entregaba un E0 (contenido en un E1 a 32 canales), así como la infraestructura necesaria para RDI: tierra física, site RDI, ajustes a la instalación, equipo, etc.

Hoy en día, TELMEX opera modificaciones en los precios, y una fuerte promoción de lo DSO's (enlaces a 64 Kbps), para optimizar el costo de adquisición a los usuarios del mínimo ancho de banda requerido para sus aplicaciones.

Es importante comprender básicamente la ISDN, ya que los estándares explicados en este capítulo se basan en la transmisión sobre la ISDN.

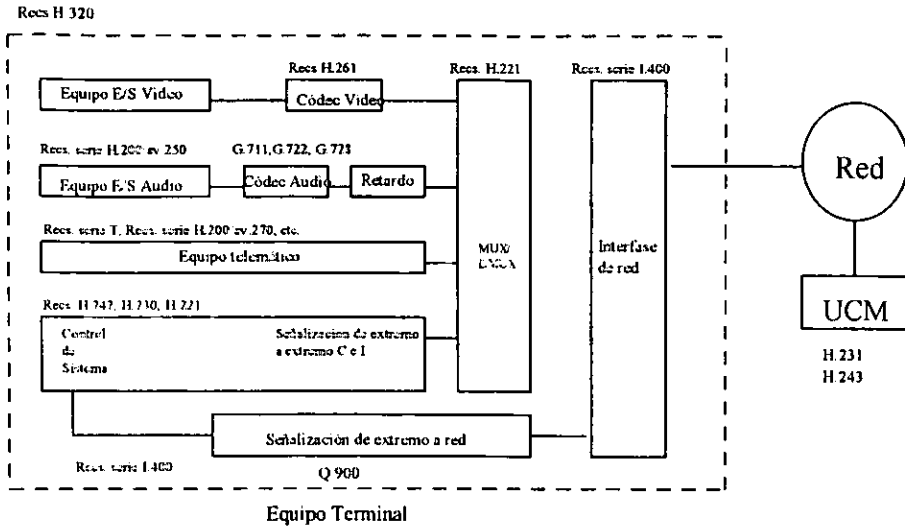
3.1. ESTÁNDAR H.320: SEÑALES VIDEOTELEFÓNICAS.

3.1.1. Descripción del Sistema

La recomendación H.320 abarca los requisitos que deben de cumplir los sistemas videotelefónicos¹ de banda estrecha, en los que las velocidades del canal no exceden de 1920 Kbps.

¹ Estos sistemas videotelefónicos están definidos en las recomendaciones H.200/AV.120, la explicación de dichos sistemas queda fuera del alcance de esta Tesis, por lo que es necesario referirse a dicha norma para más explicación.

Dichos requisitos se muestran en la figura 3.1.1 en la que podemos observar las normas que deben de cumplir cada uno de los elementos del sistema videotelefónico.



UCM Unidad de control multipunto

Fig.3.1.1 Recomendación H.320

Códec de vídeo.- De acuerdo a la recomendación H.261

Códec de audio.- De conformidad con las recomendaciones G.711, G.722, H.200/AV.254 y AV.253²

Estructura de trama.- De acuerdo a la recomendación H.221

Control e Indicaciones (C & I).- Se utiliza la recomendación H.230

²La explicación de estas normas queda fuera del alcance de esta tesis, ver directamente la norma.

Procedimiento de comunicación.- De acuerdo con la recomendación H.242.

El equipo de E/S de vídeo comprende monitores, cámaras y unidades de tratamiento del vídeo que realizan funciones como la división de pantalla. El equipo de E/S de audio comprende micrófonos y altavoces, así como, unidades de tratamiento de audio que cancelan el eco acústico. El equipo telemático comprende pizarras electrónicas, tranceptores de imágenes fijas. La unidad de control efectúa funciones tales como el acceso a red por medio de una señalización de extremo a red y un control de extremo a extremo para establecer un modo común de funcionamiento y su señalización necesaria. El códec de vídeo codifica y decodifica las señales de vídeo con reducción de redundancia, y el códec de audio hace lo mismo con las señales de audio. El retardo en el trayecto de audio compensa el retardo del códec de vídeo para mantener la sincronización con el movimiento de los labios. La unidad mux/demux multiplexa las señales de audio, datos, vídeo y de control que han de transmitirse como un tren binario único. La interfase de la red efectúa la adaptación necesaria entre la red y la terminal.

3.1.2. Clasificación de las señales videotelefónicas

- Señales de audio.- representan un tráfico continuo y exigen transmisión en tiempo real.
- Señales de vídeo.- representan un tráfico continuo, es conveniente dar a las señales de vídeo la velocidad más alta posible, a fin de obtener mejor calidad con la capacidad disponible del canal.

- Señales de datos.- comprenden imágenes fijas, facsimiles de documentos y otras facilidades que están relacionadas únicamente con las mejoras optativas del sistema videotelefónico básico.
- Señales de control.- el trayecto de las señales de control de terminal a red se establece mediante el canal D, mientras que el trayecto de las señales de terminal a terminal se establece mediante la SAB¹ o el canal de servicio y únicamente cuando así lo exige el mecanismo de la recomendación H.221.

3.1.3. Establecimiento de una comunicación videotelefónica

Fase A: establecimiento de la comunicación, señalización fuera de banda;

Fase B1: iniciación del modo del canal inicial;

Fase CA: establecimiento de la comunicación con uno o más canales adicionales;

Fase CB1: iniciación de uno o más canales adicionales;

Fase CB2: establecimiento de los parámetros comunes;

Fase C: comunicación videotelefónica;

Fase D: terminación;

Fase E: liberación de llamada.

3.1.4. Mejoras Facultativas

Puerto de datos.- Los puertos de datos, como puertos físicos de E/S de la terminal para la conexión del equipo telemático y de otra naturaleza, se activan y desactivan por medio de instrucciones SAB. Según cual sea la capacidad de transmisión de una conexión, en estos puertos se disponen de diversas velocidades binarias (múltiplos de los canales B/H₀). La atribución de trenes binarios a los puertos se efectúa por medio de la señalización en banda. Los datos transmitidos por los

¹ SAB: Señal de asignación binaria (Ver punto 3.2. de este capítulo).

puertos son transparentes; las velocidades de transmisión de datos se indican en el anexo A de la recomendación H.221.

Encriptación.- Se puede aplicar encriptación a las señales de audio y video por separado (de preferencia para conexiones multipunto) o a las señales audio y video multiplexadas. La activación y desactivación del proceso de encriptación debe señalizarse entre las terminales (o entre terminal y el UCM) por medio de señalización dentro de banda.

3.1.5. Intercomunicación¹

Intercomunicación con teléfonos para ISDN.- La comunicación de un videoteléfono a un teléfono para ISDN se establece primeramente como una comunicación audiovisual, pero el teléfono para ISDN devuelve la indicación "destino incompatible" o bien la red devuelve la indicación "recuperación" al expirar el temporizador si no hay respuesta del lado llamado, tras lo cual, el videoteléfono puede pasar la llamada a un servicio portador de voz o a un servicio portador de audio a 7 Khz.

La llamada de un teléfono para ISDN a un videoteléfono será aceptada por este último, porque toda terminal audiovisual posee la capacidad de telefonía a modo mínimo. En ambos casos, el modo operacional de comunicación es el de conversación, de acuerdo con la recomendación G.711, o el de audio, de acuerdo a la recomendación G.722.

Intercomunicación con Red Telefónica Pública Conmutada (PSTN).- La llamada de un videoteléfono a un teléfono de la PSTN puede iniciarse como una llamada audiovisual, pero la red devuelve la indicación "no hay ruta hacia destino", tras lo cual el videoteléfono puede pasar la llamada a un servicio portador de voz o a un servicio

¹ La descripción que aparece en este punto se aplica a las comunicaciones efectuadas mediante un canal B.

portador de audio de 3.1 KHz. El modo de operacional es de acuerdo a la recomendación G.711.

La llamada procedente de teléfono para PSTN se dirige hacia la ISDN en forma de una llamada de audio a 3.1 KHz, que puede ser contestada por el videoteléfono. El modo de operación es el de audio a 3.1 KHz.

3.2. H.221 TELESERVICIOS AUDIOVISUALES

La finalidad de esta recomendación es definir la estructura de trama para los teleservicios audiovisuales transmitidos por uno o múltiples canales B, H₀ o por un solo canal H₁₁ o H₁₂, utilizando de la mejor manera posible las características y propiedades de los algoritmos de codificación de audio y de vídeo.

3.2.1. Principio básico

Esta recomendación permite subdividir dinámicamente un canal de transmisión global de 64 a 1920 Kbps, en velocidades inferiores adecuadas para fines de audio, vídeo, datos y telemáticos. El canal de transmisión global se obtiene sincronizando y ordenando transmisiones de 1 a 6 conexiones B, de 1 a 5 conexiones H₀, de una conexión H₁₁ o H₁₂. La velocidad total de la información transmitida se denomina "velocidad de transferencia" y es posible fijar la velocidad de transferencia en un valor inferior a la capacidad del canal de transmisión global.

Un canal único de 64 Kbps está estructurado en octetos transmitidos a 8 KHz. Cada posición de bit de los octetos puede considerarse un subcanal de 8 Kbps (ver figura 3.2.1). El octavo subcanal se denomina canal de servicio (CS), y consta de varias partes que se describen más adelante.

Número de bit								Número de octeto
1	2	3	4	5	6	7	8(CS)	
S	S	S	S	S	S	S	SAT	1
u	u	u	u	u	u	u		SAB
b-	b-	b-	b-	b-	b-	b-	SCE	
C	C	C	C	C	C	C		#8
a	a	a	a	a	a	a	#8	
n	n	n	n	n	n	n		#8
a	a	a	a	a	a	a	#8	
l	l	l	l	l	l	l		#8
#1	#2	#3	#4	#5	#6	#7		

SAT Señal de asignación de trama
SAB Señal de asignación de velocidad binaria
SCE Señal de control de encriptación

Fig 3.2.1 Estructura de trama de un canal único de 64 Kbps (canal B)

Un canal H_0 , H_{11} , H_{12} , puede considerarse constituido por intervalos de tiempo a 64 Kbps (Ver figura 3.2.2). La estructura del intervalo de tiempo de número más bajo es la misma que la de un canal único de 64 Kbps; los otros intervalos no tienen esa estructura. En el caso de múltiples canales B o H_0 , todos los canales tienen una estructura de trama; la estructura de trama del canal inicial controla la mayor parte de las funciones de la transmisión global, mientras que la estructura de trama de los canales adicionales se utiliza para sincronización, numeración de canales y controles conexos.

complementados por un <<1>> en el bit 2 de la trama impar subsiguiente (Ver figura 3.2.3).

- Bit A: El <<bit A>> del canal I se pone a cero cuando el receptor está en alineación de multitrama y a <<1>> en caso contrario.

- Bit E y Bit C: Con el fin de proporcionar una supervisión de calidad de extremo a extremo de la conexión de 64 Kbps, se puede utilizar un procedimiento de Verificación por redundancia cíclica de 4 bits (VRC4) y los cuatro bits C1, C2, C3, C4 computados en la posición de origen se insertan en las posiciones de bit 5 a 8 de las tramas impares. Además, el bit 4 de las tramas impares (el bit E), se utiliza para transmitir una indicación de si el bloque VRC más reciente recibido en sentido entrante, contiene o no errores. Cuando no se utilice el procedimiento VRC4, el transmisor pondrá a 0 el bit E y a 1 los bits C1, C2, C3 y C4.

	No. de bit	1	2	3	4	5	6	7	8
Tramas sucesivos									
Tramas pares	Nota 1		<div style="border: 1px solid black; padding: 2px; display: inline-block;"> 0 0 1 1 0 1 1 </div>						
			Palabra de alineación de trama- Nota 2						
Tramas impares	Nota 1	1	A	E	C1	C2	C3	C4	
		Nota 2	Nota 3	Nota 4					

Figura 3.2.3 Asignación de los bits 1-8 del canal de servicio, estructura de la SAT.

Además de la información de alineación de trama y multitrama, en la señal SAT se puede insertar información de control y alarma, así como información de verificación de error. Los otros intervalos de tiempo se alinean al primer intervalo. Los bits se transmiten a la línea por su orden, comenzando por el bit 1.

Cuando se cuenta con un reloj de red de 8 KHz, la SAT se transmite y se recibe en el bit menos significativo del octeto de cada período de 125 microsegundos, por ejemplo, en una interfase ISDN de velocidad primaria.

3.2.3. Señal de Asignación de Velocidad Binaria (SAB)

Los bits 9 a 16 del canal de servicio (CS) en cada trama constituyen la señal de asignación de velocidad binaria (SAB). Esta señal permite la transmisión de palabras de código para describir las características de una terminal y así estructurar la capacidad del canal, o de múltiples canales sincronizados de diversas formas, y para ordenar a un receptor que demultiplexe y utilice las señales constitutivas de esas estructuras. La SAB se utiliza también para las señales de control e indicación (C&I).

Un código SAB de 8 bits (b_0, b_1, \dots, b_7) está complementado por ocho bits de corrección de errores (p_0, p_1, \dots, p_7), el código SAB se envía en la trama par, mientras que los bits de corrección de errores asociados se envían en la trama impar subsiguiente.

3.2.4. Señal de Control y Encriptación (SCE)

Esta capacidad de encriptación puede requerir un canal de transmisión dedicado. Se prevé que se tendrán que proporcionar 800 bps cuando sea necesario, atribuyendo los bits 17-24 del canal de servicio.

3.2.5. Capacidad Restante

La capacidad restante (incluyendo el canal de servicio), aportada por los bits 1 al 8 de cada octeto en caso de una conexión única a 64 Kbps, puede transportar diversas señales en el contexto de un servicio multimedia, bajo el control de la SAB.

3.3 H.242. ESTABLECIMIENTO DE COMUNICACIÓN ENTRE TERMINALES AUDIOVISUALES POR CANALES DIGITALES DE HASTA 2MBPS.

Se han identificado cierto número de aplicaciones que utilizan señales vocales de banda estrecha y de banda ancha, junto con vídeo y/o datos, incluida la telefonía de alta calidad, audio y videoconferencia (con o sin diversos tipos de ayuda telemática). En el futuro surgirán seguramente otras aplicaciones.

Para proporcionar estos servicios se recomienda un esquema en el cual un canal transmite señales vocales, y optativamente señales de vídeo y/o datos a diversas velocidades, en varios modos diferentes. Se necesitan procedimientos de señalización para establecer un modo compatible de llamada, para conmutar entre modos durante una llamada, y permitir la transferencia de llamada.

Todas las terminales de audio, y audiovisuales que utilizan codificación G.722 y/o G.711 y otras terminales con codificaciones normalizadas de audio deberán ser compatibles, para que sea posible la conexión de dos terminales de cualquier tipo. Por eso surge la necesidad de establecer un modo común de funcionamiento para la llamada. Es posible que sólo se utilice el modo inicial, o que se conmute a otro modo dependiendo de las capacidades de la terminal.

3.3.1. Capacidades de la terminal

La capacidad total de la terminal para recibir y decodificar diversas señales se comunica a la otra terminal mediante la transmisión de su conjunto de capacidades que consiste en el marcador de capacidad SAB seguido de todas las capacidades actuales. Los códigos se especifican en el anexo A de la recomendación H.221; el orden de transmisión es intrascendente excepto que los valores de formato de imagen tienen que ir seguidos por valores mínimos de intervalo de imagen.

Las capacidades de vídeo incluyen:

- Formato de imagen: CIF (un cuarto del formato intermedio común), o FIC (formato intermedio común);
- Intervalo de imagen mínimo (IIM): 1/29,97;2/29,97;3/29,97;4/29,97 segundos.

El valor de CIF tiene que ir seguido de un valor de IIM; el valor de FIC completo tiene que ir seguido de dos valores de IIM.

Capacidades de velocidad de transferencia.- La capacidad para recibir un número dado de canales de 64 Kbps incluye la capacidad para recibir un número menor de estos canales; de igual manera para los canales H_0 ; en ambos casos la terminal receptora sincronizará al canal inicial los canales adicionales conectados y mantendrá ese sincronismo durante todo el período de conexión.

Capacidades de datos.- Si una terminal puede aceptar más de una velocidad de datos de cualquier tipo, hay que incluir todos los valores pertinentes en el conjunto de capacidades; la declaración de un valor no incluye ningún otro.

3.3.2. Transmisión

Los modos de transmisión de audio, vídeo, velocidad de transferencia y datos se definen en la recomendación H.221.

La transmisión de vídeo está gobernada por las instrucciones de vídeo activado y vídeo desactivado. Cuando está activada, la señal de vídeo ocupa toda la capacidad, tanto en el canal inicial como en los canales adicionales, que no está asignada a otras señales o instrucciones. Por consiguiente, unas instrucciones de audio, velocidad de transferencia, velocidad de control de encriptación (SCE) y datos producen diferentes velocidades binarias de vídeo; la velocidad que efectivamente se obtiene, viene dada por la expresión: {velocidad de transferencia, menos velocidad de audio, menos velocidad de datos, menos SAT y SAB en todos los canales e intervalos en los que están presentes}. Los modos de velocidad de transferencia especifican la capacidad total de la submultitrama subsiguiente.

Establecimiento de los modos de funcionamiento compatibles

Este procedimiento consiste en:

- la transmisión de información sobre las capacidades de las respectivas terminales para la recepción y la decodificación de las capacidades de audio, vídeo, velocidad de transferencia, velocidades de datos y otras capacidades;
- la determinación de un modo de transmisión adecuado, concordante con las capacidades conocidas de ambas terminales.
- conmutación al modo antes determinado, estableciendo si son necesarios los canales adicionales.

Las terminales conectadas en una llamada pueden cambiar en el transcurso de una llamada. Esto puede requerir la reinicialización para determinar el tipo de terminal y restablecer el modo deseado de operación.

3.3.3. Estructura de trama

La estructura de trama definida en la recomendación H.221 se emplea para definir el multiplexaje de los diversos trenes de bits (audio, video, señal de control de encriptación y datos) dentro de la trama.

3.3.4. Secuencias básicas de señalización dentro del canal

Se definen tres secuencias de señalización, estas son:

- *Secuencia A- Intercambio de capacidad*

La secuencia de intercambio de capacidad obliga a la alineación de trama en ambos sentidos de la transmisión, y al intercambio de códigos de capacidad de terminal. Cualquiera de las dos terminales puede iniciar la secuencia.

La terminal X que inicia la secuencia de intercambio de capacidad tiene que, primero conmutar a un modo de alineación de trama si antes estaba transmitiendo sin alineación de trama; después arranca un temporizador de 10 segundos y transmite el conjunto actual de capacidades repetitivamente.

Cuando la terminal Y detecta por primera vez un código cualquiera de capacidad entrante, excepto uno neutro¹, comienza la transmisión de su propio conjunto de códigos de capacidad; para esto es necesario

¹ Canal I neutralizado, que contiene solamente SAT y SAB; todos los demás bits se ignoran en el receptor.

conmutar a un modo de alineación de trama. Para asegurar que cada uno recibe el conjunto completo de capacidades del otro, ésta tiene que continuar la transmisión repetitiva aún después de haber detectado $A=0$ por lo menos durante un conjunto completo y el marcador.

- *Secuencia B- Conmutación de modo*

La conmutación de modo se lleva a cabo utilizando códigos de instrucción SAB, cada uno de los cuales se hace efectivo a partir de la trama par que sigue a la submultitrama en que se transmitió el código por primera vez. La conmutación de modo es posible en cualquier momento de la comunicación, después de que se haya seguido el procedimiento de inicialización.

La terminal receptora decodifica y válida los códigos SAB y conmuta a su modo de funcionamiento en recepción como corresponda. Si por alguna razón cualquiera una terminal recibe una instrucción SAB que no pueda ejecutar, se puede producir la desadaptación del modo.

Además de la conmutación del modo audio, la conmutación de modo incluye: la activación o desactivación de video, la adopción o el cese de canales adicionales; la apertura o cierre del canal de control de encriptación; la apertura o cierre de un canal de datos.

- *Secuencia C- Restablecimiento de trama*

Si la terminal X esta transmitiendo sin alineación de trama pero recibiendo con alineación de trama, el restablecimiento de trama consiste en la inserción de SAT y SAB en los primeros 16 bits del canal de servicio, en espera de A entrante igual a 0; la trama superpuesta

puede contener capacidad SAB neutra, para no provocar un intercambio completo de capacidades.

3.3.5. Procedimiento de inicialización de modo

El procedimiento de inicialización comienza tan pronto como se ha recibido de la red un mensaje de conexión, o cualquier indicación de conexión física.

La parte receptora deberá estar en condición de búsqueda de trama, y la recepción de audio en modo OF. Se comienzan los pasos mencionados en el punto 3.3.4.

El procedimiento de inicialización se completa cuando ambas terminales han conmutado al (a los) modo(s) de trabajo deseado(s).

3.3.6. Procedimientos para la activación y desactivación de canales de datos

Una terminal puede transmitir datos a cualquier velocidad indicada en los códigos de capacidad de velocidad recibidos de la otra terminal. Se envía la instrucción de datos adecuada, y en la siguiente submultitrama comienza la transmisión de datos en los bits de cada trama definidos en la recomendación H.221. Sin embargo, en el momento que se envía por primera vez la instrucción de datos, estos bits a deben de estar desocupados o contener únicamente información de video, por lo que deben eliminarse señales de audio u otras de esta parte de la trama. En caso de ocupación por información de video, no se dispone de instrucciones para reducir la velocidad de video; no obstante, si la velocidad de video se reduce mucho (por ejemplo, a menos de 30.4 Kbps), o se anula por la introducción de un tren de datos, se envía

primero la petición de congelar imagen seguida de la instrucción de vídeo desactivado.

Al concluir la transmisión de datos se envía la instrucción de datos desactivado; si el vídeo esta activado, ocupará entonces los bits liberados en la siguiente submultitrama; y de allí en adelante, en caso contrario, estos bits seguirán desocupados hasta que se envíe otra instrucción.

3.4. H.230. SEÑALES DE CONTROL E INDICACIÓN PARA SISTEMAS AUDIOVISUALES.

El suministro de señales audiovisuales digitales se efectúa por un sistema de transmisión en el que las señales son multiplexadas, que utiliza una estructura de trama definida en la recomendación H.221. Además de vídeo, audio e información telemática, estas señales incluyen información sobre el funcionamiento del sistema llamadas señales de **Control e Indicación C&I**. Esta recomendación sólo trata los C&I de control de transmisión con sincronismo de trama, es decir, de respuesta rápida.

La presente recomendación detalla los C&I relacionados con el vídeo, audio, modos de transmitir números y caracteres, C&I para fines de mantenimiento para conferencias multipunto simples que no utilizan protocolo en el canal MLP (*multilayer protocol*), para agregación de canales y para transferencia de direcciones de red.

Las señales de control e indicación (C&I) pueden dividirse en tres grupos:

- a) Control de llamada (estas señales se tratan en las recomendaciones de la serie Q);
- b) Control de transmisión con sincronismo de trama o que requieren una respuesta rápida;
- c) Control de conferencia , datos y telemáticos que no requieran sincronismo de trama.

Esta recomendación sólo trata los (C&I) de la categoría b), que incluyen un conjunto simplificado de C&I de conferencia para conexiones multipunto de terminales simples.

A continuación se dará una breve explicación de las señales C&I más importantes.

3.4.2. Definiciones de los símbolos C&I

3.4.2.1. C&I relacionados con el vídeo

- **Indicación de vídeo suprimido (VIS, *video indicate suppressed*):** Este simbolo se utiliza para indicar que el contenido del canal de vídeo no representa una imagen de cámara normal; el codificador de vídeo puede haber quedado sin entrada de vídeo, o ésta puede haberse sustituido por un patrón generado electrónicamente.
- **Indicación de vídeo activo (VIA, *video indicate active*):** Es complementario de VIS. Cuando deba distinguirse entre varias fuentes de vídeo, la fuente es la designada como "VIA N° 1".
- **VIA2:** Es equivalente a VIA, pero la fuente es la designada como "vídeo N° 2".

- **Indicación de vídeo preparado para activación (VIR, *video indicate ready -to-activate*):** Este símbolo lo transmite una terminal cuyo usuario ha decidido no enviar vídeo a menos que el también reciba vídeo del otro extremo.
- **Instrucción de vídeo de petición de congelación de imagen (VFC, *video command «freeze picture request»*):** Este símbolo puede transmitirse antes de la conmutación al modo «vídeo desactivado», para preparar al decodificador de vídeo para este suceso¹; lo transmite también una unidad de control multipunto (MCU, *multipoint control unit*) antes de la conmutación a vídeo; cuando un decodificador terminal de vídeo recibe este símbolo, debe completar la actualización de la trama de vídeo en curso pero seguidamente tiene que visualizar la imagen congelada hasta que se reciba la instrucción de liberación de imagen congelada.
- **Instrucción de vídeo «petición de actualización rápida» (VCU, *video command «fast update request»*):** Este símbolo lo transmite un MCU después de haber conmutado a vídeo; también puede transmitirlo una terminal al principio de la comunicación cuando el decodificador de vídeo esta preparado para recibir por primera vez.

3.4.2.2. C&I relacionados con el audio

- **Indicación de audio silenciado (AIM, *audio indicated muted*):** Este símbolo se utiliza para indicar que el contenido del canal de audio no representa una señal de audio normal; el codificador de audio puede

¹ Si un decodificador recibe petición de congelación de imagen , congela las imágenes hasta que se recibe una señal de liberación de congelación de imagen o hasta que expira un período de temporización de al menos seis segundos.

haber quedado sin entrada de audio, o ésta se substituyó por un tono generado electrónicamente.

- **Indicación de audio activo (AIA, *audio indicate active*):** Es el complementario de AIM.

Instrucción de ecualización de audio (ACE, *audio command equalise*): La envía una terminal para pedir la ecualización del retardo de la señal de vídeo (sincronización con el movimiento de los labios).

- **instrucción de retardo nulo (ACZ, *audio command zero-delay*):** La envía una terminal llamante para pedir que la señal de audio no sea retardada y adaptarla al vídeo.

3.4.2.3. C&I para fines de mantenimiento

- **Instrucción de petición de bucle de vídeo (LCV, *loopback command, video loop request*):** Al recibir este símbolo, la terminal debe conectar la salida del decodificador de vídeo a la entrada del codificador de vídeo.
- **instrucción de petición de bucle digital (LCD, *loopback command, digital loop request*):** Al recibir esta instrucción, la terminal debe desconectar la salida del multiplexor del trayecto de salida, remplazándola por la entrada al demultiplexor; en el caso de múltiples B o H₀, se activa el bucle en cada conexión.
- **Instrucción de petición de bucle audio (LCA, *loopback command, audio loop request*):** Al recibir este símbolo, la terminal debe, si es posible, conectar la salida del decodificador de audio al codificador de audio.

Instrucción de bucle desactivado (LCO, *loopback command off*): Al recibir este símbolo el terminal debe desconectar todos los bucles y restaurar los trayectos de audio/vídeo y de datos a su estado normal.

3.4.3. Requisitos de los C&I

Las funciones C&I están definidas de tal manera que, en circunstancias especiales, el sistema audiovisual funcionará sin el menor contratiempo y será posible una presentación agradable al usuario. Por tanto, algunas funciones tienen que ser obligatorias y otras optativas. Aquí se aclaran las circunstancias en las cuales las funciones C&I son obligatorias.

- **CM Obligatoria condicionalmente (conditionally mandatory):** Si la terminal (o UCM) tiene la capacidad de entrar en un estado dado, tiene que transmitir el código dado y, al salir de ese estado, el código complementario; si no tiene dicha capacidad puede ignorar ambos códigos.
- **M Obligatoria (mandatory):** Para todos los equipos, sean del tipo terminal o UCM.
- **X No obligatoria:** Cuando se recibe este código, puede no ser reconocido, o ser reconocido y no provocar ninguna acción, lo que queda a la discreción del constructor o usuario.
- **NA El código no es aplicable en este caso.**
- **# Directividad de la señal C&I:** Véase la recomendación H.243 para saber si es obligatorio u opcional para la terminal o MCU.

Debe señalarse que, en la mayor parte de las terminales, sólo unos pocos requisitos son obligatorios. Todas las terminales audiovisuales tienen que reconocer y obedecer las instrucciones de establecer y de suprimir el bucle digital, y también los de establecer y de suprimir el bucle de vídeo si tienen la capacidad de vídeo.

GLOSARIO

Canal B.- Canal portador que permite voz real o tráfico de datos. Cada canal B puede funcionar a velocidades de hasta 64 Kbps.

Canal D.- Canal que corre a 16 Kbps. Usado para soportar funciones complementarias de ISDN, tales como señalamiento entre lugares del usuario y la compañía de Telecomunicaciones.

2B+D.- Abreviatura común para una línea ISDN de interfase de tarifa básica. Soporta dos canales B más un canal D.

Bonding.- Es un protocolo de norma industrial para amplitud de banda a demanda. El Bonding usa la técnica de multiplexado invertido para enlazar los dos canales B en una sola línea ISDN para formar un circuito de alta velocidad capaz de correr a 128 Kbps.

NT-1.- Es un dispositivo que convierte la interconexión ISDN (dos hilos) de la compañía telefónica a una conexión de cuatro hilos S/T para compatibilidad con los adaptadores terminales del local del cliente.

Interfase S/T.- Se refiere a la interconexión cuatro hilos de un adaptador terminal.

Interfase U.- La interconexión ISDN de dos hilos del intercambio local que es terminado en el local del cliente.

Adaptadores Terminales.- Equipo ISDN principal fabricado para el cliente, que permite al equipo del cliente (teléfono o PC), unirse a ISDN.

Banda estrecha.- Velocidades binarias comprendidas entre 64 Kbps y 1920 Kbps.

C&I.- Señalización de extremo a extremo entre terminales compuesta por un control que produce un cambio de estado en el receptor y una indicación que facilita información sobre el funcionamiento del sistema.

Señal de Asignación Binaria (SAB).- Posición de bit en la estructura de trama definida en la recomendación H.221, que se utiliza para transmitir, por ejemplo, instrucciones o señales de control e indicación.

Señalización dentro de banda.- Señalización por medio de la SAB de la estructura de trama de la recomendación H.221.

Señalización fuera de banda.- Señalización por medio de un canal que no forma parte del canal (corresponde a las Recomendaciones de la serie I.400).

UCM (Unidad de Control Multipunto).- Una parte del equipo situada en un nodo de la red o en una terminal que recibe varios canales de los puertos de acceso y que, de acuerdo con determinados criterios,

procesa las señales audiovisuales y las distribuye a los canales conectados.

Bibliografía.

1. UIT, H.320
Recomendación; Sistemas y equipos terminales videotelefónicos de banda estrecha.
Unión Internacional de Telecomunicaciones (UIT-CCITT);
Ginebra, 1990.
2. UIT, H.221
Estructura de trama para un canal de 64 a 1920 Kbps en Teleservicios Audiovisuales.
Unión Internacional de Telecomunicaciones (UIT-CCITT)
Ginebra, 1990.
3. UIT, H.242.
Sistema de establecimiento de comunicación entre terminales audiovisuales por canales digitales de hasta 2 Mbps
Unión Internacional de Telecomunicaciones (UIT-CCITT)
Ginebra, 1990.

4. UIT, H.230
Señales de control e indicación con sincronismo de trama para sistemas audiovisuales (anteriormente recomendación de la CCITT)
Unión Internacional de Telecomunicaciones (UIT-T)
Ginebra, 1995.
5. Catalogo Black Box; Invierno 1997
Black Box Corporation
Pittsburg, 1996.
6. HDSL, High Bit Rate Digital Subscriber Line
Distech Report
México D.F., 1996.

CAPÍTULO 4

COMUNICACIONES MULTIMEDIA VÍA RED TELEFÓNICA PÚBLICA CONMUTADA Y RADIO MÓVIL [ABL, 1996]

Las telecomunicaciones consisten en la transmisión de información a larga distancia. Actualmente convergen los diversos tipos de telecomunicaciones. La tarea de poner todas las piezas juntas para convertirlas en una sola red funcional, requiere combinar muchas señales que proceden de diversas fuentes.

En este proceso están implicados dos conceptos fundamentales: el multiplexaje o la combinación de muchas señales para transmitir las por una sola portadora y la conversión de las diversas señales a un formato común. Por ejemplo, las compañías telefónicas organizan sus redes digitales en arquitecturas jerárquicas altamente estructuradas y multiplexadas.

Nombre del Valor Nominal Jerarquía Digital Norteamericana	Velocidad de datos Binaria	Número de Circuitos
Circuito único	56.00 Kbps	1
T1 o DS1	1.5 Mbps	24
T2 o DS2	6.34 Mbps	96
T3 o DS3	45 Mbps	672
T3C o DS3C	90 Mbps	1,344
T4 o DS4	274 Mbps	4,032
400 Mbps	405 o 417 Mbps	6,048
565 Mbps	565 Mbps *	8,064 (56 Kbps equiv)
810 Mbps	810 Mbps *	12,098
1,700 Mbps	1,700 Mbps *	24,192
2,400 Mbps	2,400 Mbps *	36,290
Jerarquía Digital Norteamericana		
Circuito único	64.00 Kbps	1
E1	2.048 Mbps	30
E2	8.448 Mbps	120
E3	34.304 Mbps	480
E4	139.264 Mbps	1,920
E5	565.148 Mbps	7,680
SONET Jerarquía Digital Síncrona		
STS - 1/OC - 1	51.84 Mbps	672 (28 DS1 o DS3)
STS - 3/OC - 3 (STM-1)	155.52 Mbps	2,016
STS - 12/OC (STM-4)	622.08 Mbps	8,064
STS - 48/OC - 48 (STM-16)	2,488.32 Mbps	32,256
STS - 96/OC - 96 (STM-32)	4,976.64 Mbps	64,512
STS - 192/OC - 192 (STM-64)	9,953.28 Mbps	129,024

*La velocidad de transmisión real depende del diseño y de los bits suplementarios y no esta normalizada.

Tabla 4.1 Velocidad de transmisión binaria en Norteamérica y Europa.

Cuando AT&T estableció la jerarquía de teléfonos digitales norteamericana, nunca se convirtió en una norma mundial. El *Comité Consultivo Internacional de Telefonía y Telegrafía* (CCITT) estableció su propia familia de normas que giran en torno a la digitalización de un solo circuito de telefonía, a 64 Kbps, en lugar de 56 Kbps en América del Norte. Se efectuó la multiplexación de un diverso número de canales conjuntamente, utilizando diversas velocidades de transmisión de datos.

Descontenta por la falta de normas, que era caótica, la industria de telecomunicaciones redactó una nueva generación de normas para el funcionamiento a velocidades más altas y que puede dar cabida a muchas clases de señales, además del tráfico telefónico, a la que denominó *Synchronous Optical Network* (SONET) en América del Norte (Red Óptica Síncrona), en la tabla 4.1 se muestra la tabla con la lista de velocidades de transmisión con respecto a éstas normas.

Los sistemas de televisión analógica de hoy en día constan de un gran número de sistemas de transmisión digital. Para asegurar el uso eficaz de estas instalaciones, la velocidad binaria total para el vídeo digital debe por consiguiente, diseñarse de forma que sea compatible con la jerarquía digital.

Aunque la mayoría de los expertos sienten que la influencia del vídeo digital es realmente la llave del éxito final de las comunicaciones multimedia, el vídeo digital tiene aún un cierto camino que recorrer antes de reemplazar por completo al vídeo analógico. [Jeff, 1993]

Es importante señalar que la transmisión de señales de televisión es diferente a la transmisión de vídeo a través de las líneas telefónicas, por lo que se diferenciarán los términos televisión y vídeo.

Por un lado la televisión se encuentra asociada con el concepto de "broadcast", o bien la distribución por cable con su programación asociada a algún tipo de horario, y por otra parte el video esta asociado con la administración de los archivos, a la edición o bien a una programación de acuerdo a las necesidades personales y sus respectivos horarios.

Además, hasta que la televisión interactiva llegue a ser una realidad, los multimedios ofrecen en un corto periodo una promesa para extender el paradigma del video, porque dentro de este corto periodo de tiempo todas las computadoras tendrán incorporado el requisito de hardware y software para trabajar en multimedia.

La más simple y común visión de multimedia es la combinación de la potencia de las computadoras en conjunto con la capacidad de comunicación de video.

El término multimedia no se ha definido muy bien, tal vez la mejor descripción técnica es "la integración de dos o más diferentes medios de comunicación por un solo medio de transmisión a través de una computadora personal localizada en ambos extremos". Los componentes de los diferentes medios incluyen texto, gráficas, animación, voz, música y video.

Pero, la definición de multimedia en el contexto de la tecnología no hace justicia al concepto que describen las páginas de libros completos, hojas de papel con texto impreso en ellas, gráficas en movimiento, figuras animadas con voz y música de fondo, la transmisión de imágenes para aplicaciones en medicina, etc., ya que

los usos de la multimedia son tan diversos que se pueden hacer combinaciones o permutaciones de los mismos.

Las señales de vídeo codifican de forma continua imágenes y sonidos cambiantes, y como resultado, se exige un ancho de banda considerablemente superior a la correspondiente al de las señales telefónicas. Mediante la compresión digital se pueden reproducir las velocidades de transmisión de datos, por factores de 10 a 60 para la transmisión de vídeo y, por consiguiente, pueden disminuir los requisitos de ancho de banda. Cada canal de vídeo debe ser digitalizado, comprimido y compactado en una representación digital eficiente para que la transmisión tenga lugar a una velocidad digital binaria inferior.

Por otra parte, para la transmisión de las señales de televisión, el ancho de banda disponible debe dar cabida no solamente al vídeo, sino también a varias señales auxiliares, tales como las de audio asociadas, las señales de ensayo, la formación de ráfagas de color, los requisitos de entramado y los códigos de corrección de errores.

Estos factores, junto con los requisitos para un número de campo en exploración, de velocidades de transmisión de campo, de técnicas de modulación y de frecuencias portadoras, están predeterminados por acuerdos de telecomunicaciones que aseguran la calidad de la imagen en el extremo receptor.

Algunas de estas normas, que se describen a continuación, forman parte de las ya analizadas en los capítulos anteriores, siendo algunas de ellas complemento de las presentes:

4.1 H.223 Protocolo de Multiplexación para comunicación Multimedios a baja velocidad Binaria [UIT, Septiembre de 1995]

Esta recomendación especifica la estructura de las tramas, el formato de los campos y los procedimientos del protocolo de multiplexación de paquetes para comunicación multimedios de baja velocidad binaria y una unidad de control multipunto (MCU) o un adaptador de interfuncionamiento (IWA). En la recomendación H.245 se especifican los procedimientos de control necesarios para realizar este protocolo de multiplexación.

En esta recomendación, la comunicación entre diferentes capas de protocolo está modelada como un conjunto de primitivas abstractas, que representan un intercambio de información lógico. La descomposición de funcionalidad en (sub)capas, así como la descripción de las primitivas, no implica un método particular de realización. En particular, las capas pueden intercambiar el contenido de una unidad lógica (una SDU) en un modo "tren" en el que el intercambio de información puede iniciarse antes de que la capa que efectúa la transferencia posea la unidad completa.

Convenios de formato

Los convenios sobre numeración y correspondencia de campos de transmisión de bits utilizados en esta recomendación concuerdan con los utilizados en la recomendación V.42

Convenios de formato

En la figura 4.1.1 se ilustra el convenio de numeración básico utilizado en esta recomendación. Los bits agrupados en octetos. Los bits de un

octeto aparecen horizontalmente y están numerados de 1 a 8. Los distintos octetos se indican verticalmente y están numerados de 1 a n.

Orden de transmisión de los bits

Los octetos se transmiten por orden numérico ascendente, dentro de un octeto, el bit 1 es el primer bit que se transmite.

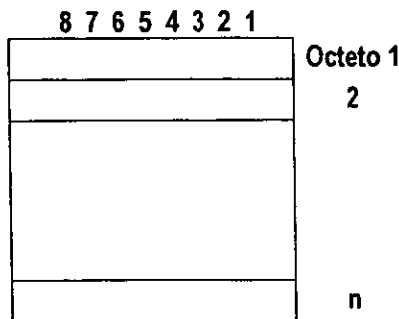


Figura 4.1.1 Convenio de numeración básico utilizado en esta recomendación

Convenio de correspondencia de campos

Cuando un campo está contenido dentro de un solo octeto, el bit de número más bajo del campo representa el valor de orden más bajo (o el bit menos significativo).

Cuando un campo abarca más de un octeto, el bit de número más alto del primer octeto representa el valor de orden más alto, el bit de número más bajo del último octeto representa un valor de orden más bajo.

El código de redundancia cíclica (CRC), constituye una excepción al convenio de correspondencia de campos precedente. En este caso, el

bit de número más bajo del primer octeto es el término de orden más alto del polinomio que representa el campo CRC, el bit de número más alto del último octeto en el término de orden más bajo del polinomio que representa el campo CRC.

Panorama General

Esta recomendación especifica un protocolo de multiplexación de paquete diseñado para el intercambio de uno o más trenes de información entre entidades de capa superior, tales como los protocolos de control y datos, también los códecs de audio y video.

En esta recomendación cada tren está representado por un canal lógico unidireccional que está identificado por un número de canal lógico (LNC) único, que es un entero comprendido entre 0 y 65535. El LNC 0 es un canal lógico permanente asignado al canal de control indicado en la recomendación H.245. Todos los demás canales lógicos son abiertos y cerrados dinámicamente por el transitor que utiliza los mensajes *abrir canal lógico* y *cerrar canal lógico* de la recomendación H.245.

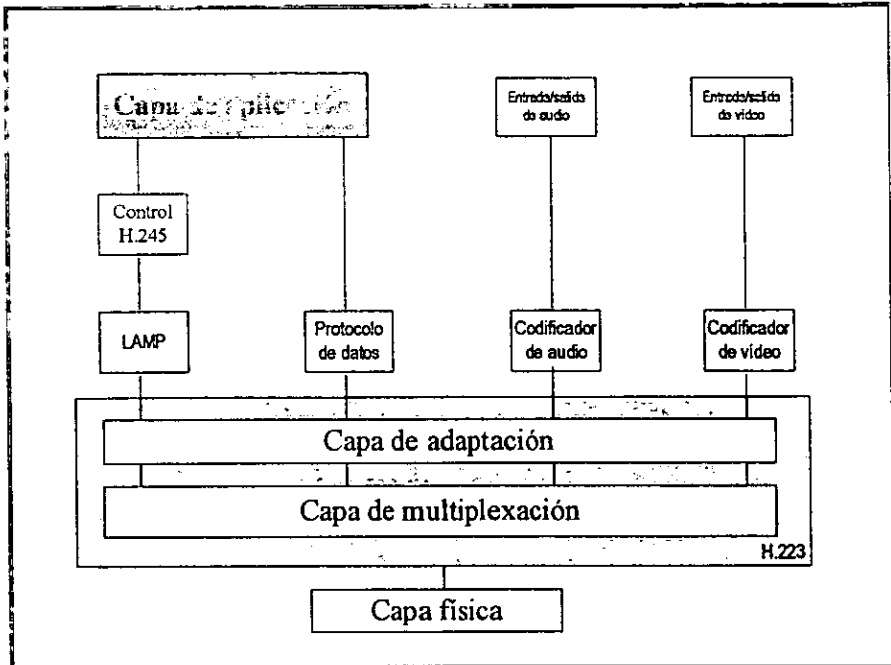


Figura 4.1.2 Pila de protocolos de la recomendación H.223

Todos los atributos necesarios del canal lógico se especifican en el mensaje abrir canal lógico. En cuanto a las aplicaciones que requieren un canal inverso, en la recomendación H.245 se define también un procedimiento para abrir canales lógicos bidireccionales.

En la figura 4.1.2 se indica la estructura general del multiplexor. Éste consta de dos capas distintas: una capa de multiplexación (MUX) y una capa de adaptación (AL).

Capa de multiplexación.

La capa MUX se encarga de transferir la información recibida de la AL en extremo distante, utilizando los servicios de una capa física

adyacente. La capa de multiplexación intercambia información con la AL en unidades lógicas llamadas MUX-SDU. Las MUX-SDU contienen siempre un número entero de octetos que pertenecen a un solo canal lógico. Las MUX-SDU representan generalmente bloques de información cuyo inicio y terminación marca la localización de campos que necesitan ser interpretados en el receptor.

Capa de adaptación.

La unidad de información intercambiada entre la AL y los usuarios de la AL de capa superior es una AL-SDU. El método de aplicación de trenes de información de capas superiores en las AL-SDU está fuera del alcance de esta recomendación y se especifica en la recomendación sobre sistemas que utiliza la recomendación H.223.

Las AL-SDU contienen un número entero de octetos. La AL adapta las AL-SDU a la capa MUX añadiendo, cuando sea conveniente, octetos adicionales para fines como la detección de errores, la numeración de secuencias y la retransmisión. La unidad de información lógica intercambiada entre entidades AL pares se llama AL-PDU. Una AL-PDU transporta exactamente la misma información que una MUX-SDU.

En esta recomendación se especifican tres tipos diferentes de AL, denominadas AL1, AL2 y AL3:

- La AL1 está diseñada básicamente para la transferencia de datos o información de control. Dado que AL1 no proporciona ningún control de errores, toda la protección contra errores necesaria debería ser proporcionada por el usuario AL1.
- La AL2 está diseñada básicamente para la transferencia de señales de audio digitales.

La AL2 recibe tramas, posiblemente de longitud variable, de su capa superior (por ejemplo un codificador de audio) en las AL-SDU, y pasa éstas a la capa MUX en las MUX-SDU, tras añadir un octeto para un CRC de 8 bits, y optativamente, 1 octeto para la numeración de secuencias.

- La AL3 está diseñada básicamente para la transferencia de señales de video digitales.

La AL3 recibe tramas de longitud variable de su capa superior (por ejemplo, un codificador de vídeo) en las AL-SDU, y pasa éstas a la capa MUX en las MUX-SDU tras añadir 2 octetos para un CRC de 16 bits, y optativamente 1 ó 2 octetos de control. La AL3 incluye un protocolo de retransmisión diseñado para vídeo.

Estructura de la capa de multiplexación.

La capa de multiplexación proporciona las capacidades para transferir las MUX-SSDU de la AL emisora a la AL receptora utilizando los servicios de una capa física que está por debajo. Las MUX-SDU deberán contener siempre un número entero de octetos. La capa MUX deberá transferir todas las MUX-SDU que pertenecen a un canal lógico dado, en el mismo orden en que son recibidas de la AL que está por encima.

Primitivas intercambiadas entre la capa de multiplexación y la AL.

La capa de multiplexación puede interconectar una o más AL. La información intercambiada entre la capa MUX y cada AL incluye las siguientes primitivas:

- Petición MUX-DATOS (MUX-SDU): Esta primitiva es emitida a la capa de multiplexación por una entidad emisora de AL para solicitar la transferencia de una MUX-SDU a la entidad receptora correspondiente.

- Indicación MUX-DATOS (MUX-SDU): Esta primitiva es emitida por la capa de multiplexación a una entidad receptora de la AL para indicar la llegada de una MUX-SDU de la entidad emisora correspondiente.
- Petición MUX-ABORTO: Esta primitiva es emitida a la capa de multiplexación por una entidad emisora de la AL para señalar que ha de ser descartada una MUX-SDU parcialmente entregada.
- Indicación MUX-ABORTO. Esta primitiva es emitida por la capa de multiplexación a una entidad receptora de la AL para señalar que ha de ser descartada una MUX-SDU parcialmente entregada.

4.2 H.245 Protocolo de control para comunicaciones multimedios [UIT, Julio 1995]

La presente recomendación especifica la sintaxis y la semántica de los mensajes de información de terminal así como los procedimientos para utilizarlos en la negociación en la banda al comienzo de la comunicación o durante ésta. Los mensajes comprenden capacidades de transmisión y de recepción así como preferencia de modos desde el extremo de recepción, señalización del canal lógico y la indicación y control. Se especifican procedimientos de señalización con acuse de recibo para garantizar comunicaciones fiables audiovisuales y de datos.

Esta recomendación abarca una amplia gama de aplicaciones que incluyen servicios telefónicos, así como los de almacenamiento/recuperación, mensajería y distribución. Se aplica a sistemas multimedios que utilizan las multiplexaciones definidas en las

recomendaciones H.222.0 y H.223 aunque no se limita a tales sistemas únicamente. Estos sistemas comparten la misma sintaxis y semántica, por lo que son compatibles en términos de bits. Algunos de los procedimientos se aplican a la totalidad de los sistemas, en tanto que otros son más específicos de sistemas concretos.¹

Consideraciones Generales

Esta recomendación proporciona diversos servicios, algunos de los cuales se espera que sean aplicables a todas las terminales que las utilizan, siendo otros más específicos para algunas terminales en concreto. Se definen procedimientos para permitir el intercambio de capacidades audiovisuales y de datos; para solicitar la transmisión de un modo audiovisual y de datos determinado; para gestionar los canales lógicos utilizados para transportar la información audiovisual y de datos; para establecer cual terminal es la principal y cual la subordinada con fines de gestión de los canales lógicos bidireccionales; para transportar distintas señales de control e indicación; para controlar la velocidad de bits de los canales lógicos individuales y la totalidad del multiplexor, y para medir el retardo de ida y vuelta entre un par de terminales.

Aunque no todos los mensajes y procedimientos definidos en esta recomendación se aplicarán a la totalidad de las terminales, no se indican aquí tales restricciones, las cuales competen a las recomendaciones que hagan uso de esta recomendación.

Se ha definido esta recomendación de forma que sea independiente del mecanismo de transporte subyacente, aunque se ha previsto su

¹ Nota: No debería de existir ninguna confusión con el sistema de gestión T.120 transportado dentro del tren de datos y que comprende funcionalidades distintas de las que aquí se describen. El tren H.245 y el tren de datos T.120 son complementarios.

empleo con una capa de transporte fiable, es decir aquella que proporcione una entrega garantizada de datos correctos.

Intercambio de capacidad.

Los procedimientos de intercambio de capacidad tienen por finalidad asegurar que únicamente las señales multimedios que deben transmitirse son aquellas que la terminal de recepción puede recibir y manejar adecuadamente. Esto exige que las capacidades de cada terminal para recibir y decodificar sean conocidas por la otra terminal. No es necesario que la terminal comprenda o almacene todas las capacidades entrantes. Podrán ignorarse capacidades no comprendidas o no utilizadas, sin que esto implique la consideración de que existen errores.

Mediante la transmisión de este juego de capacidades, se pone en conocimiento de una terminal la capacidad total de la otra terminal para recibir y decodificar diversas señales.

Las capacidades de recepción describen la aptitud de la terminal para recibir y procesar los trenes de información entrantes. Los transmisores deberán limitar el contenido de la información transmitida al valor que el receptor haya indicado que es capaz de recibir. La ausencia de una capacidad de recepción indica que la terminal es incapaz de recibir información (se trata de un transmisor únicamente).

Las capacidades de transmisión describen la aptitud de la terminal para la transmisión de trenes de información. Estas capacidades permiten ofrecer a los receptores la posibilidad de elección entre distintos modos de funcionamiento, de modo que el receptor puede solicitar el modo en el que prefiere efectuar la recepción. La ausencia de una capacidad de transmisión indica que la terminal no ofrecerá al

receptor la elección de modos preferidos (pero puede, sin embargo, transmitir alguna información compatible con la capacidad del receptor).

Estos juegos de capacidades permiten la transmisión simultánea de más de un tren de un tipo de medio determinado. Por ejemplo, una terminal puede declarar su aptitud para recibir (o enviar) dos flujos de vídeo H.262 independientes y dos trenes de audio G.722 independientes. Se han definido los mensajes de capacidad para permitir que una terminal indique que no posee capacidades fijas, sino que dependen de los demás modos que se estén utilizando simultáneamente. Por ejemplo, es posible indicar que se puede decodificar una señal de vídeo de elevada resolución cuando se utiliza un algoritmo de audio más simple o que pueden decodificarse cualquiera de las dos secuencias de vídeo de baja resolución o una sola de alta resolución. Es posible también expresar compromisos entre la capacidad de transmisión y la capacidad de recepción.

Pueden emitirse capacidades y mensajes de control no normalizados utilizando la estructura *"Non StandardParameter"*. Debe observarse que aunque el significado de los mensajes no normalizados lo definen organizaciones individuales, los equipos construidos por cualquier fabricante pueden señalar cualquier mensaje no normalizado si conocen su significado.

Las terminales pueden remitir juegos de capacidades en cualquier momento.

Petición de modos audiovisual y de datos.

Cuando se ha completado el protocolo de intercambio de capacidad, ambas terminales conocen la capacidad de la otra para poder

transmitir y recibir según se ha especificado en los descriptores de capacidad intercambiados. No es obligatorio que una terminal declare todas sus capacidades sino únicamente las que desea utilizar.

Una terminal puede indicar sus capacidades de transmisión. Una terminal que reciba capacidades de transmisión desde una terminal distante, puede pedir el mismo modo determinado de transmisión. Una terminal indicará que no desea que la terminal distante controle su modo de transmisión mediante el envío de capacidades de no transmisión.

Procedimientos de señalización de canal lógico.

Se define un protocolo de acuse de recibo para la apertura y el cierre de canales lógicos que transportan información audiovisual y de datos. La finalidad de estos procedimientos es garantizar que la terminal es capaz de recibir y decodificar los datos que se transmitirán por un canal lógico en el momento en que se transmita el primer dato por ella y para asegurar que la terminal de recepción está preparada para recibir y decodificar los datos que se transmitirán por el canal lógico antes del comienzo de la transmisión. El mensaje de apertura del canal lógico incluye una descripción de los datos que se transportarán, por ejemplo *H.262 MP@ML a 6 Mbits/s*. Los canales lógicos únicamente se abrirán cuando exista capacidad suficiente para recibir datos sobre todos los canales lógicos abiertos simultáneamente.

Una parte de este protocolo se aplica a la apertura de canales bidireccionales. Para evitar problemas de temporización, se define una terminal como terminal principal y la otra terminal como subordinada, únicamente la terminal principal puede iniciar la apertura de canales bidireccionales. Sin embargo, la terminal subordinada puede solicitar a la terminal principal que realice esa apertura.

Se ha definido un protocolo para establecer cual terminal será la principal y cual la secundaria. Sin embargo, los sistemas que hagan uso de esta recomendación deberán especificar otros modos de determinar cual terminal es la principal y cual la subordinada.

Instrucciones e indicaciones

Se proporcionan instrucciones e indicaciones con varios fines, por ejemplo bucles para mantenimiento, señales video/audio, activo/inactivo para informar al usuario, petición de actualización rápida para conmutación en aplicaciones multipunto. Ni las instrucciones ni las indicaciones requieren mensajes de respuesta de terminal distante. Las instrucciones implican una acción en la terminal distante, en tanto que las indicaciones se limitan a proporcionar información y no desencadenan ninguna acción.

Se define una instrucción para permitir que la terminal distante controle la velocidad de bits de los canales lógicos y la totalidad de los multiplexores. Esto tiene diversas finalidades: interfuncionamiento con las terminales que utilizan multiplexaciones en las que únicamente se dispone de un número finito de velocidades de bits, aplicaciones multipunto en las que deben adaptarse velocidades de señales de diferentes fuentes y control de flujo en redes congestionadas.

Mensajes

- ***Mensajes de determinación principal-subordinada.***

El protocolo utiliza este conjunto de mensajes para determinar cual terminal actúa como principal y cual como subordinada.

Acuse de recibo de determinación principal-subordinada

Se utiliza para confirmar si la terminal en cuestión es principal o subordinada, tal y como lo indica la decisión. Cuando la decisión es principal, la terminal receptora de este mensaje será la terminal principal y cuando la decisión es subordinada, se tratará de la terminal subordinada.

Rechazo de determinación principal-subordinada

Se utiliza para rechazar el mensaje "*MasterSlaveDetermination*". Cuando la causa es del tipo *IdenticalNumbers*, el rechazo se produjo porque ambos números aleatorios eran iguales.

- **Mensajes de capacidad de terminal**

Este juego de mensajes tiene por objeto el intercambio seguro de capacidades entre dos terminales.

Visión de conjunto

La terminal que transmite, asigna cada modo individual en el que puede funcionar en varias "*CapabilityTable*". Por ejemplo, audio G.723, audio G.728 y video CIF H.263, a las que se les asignarán números diferentes.

Estos números indicativos de capacidad se agrupan en estructuras "*AlternativeCapabilitySet*" e indican que la terminal puede funcionar exactamente en uno de los modos enumerados en el juego. Por ejemplo, un listado "*AlternativeCapabilitySet*" {G.711, G.723, G.72} indica que la terminal puede funcionar solamente en uno de los modos de audio.

Conjunto de capacidades de terminal.

Este mensaje contiene información sobre la capacidad de la terminal para transmitir y recibir. Indica también la versión en uso de esta recomendación. Se envía desde una CESE de salida a una CESE de entrada par.

Cuadro de capacidades

Un cuadro de capacidades es una lista numerada de capacidades. Una terminal deberá poseer todas las capacidades enumeradas en su cuadro de capacidades, aunque no es necesario que se tenga la aptitud de ejecutar simultáneamente más de una de ellas.

Descriptores de capacidad

Se utilizan los "*CapabilityDescriptors*" para expresar la capacidad que tiene una terminal para transmitir y recibir. Cada "*CapabilityDescriptor*" proporciona una declaración independiente acerca de las capacidades de la terminal.

Capacidades del multiplexor

"*MultiplexCapability*" expresa capacidades relacionadas con la multiplexación y la adaptación de la red. Toda terminal deberá enviar *MultiplexCapability* al menos en un "*TerminalCapabilitySet*", incluido el primero que se transmita.

Capacidades de vídeo.

Indica las capacidades de vídeo. La indicación de más de una capacidad en un único elemento "*VideoCapability*" no implica la capacidad de tratamiento simultáneo. La capacidad de tratamiento simultáneo puede expresarse mediante casos de "*VideoCapability*" en "*AlternativeCapabilitySets*" diferentes dentro de un único "*CapabilityDescriptor*".

Capacidades de audio.

Este parámetro expresa las capacidades de audio. La indicación de más de una capacidad dentro de un "AudioCapability" único no implica una capacidad de tratamiento simultánea. Tal capacidad puede expresarse mediante casos de "AudioCapabilitySets" diferentes en un único "CapabilityDescriptor".

Capacidades de aplicación de datos.

Este parámetro expresa capacidades de datos. La indicación de más de una capacidad dentro de un solo "DataApplicationCapability" no implica una capacidad de procesamiento simultáneo. Tal capacidad puede expresarse mediante casos de "DataApplicationCapability" en "AlternativeCapabilitySets" diferentes dentro de un único "CapabilityDescriptor".

- **Mensajes de señalización de canal lógico.**

Este juego de mensajes se utiliza para la señalización del canal lógico.

Apertura del canal lógico.

Se utiliza para intentar la apertura de una conexión de canal lógico entre una LCSE de salida y su LCSE de entrada par.

Acuse de recibo de apertura del canal lógico.

Se utiliza para confirmar la aceptación de la petición de conexión del canal lógico procedente de la LCSE par.

Rechazo de apertura del canal lógico

Se utiliza para rechazar la petición de conexión del canal lógico procedente de la LCSE par. "LogicalChannelNumber" indica el número de canal lógico del canal lógico en curso de rechazo.

Cierre de canal lógico.

Se utiliza por la LCSE de salida para cerrar una conexión de canal lógico entre dos LCSE pares. "LogicalChannelNumber" expresa el número de canal lógico que debe cerrarse.

Acuse de recibo de cierre de canal lógico.

Se utiliza para confirmar el cierre de una conexión de canal lógico

• **Mensajes de señalización de cuadro de multiplexación.**

Este conjunto de mensajes está destinado a la transmisión segura de inscripciones del cuadro de multiplexación H.223 del transmisor al receptor.

Envío de inscripciones al multiplexor.

Se utiliza para el envío de inscripciones de cuadro de multiplexación H.223 del transmisor al receptor.

Acuse de recibo de envío de inscripciones al multiplexor.

Se utiliza para confirmar la recepción de uno o más "MultiplexEntryDescriptor" desde una "MultiplexEntrySend" de la MTSE par.

Rechazo de envío de inscripciones al multiplexor.

Se utiliza para rechazar uno o más "MultiplexEntryDescriptors" desde una "MultiplexEntrySend" de la MTSE par.

- **Mensajes de petición de modo.**

Una terminal receptora utiliza este conjunto de mensajes para solicitar a la terminal de transmisión modos de transmisión determinados.

Petición de modo.

Se utiliza para solicitar modos concretos de transmisión de la terminal transmisora. Se trata de una lista, en orden de preferencia (el más preferible primero), de los modos en que la terminal desearía recibir. Se describe cada modo mediante un "*ModeDescription*".

Modo Video.

Se trata de una selección de "*VideoModes*".

H.261 VideoMode: Expresa la resolución de imagen solicitada (QCIF o CIF).

H.262 VideoMode: Expresa el perfil y nivel solicitados y los campos facultativos que, si están presentes, indican los valores solicitados de los parámetros proporcionados.

H.263 VideoMode: Expresa la resolución de imagen solicitada (SQCIF, QCIF, CIF, 4CIF y 16CIF).

Modo Audio.

Se trata de una selección de "*AudioModes*".

En la figura 4.2.1 se puede observar el significado exacto de los elementos de codificación de audio de la serie G. Hay cuatro posibilidades para el audio G.723, que permiten la elección de velocidades de bits (velocidad de bits reducida de 5.3 Kbps o la

velocidad de bits elevada de 6.3 Kbps) con el empleo o no de la supresión de los silencios.

Elemento de codificación ASN.1	Significado del elemento de codificación
g711Ala64k	Audio G.711 a 64 Kbps, ley A.
g711Ala56k	Audio G.711 a 56 Kbps, ley A., Truncado a 7 bits.
g711Ula64k	Audio G.711 a 64 Kbps, ley μ
g711Ula56k	Audio G.711 a 64 Kbps, ley μ . Truncado a 7 bits.
G722-64K	Audio G.722 7KHz a 64 Kbps.
G722-56K	Audio G.722 7KHz a 56 Kbps.
G722-48K	Audio G.722 7KHz a 48 Kbps.
G723	Audio G.723 a 5.3 o 6.4 Kbps.
G728	Audio G.728 a 16 Kbps.
G729	Audio G.729 a 8 Kbps.
G-dsvd	Audio G.dsvd (en elaboración).

Figura 4.2.1 Elementos de codificación de audio de la serie G

Modo datos.

Se trata de la selección de aplicaciones de datos.

- **Mensajes de retardo de ida y vuelta.**

La terminal utiliza este conjunto de mensajes para determinar el retardo de ida y de vuelta entre dos terminales en comunicación. Permite también a un usuario H.245 determinar si está activa la entidad de protocolo H.245 par.

- **Mensajes de bucle de mantenimiento.**

Una terminal utiliza este conjunto de mensajes para ejecutar funciones de bucle de mantenimiento.

Petición de bucle de mantenimiento.

Se envía para solicitar una clase determinada de puesta en bucle. Los tipos "***MediaLoop***" y "***LogicalChannelLoop***" solicitan la puesta en bucle de un único canal lógico indicado por "***LogicalChannelNumber***", en tanto que el tipo "***SystemLoop***" se refiere a la totalidad de los canales lógicos. La definición exacta de estos tipos es característica del sistema y queda fuera del alcance de esta recomendación.

Respuesta del bucle de mantenimiento.

Se utiliza para indicar como responderá la terminal a la petición de un bucle de mantenimiento.

Instrucción desconexión de bucle de mantenimiento.

Al recibir esta instrucción, la terminal desconectará todos los bucles y restaurará todos las trayectorias de vídeo, audio y datos a su condición normal.

• ***Instrucciones.***

Un mensaje de instrucción requiere una acción pero no una respuesta implícita.

Envío del conjunto de capacidades de terminal.

Esta instrucción ordena a la terminal del extremo distante que exprese sus capacidades de transmisión y recepción enviando uno o más "***TerminalCapabilitySets***" que contienen la información solicitada, como se indica a continuación. Esta instrucción puede enviarse en cualquier momento para expresar las capacidades de la terminal distante, por ejemplo tras una interrupción o cualquier otro motivo de incertidumbre. Sin embargo no deberán enviarse repetidamente esos mensajes sin una causa justificada.

Cifrado.

Esta instrucción se utiliza para intercambiar capacidades de cifrado y ordenar la transmisión de un vector de inicialización (IV).

Control de flujo.

Esta instrucción se utiliza para especificar el limite superior de la velocidad de bits en un solo canal lógico o en la totalidad del multiplexor. Una terminal puede enviar esta instrucción para limitar la velocidad de bits que envía la terminal distante. Toda terminal que reciba esta instrucción deberá cumplirla.

Finalización de sesión.

Esta instrucción indica el final de la sesión H.245. Tras la transmisión de "EndSessionCommand", la terminal no enviará ninguno de los mensajes definidos en esta recomendación. "Disconnect" indica que se eliminará la conexión.

• **Indicaciones.**

Una indicación contiene informaciones que no requieren acción ni respuesta.

Función no sustentada.

Se utiliza para devolver peticiones, respuestas e instrucciones no interpretadas al transmisor de las mismas.

Indicación de fluctuación.

Se utiliza para expresar el valor de la fluctuación de un canal lógico estimado por la terminal receptora. Es útil para la elección de la velocidad de bits y el control de la memoria tampón en los canales de video o para determinar una velocidad apropiada de transmisión

de la información de temporización, etc. El codificador de video tendrá entonces la posibilidad de emplear esta información para limitar la velocidad de bits de video o las fluctuaciones de la memoria tampón del codificador para la fluctuación existente. Si el decodificador hace uso de ésta opción, podrá corregir el funcionamiento de los diseños existentes de memoria tampón de decodificación de video, independientemente de la amplitud de la fluctuación recibida, así como corregir el funcionamiento con un retardo mínimo.

Indicación de asimetría H.223.

Se utiliza para indicar a la terminal del extremo distante el valor medio de la asimetría temporal entre dos canales lógicos.

Entrada de usuario.

Se utiliza para los mensajes de entrada de usuario.²

4.3 H.324 Transmisión en Línea de Señales No Telefónicas

[UIT, 1995]

La recomendación describe las terminales para comunicación multimedia a baja velocidad binaria, que utilizan los módems de la recomendación V.34 que funcionan en la Red Telefónica Pública Conmutada. Las terminales de la recomendación H.324 pueden transportar voz, datos y video en tiempo real o en cualquier combinación, incluida la telefonía.

² No se cifrará ningún dato transportado en H.245, incluidos los mensajes de entrada de usuario.

Las terminales de la recomendación H.324 pueden integrarse en computadoras personales o utilizarse en dispositivos autónomos, tales como videotelefonos. El soporte para cada tipo de medio (voz, datos, video) es facultativo, pero se requiere la posibilidad de utilizar un modo de funcionamiento común especificado, para que todas las terminales que sustentan este tipo de medio puedan interfuncionar. Además permite la utilización de más de un canal de cada tipo. Otras recomendaciones de la serie H.324 incluyen el múltiplex de la recomendación H.223, el control de la recomendación H.245, el códec de video de la recomendación H.263 y el códec de audio de la recomendación G.723.

La recomendación H.324 utiliza los procedimientos de señalización de canal lógico de la recomendación H.324, en la cual el contenido de cada canal lógico se describe cuando abre el canal. Se proporcionan procedimientos para indicar las capacidades del receptor y del transmisor, de manera que las transmisiones estén limitadas a aquellas en que los receptores puedan decodificar y que dichos receptores puedan solicitar a los transmisores un modo de funcionamiento determinado deseado. Dado que los procedimientos de la recomendación H.245 también están previstos para ser utilizados por los sistemas de la recomendación H.31 para redes de área local de anchura de banda no garantizada, el interfuncionamiento con estos sistemas debe ser directo.

Las terminales de la recomendación H.324 se pueden utilizar en configuraciones multipunto a través de las unidades de control multipunto y pueden interfuncionar con las terminales en redes inalámbricas.

Las terminales de la recomendación H.324 proporcionan video, audio o datos en tiempo real o cualquier combinación de estos, entre dos terminales telefónicas multimedios en una conexión de red de banda vocal de la *Red Telefónica Pública Conmutada*. La comunicación puede ser unidireccional o bidireccional. La comunicación multipunto es posible utilizando una *Unidad de Control Multipunto* (MCU) separada entre más de dos terminales de la recomendación H.324. Las MCU's y otros dispositivos que no son terminales no están vinculados por los requisitos de esta recomendación, pero deben cumplirlos cuando se considere práctico.

Las terminales telefónicas multimedios definidos en esta recomendación pueden integrarse en computadoras personales (PC) o estaciones de trabajo o ser unidades autónomas.

La presente recomendación trata también del interfuncionamiento con sistemas videotelefonicos en la RDSI (descrita en las recomendaciones de la serie H.320) y en redes de radiocomunicación móviles (ver proyecto de recomendación H.324).

En la figura 4.3.1 se muestra un sistema videotelefonico multimedios genérico de la H.324, que consiste en el equipo terminal, módem de *Red Telefónica Pública Conmutada* (RTPC), la *Unidad de Control Multipunto* (MCU) y otras entidades de funcionamiento del sistema. No es necesario que las realizaciones conforme a la recomendación H.324 tengan cada elemento funcional.

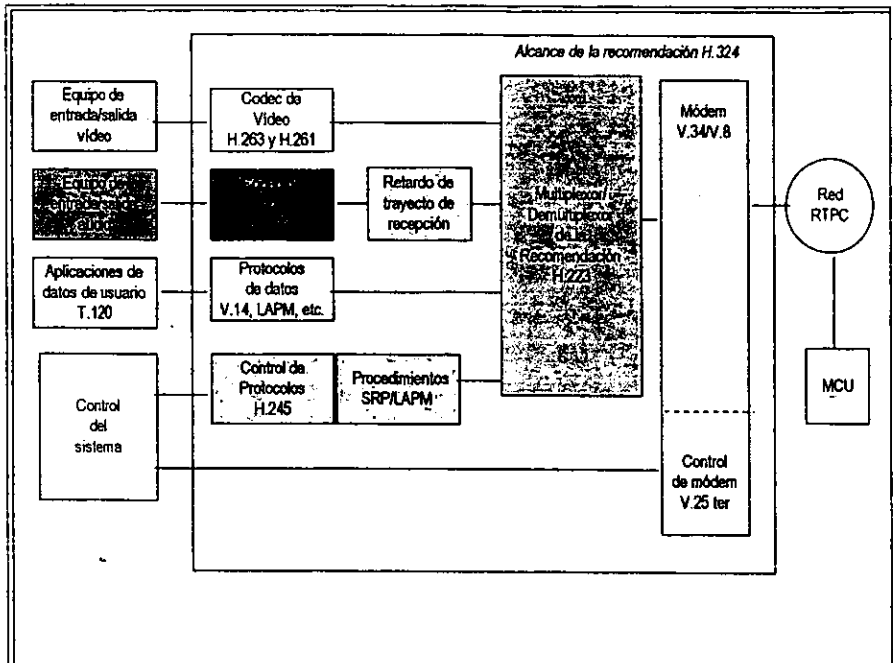


Figura 4.3.1 Diagrama de bloques para el sistema multimedia de la recomendación H.324.

Los siguientes elementos del sistema se tratan en otras recomendaciones o no están sujetos a normalización, por lo que no se definen en esta recomendación:

- Equipo de entrada/salida de vídeo, incluidos cámaras y monitores, su control y selección, procesamiento de vídeo para mejorar la compresión o proporcionar funciones de pantalla dividida.
- Equipo de entrada/salida de audio, incluidos micrófonos y altavoz, instrumento telefónico o equivalente, dispositivos de audio adjuntos que proporcionan detección de activación de voz, mezcladores de múltiples micrófonos, compensación del eco acústico.

- Equipos de aplicación de datos, tales como computadoras, protocolos de aplicaciones de datos no normalizadas, auxiliares visuales telemáticos, por ejemplo tableros electrónicos, etc.
- Interfase de la Red TPC que admite funciones adecuadas de señalización, tono de niveles de tensión, de acuerdo con las normas internacionales.
- Control interfase de usuario y funcionamiento del sistema de usuario.

El alcance de la recomendación H.324 esta indicado por los elementos que figuran dentro de la línea de trazo interrumpido de la figura 4.1, que incluye:

- El códec de video (recomendación H.263 o H.261) lleva a cabo codificación de la reducción de redundancia para trenes de video.
- El códec de audio (recomendación G.723) codifica la señal de audio desde el micrófono para transmisión y decodifica el código de audio de salida al hablante. El retardo facultativo del trayecto de audio, para mantener la sincronización de audio y video.
- El protocolo de datos admite aplicaciones telemáticas tales como: tableros electrónicos, transferencia de imágenes fijas, intercambio de ficheros, accesos a bases de datos, conferencias audiográficas, etc. Las aplicaciones de datos normalizados incluyen la recomendación T.120 para conferencia audiográfica en tiempo real, la transferencia de ficheros de imágenes fijas punto a punto simple de la recomendación T.84, la transferencia de ficheros punto a punto simple de la recomendación T.434 y el transporte de datos de

usuario utilizando la recomendación V.14 con memoria tampón o LAPM/V.42. También se pueden utilizar otras aplicaciones y protocolos, mediante negociación con el protocolo de la recomendación H.245.

- El protocolo de control (Recomendación H.245) proporciona señalización extremo a extremo para asegurar un funcionamiento correcto del terminal H.324, y señala todas las demás funciones del sistema extremo a extremo, incluida la inversión al modo de telefonía analógica de voz solamente. Proporciona el intercambio de capacidades, la señalización de instrucciones e indicaciones y los mensajes para abrir y describir completamente el contenido de los canales lógicos.
- El protocolo múltiplex (Recomendación H.223) multiplexa el vídeo, el audio, los datos y los trenes de control transmitidos en un tren binario único, y demultiplexa un tren de bits recibido en varios trenes multimedios. Además, realiza la alineación de trama lógica, numeración de secuencias, detección de errores y corrección de errores mediante retransmisión, según convenga a cada tipo de medio.
- El módem (recomendación V.34) convierte el tren de bits multiplexado síncrono de la recomendación H.223 en una señal analógica que puede transmitir por la RTPC, y convierte la señal lógica recibida en un tren de bits síncrono que se envía a la unidad de protocolo múltiplex/demúltiplex. La recomendación V.25ter se utiliza para proporcionar control/detección de la interfase de módem/red cuando el módem con señalización de red y los elementos funcionales de las recomendaciones V.8 y V.8Bis constituyen un elemento físico separado.

No es necesario que las aplicaciones conforme a la recomendación H.324 tengan cada elemento funcional, excepto en el caso del módem de la recomendación V.34, el múltiplex de la recomendación H.223 y el protocolo de control del sistema de la recomendación H.245, que serán soportadas por todas las terminales H.324.

Las terminales H.324 que ofrecen comunicación de audio admitirán el códec de audio de la Recomendación G.723. Las terminales H.324 que ofrecen comunicación de vídeo admitirán los códecs para vídeo de las recomendaciones H.263 y H.261. Las terminales de las recomendación H.324 que ofrecen conferencia audiográfica en tiempo real admitirán la serie de protocolos de la recomendación T.120. Además facultativamente se pueden utilizar otros códecs de vídeo y audio además de otros protocolos de datos mediante negociación por el canal de control de la recomendación H. 245.

Si se utiliza un módem externo a la terminal H.324, el control de la terminal/módem será conforme a la recomendación V.25^{ter}.

La presencia de facilidades facultativas se señala por el canal de control de la recomendación H.245. Si ambos extremos admiten una facilidad facultativa y eligen usarla, se negocia la apertura de un proyecto para transportar dichos trenes de información con arreglo a los procedimientos de la recomendación H.245.³

³ Nota: Esta recomendación no especifica una aplicación particular. Cualquier realización que proporcione la funcionalidad requerida y que se ajuste al formato del tren de bits descrito en esta recomendación, se considera conforme.

Trenes de Información

Los trenes de información multimedia se clasifican en video, audio, datos y control, de la siguiente manera:

- Los trenes de video constituyen tráfico continuo que transporta imágenes de color en movimiento. Cuando se utiliza la velocidad binaria disponible para los trenes de video puede variar según las necesidades de los canales de audio y datos.
- Los trenes de audio son en tiempo real pero pueden retardarse facultativamente en el trayecto de procesamiento del receptor para mantener la sincronización con los trenes de video. Con el fin de reducir la velocidad binaria media de los trenes de audio, se puede proporcionar la activación por la voz.
- Los trenes de datos pueden representar imágenes fijas, facsimil, documentos, archivos de computadora, datos de aplicaciones por computadora, datos de usuario no definidos y otros trenes de datos.
- Los trenes de control pasan instrucciones e indicaciones de control entre elementos funcionales distantes. El control de terminal a módem es conforme a la recomendación V.25_{ter} para terminales que utilizan módems externos conectados por una interfase física separada. El control de terminal es conforme a la Recomendación H.245.

Módem

Los módems utilizados para los terminales H.324 funcionarán en modo dúplex y síncrono, conformándose con las recomendaciones UIT-T V.35 y UIT-T V.8bis, hasta que el UIT-T apruebe la recomendación V.8 será sustituida por la Recomendación V.8 bis. La salida del multiplexor de la recomendación V.34. Cuando se utiliza un módem V.34 externo

no integrado, el control entre el módem y la terminal se realizará mediante la Recomendación UIT-T V.25^{ter}. En tales casos, la interfase física es específica de la aplicación.

Multiplexor

Los canales lógicos de video, audio, datos o información de control pueden ser transmitidos de acuerdo con los procedimientos de la recomendación H.245. Los canales lógicos son unidireccionales e independientes en cada sentido de transmisión. Se puede transmitir cualquier número de canales lógicos de cada tipo de medio, excepto en el caso del canal de control de la recomendación H.245, que sólo habrá uno. El método de multiplexación utilizado para transmitir estos canales lógicos será conforme a la Recomendación H.223

El multiplexor de la recomendación H.223 consiste en una capa de multiplexación, que mezcla los diversos canales lógicos en un tren de bits, y una capa de adaptación que trata el control de errores y la numeración de secuencias, según proceda para cada tren de información. La capa múltiplex transfiere información del canal lógico en paquetes denominados MUX-PDU, delimitados por banderas HDLC y utilizando la inserción de bits cero HDLC para transparencia. Cada MUX-PDU contiene una cabecera de un octeto seguida de un número variable de octetos de campos de información. El octeto de la cabecera incluye un código múltiplex que especifica, mediante una referencia a una tabla múltiplex, la correspondencia de los octetos del campo de información con los diversos canales lógicos. Cada MUX-PDU puede contener un código múltiplex diferente y por lo tanto una mezcla diferente de canales lógicos.⁴

⁴ Nota Las terminales de la recomendación H.324 indicarán su capacidad de la recomendación H.223 mediante el mensaje *Capacidad H.223* de la recomendación H.245.

Canal de Control

El canal de control transporta mensajes de control externo que determinan el funcionamiento del sistema H.324, incluidos el cambio de capacidades, la apertura y cierre de canales lógicos, las peticiones de preferencia de modo, la transmisión de entradas de la tabla múltiplex, los mensajes de control de flujo y las instrucciones e indicaciones generales.

Habrà exactamente un canal de control en cada sentido dentro de la recomendación H.324, que utilizarà los mensajes y procedimientos de la recomendación H.245. El canal de control será transportado por el canal lógico 0. El canal de control se considerará permanentemente abierto desde el establecimiento de la comunicación digital hasta la terminación de la misma, los procedimientos normales para la apertura y cierre de los canales lógicos no serán aplicables al canal de control.⁵

Canales de Video

Todas las terminales H.324 que ofrecen comunicación de video soportarán códecs de video de las recomendaciones H.263 y H.261, excepto que los adaptadores para interfuncionamiento de la recomendación H.320 (que no son terminales) no tendrán que soportar la recomendación H.263. Los códecs de las recomendaciones H.261 y H.263 se utilizarán sin corrección de errores BCH y sin alineación de trama de corrección de errores. Existen cinco formatos de imagen normalizados 16CIF, 4CIF, CIF, QCIF y SQCIF.

CIF y QCIF se definen en la recomendación H. 261. Para el algoritmo de la recomendación H.263, SQCIF, 4CIF y 16CIF se definen en la

⁵ Nota: Todos los mensajes del canal de control se envían por un protocolo de capa de enlace que acusa la recepción correcta. Este acuse es distinto de los mensajes de respuesta, que transportan contenido además de la recepción correcta del mensaje.

misma recomendación. Para el algoritmo de la recomendación H.261, SQCIF es de cualquier tamaño de imagen activa menor que QCIF, completado por un borde negro y codificado en el formato QCIF, para todos estos formatos, el formato de pixel es el mismo que el formato CIF.⁶

La figura 4.3.2 se muestra un cuadro con los formatos de imagen requeridos y facultativos para las terminales H.324 que admiten vídeo.

Formato de Imagen	Píxeles de Luminancia	Codificado r	Decodificador		
			H.263	H.261	H.263
SQCIF	128 para la recomendación H.263 (3)	Facultativo (3)	Requerido (1,2)	Facultativo (3)	Requerido (1)
QCIF	176 x 144	Requerido	Requerido (1,2)	Requerido	Requerido (1)
CIF	352 x 288	Facultativo	Facultativo	Facultativo	Facultativo
4CIF	704 x 576	No definido	Facultativo	No definido	Facultativo
16CIF	1,408 x 1,152	No definido	Facultativo	No definido	Facultativo

1. Facultativo para los adaptadores para el interfuncionamiento de la recomendación H.320.
2. Obligatorio para codificar uno de los formatos de imagen QCIF y SQCIF, facultativo para codificar ambos formatos.
3. SQCIF de la Recomendación H.261 es cualquier tamaño activo menor que QCIF, completado por un borde negro y codificado en el formato QCIF.

Figura 4.3.2. Cuadro con los formatos de imagen requeridos y facultativos para las terminales H.324 que admiten vídeo.

Todos los decodificadores de vídeo deberán poder procesar los trenes de bits de vídeo de la velocidad binaria máxima que pueda recibirse mediante la realización de multiplexación de la recomendación H.223.

⁶ El formato de la *imagen* resultante para QCIF de la recomendación H.263 es diferente de todos los demás formatos.

Los formatos de imagen, el número mínimo de imágenes saltadas y las opciones de algoritmos que puede aceptar el decodificador, se determinan durante el intercambio de capacidades utilizando la recomendación H.245. A continuación el codificador es libre de transmitir cualquier cosa que concuerde con la capacidad del codificador. Los decodificadores que indican capacidad para una opción de algoritmo concreto podrán aceptar también trenes de bits de video que no utilicen esa opción.

También se pueden utilizar otros códecs para video y otros formatos de imagen mediante negociación de la recomendación H.245. Puede transmitirse más de un canal de video negociando mediante el canal de control de la recomendación H.245.

Canales de Audio

Todas las terminales H.324 que ofrecen comunicación de audio admitirán las velocidades alta y baja del códec de audio de la recomendación G.723. Los receptores G.723 podrán aceptar tramas de silencio. El transmisor elige la velocidad que se ha de utilizar y se señala al receptor dentro de banda en el canal de audio, como parte de la sintaxis de cada trama de audio. Los transmisores pueden cambiar las velocidades de la recomendación G.723 trama por trama, sobre la base de la velocidad binaria, la calidad de audio u otras referencias. Los receptores pueden señalar mediante los procedimientos H.245 una preferencia por una velocidad o modo particular.

También se pueden utilizar códecs de audio alternativo mediante la negociación con H.245. Los codificadores pueden omitir el envío de señales de audio durante periodos silenciosos después de enviar una

sola trama de silencio, o pueden enviar tramas de relleno con fondo de silencio si la recomendación vigente sobre códec audio especifica estas técnicas.⁷

Canales de Datos

Todos los canales de datos son facultativos, las opciones normalizadas para aplicaciones de datos comprenden:

- Las recomendaciones de la serie T para teleconferencia audiográfica punto a punto y multipunto, incluidos el acceso a la base de datos, la transferencia y anotación de imágenes fijas, la compartición de aplicaciones, la transferencia de archivos en tiempo real, etc.
- La transferencia de imágenes fijas punto a punto según la recomendación T.84 a través de los límites de aplicaciones.
- La transferencia de archivos telemáticos punto a punto según la recomendación T.434 a través de límites de aplicaciones
- La recomendación H.224 para el control en tiempo real de aplicaciones simplex, incluido el control de cámaras en el extremo distante de la recomendación H.281.
- La capa de enlace de red, según la norma ISO/CEI TR 9577.
- Datos de usuario no especificados procedentes de puertos de datos externos.

⁷ Nota. Cada canal de audio es independiente. La agrupación de canales de audio en pares estereofónicos u otros grupos sincronizados queda pendiente para un estudio posterior.

Estas aplicaciones de datos pueden residir en una computadora externa u otro dispositivo especializado vinculado a la terminal H.324 mediante una interfase V.24 o equivalente (según la aplicación), o pueden integrarse en la propia terminal H.324. Cada aplicación de datos utiliza un protocolo de datos subyacente para transporte de la capa de enlace. Para cada aplicación de datos soportada por la terminal H.324, la presente recomendación requiere el soporte de un protocolo de datos subyacente particular, para garantizar el interfuncionamiento de aplicaciones de datos.⁸

Todas las terminales H.324 que ofrecen conferencia audiográfica en tiempo real deben admitir la serie de protocolos de la recomendación T.120.

Interfuncionamiento con otras Terminales. (Terminales de conversación solamente).

Los videoteléfonos H.324 admitirán el interfuncionamiento con teléfonos analógicos de conversación solamente.

Terminales telefónicos multimedios H.320 en la RDSI

Se puede proporcionar interfuncionamiento con terminales multimedios a través de la RDSI (H.320):

- Utilizando en la RDSI un adaptador de interfuncionamiento, o
- Utilizando terminales en modo doble (RDSI y RTGC) en la RDSI

Se coloca un adaptador para interfuncionamiento H.324/H.320 en la interfase entre señales de la RDSI y de la RTGC, que transcodifica los multiplexores H.223 y H.221, y el contenido de los canales lógicos de control, audio y datos entre los protocolos H.324 y H.320.

⁸ El canal de control de la recomendación H.245 no se considera un canal de datos.

Para facilitar la comunicación entre terminales H.324 y H.320 a través de adaptadores para interfuncionamiento, las terminales H.324 que admiten vídeo soportarán el códec de vídeo H.261 en el formato de imagen QCIF de modo que pueda evitarse el retardo adicional de transcodificación de vídeo. Cuando se utiliza este modo, los adaptadores para interfuncionamiento insertarán y suprimirán la corrección de errores BCH H.261 y alineación de trama con corrección de errores, según proceda, para cada tipo de terminal. Los terminales H.324 responderán a la instrucción de control de flujo para cada tipo de terminal.

Las terminales en modo doble (H.320 y H.324) en la RDSI enviarán señales RTGC H.324, utilizando un "módem virtual", que genera y recibe una señal analógica V.34 codificada como un tren de bits de audio G.711 por la RDSI.

4.3.1 Terminales Telefónicas multimedios en Redes Radioeléctricas Móviles. (Recomendación H.324/M)

Se prevé que las terminales telefónicas multimedios también se utilizarán en redes radioeléctricas móviles. Estas terminales cubrirán los requisitos de la Recomendación H.324/M. La adaptación de la velocidad entre terminales inalámbricas y terminales RTGC se puede efectuar mediante la instrucción de control de flujo H.245. El funcionamiento inalámbrico será objeto de estudio posterior.

4.4 Hacia los Sistemas del Futuro. [Sanjuan, 1995]

Desde los años ochenta se han visto cambios dramáticos en las comunicaciones móviles, particularmente en el sector de los negocios.

Desde radiotaxis a radiolocalizadores (*paggers*). Se fue haciendo común oír el zumbador de un buscapersonas, usar un teléfono sin hilos en casa o ver a personas utilizando el teléfono en su auto mientras conducían. En los noventa, nos estamos embarcando en la segunda generación de sistemas móviles que usan comúnmente técnicas digitales. Actualmente existe una gran proliferación de tecnologías e iniciativas de nuevos servicios, (*GSM, DCT, TELEPUNTO, PCN*, etc.). A través de estas iniciativas se va definiendo una tendencia desde la situación actual de mercado multisegmentado en términos de normas existentes hacia un único estándar europeo.

Así, en telefonía celular, en la primera generación existían multitud de estándares, en Europa (*NMT, TACS, CNETZ*), para la segunda generación solo existirá el *GSM*. También se va borrando la barrera tradicional entre entorno público de exteriores (celular) y entorno privado de interiores (negocios y teléfonos inalámbricos). Aunque inicialmente diseñados para el sector de los negocios, estos sistemas se van moviendo cada vez más hacia el usuario residencial, es decir al mercado de consumo. La iniciativa *PCN*, que emplea el estándar *DCS1800 8* (variante del *GSM* y promovido por el *ETSI*), como soporte de telecomunicaciones, es un paso en esa dirección.

Al inicio de la próxima década estos sistemas de la segunda generación habrán estimulado el mercado masivo de las comunicaciones móviles. Para ese momento, los sistemas habrán sido explotados hasta el máximo de sus capacidades y no podrán satisfacer los requerimientos de los usuarios en demanda y calidad. Las expectativas son de poder llegar a una penetración de hasta un 40% de la población activa para el año 2005, esto representa una población de unos 100 millones de usuarios o más. Será el momento para un sistema de la tercera generación, un sistema universal, multifunción,

que empleará técnicas digitales desarrolladas a lo largo de los noventa y tecnologías del año 2000. Este sistema se denomina en Europa *Universal Mobile Telecommunication System* (UMTS), y está siendo desarrollado en RACE y el estándar en ETSI, grupo SMG5 y tiene su paralelo en el *Future Public Land Mobile Telecommunication System* (FPMTS) que está estandarizado en el UIT, antes CCITT.

UMTS [Sarjuan, 1995]

El sistema *UMTS* está siendo concebido para proveer de manera universal (esto es, disponible en todos los lugares y en todo momento), una amplia variedad de servicios al usuario mediante una familia de terminales. La universalidad de un servicio dependerá tanto de condicionantes técnicas como económicas. Este término también se refiere a que existirá un estándar único para *UMTS* (por ejemplo en la interfase aire).

Dentro del *UMTS* se ha acuñado un término cuyo significado es la clave para apreciar la gran diferencia ente los sistemas de la segunda generación y el sistema *UMTS*. El término es *comunicador personal*, y se refiere a una terminal que, además de ser una terminal de voz de tamaño adecuado para ser transportado cómodamente a cualquier parte, de bolsillo, y barato para que sea un verdadero producto de consumo, pueda ser usado indistintamente en entornos interiores y exteriores, en la ciudad o áreas rurales, en la casa, oficina o calles, añadiendo a todo esto la posibilidad de *soportar servicios de imágenes*; dependiendo del avance de las técnicas de codificación de las mismas.

Desde el punto de vista de los servicios, el *UMTS* soportará una mayor variedad de servicios con mejor calidad y mayores velocidades de

transmisión que los actualmente ofrecidos por sistemas como *GSM/DCS 1800* y *DCT*.

El *UMTS* soportará voz, datos, fotografía, imágenes en movimiento, gráficos, multimedia, etc. Adicionalmente se considera como esencial para el *UMTS* ofrecer mejor calidad en términos de cobertura y probabilidad de bloqueo. En el caso de servicio de voz, la percepción global de calidad y fiabilidad de la llamada deberá ser comparable a la ofrecida por la red fija actual.

Desde el punto de vista de disponibilidad, el *UMTS* deberá ofrecer cobertura universal. Esto significa la necesidad de ofrecer la capacidad de conectividad en una gran extensión geográfica. La universalidad también significa la disponibilidad de los servicios de *UMTS* en la multiplicidad de entornos en los que el *UMTS* deberá trabajar, rural, urbano, interiores, exteriores, incluyendo peatones y vehículos tanto públicos como privados.

Basados en los aspectos anteriores, se han fijado como objetivos para el *UMTS* los siguientes puntos:

- Soportar una alta densidad de usuarios.
- Soporte a una gran variedad de servicios tanto de alta como baja velocidad, incluyendo subconjuntos de los servicios de *RDSI* de banda ancha y banda angosta, tanto por conmutación de circuitos como de paquetes.
- Que ofrezca para el servicio de voz una calidad, costo y fiabilidad semejante a la de la red fija presente.
- Capaz de soportar los teléfonos de bolsillo que puedan ser usados en cualquier lugar de Europa o del mundo.

- Diseñado para soportar una gama de terminales desde el teléfono básico de bolsillo hasta terminales más sofisticadas que ofrezcan una mayor gama de servicios o una selección específica de éstos.
- Capaz de funcionar tanto en entornos diversos, incluyendo entornos no necesariamente *RDSI* de banda ancha, como en entornos distintos.
- Disponer del estándar para 1999.
- Definir una evolución viable desde los sistemas e infraestructuras actuales.
- Diseño suficientemente flexible que permita el desarrollo de futuros sistemas a comienzos del siglo XXI.

Arquitectura del Sistema.

El sistema *UMTS* se implantará a través de un conjunto de subredes, tanto públicas como privadas, cuya combinación permita la provisión de servicio en cualquiera de los entornos ya comentados con anterioridad. Cada uno de los escenarios impone diferentes condicionantes económicas y técnicas, por lo tanto requieren distintos tipos de soluciones.

Las áreas en las que el *UMTS* deberá trabajar deben ser cuidadosamente caracterizadas particularmente en términos de tráfico esperado y en estimación de mercado. Las áreas en las que se están considerando son: puntos de alta densidad (*busy spots*), oficinas, residenciales, autopistas, áreas metropolitanas, áreas urbanas, y áreas rurales, en la figura 4.4.1 aparecen los valores de demanda con los que se está trabajando el programa *RACE*.

Adicionalmente, los puntos de acceso radio *UMTS* deberán ser capaces de conectarse o integrarse con las redes fijas existentes a lo largo de Europa y aunque principalmente ideados para su conectividad a redes *RDSI* de banda ancha, deberán ser capaces de operar otros sistemas autónomos para trabajar en otros tipos de entornos. Otro aspecto crítico para el sistema *UMTS* son las técnicas de manejo de las grandes cantidades de información, principalmente de abonados, y la inteligencia necesaria para controlar llamadas y la movilidad de los usuarios.

<i>Tipos de Área</i>	<i>Estimación de Carga Esperada</i>
Punto de alta densidad usuarios (busy spots)	1,500 - 7,000 ETE por Km ²
Interiores negocios	10,000 - 20,000 ETE por Km ²
Interiores residenciales	3 ETE por cada 64 viviendas
Autopistas	1 ETE por Km
Urbana	500 - 1,000 ETE por Km ²
Áreas metropolitanas y rurales	A definir en estudios posteriores

ETE = Erlang telefónico equivalente

Figura 4.4.1 *Tabla de valores de demanda telefónica*

La figura 4.4.2 muestra un esquema básico de un sistema *UMTS*. Aquí aparecen entornos privados, *Customer Premises Networks*, (CPN), tanto residenciales (*DCPN*) como de negocios (*BCPN*) así como redes privadas móviles (*MCPN*). Aparece por supuesto, el acceso público a *UMTS*. En esta misma figura aparecen marcadas las celdas que están siendo definidas por *UMTS*, microcélulas para autopistas y áreas densas en las ciudades, picocélulas para entornos de interiores y macrocélulas para las áreas grandes.

Las macrocélulas darán servicios a las áreas públicas metropolitanas y rurales pero también serán usadas en el interior de las ciudades sirviendo como células paraguas, asegurando la cobertura y el servicio móvil desplazándose de unas zonas atendidas por microcélulas a otras.

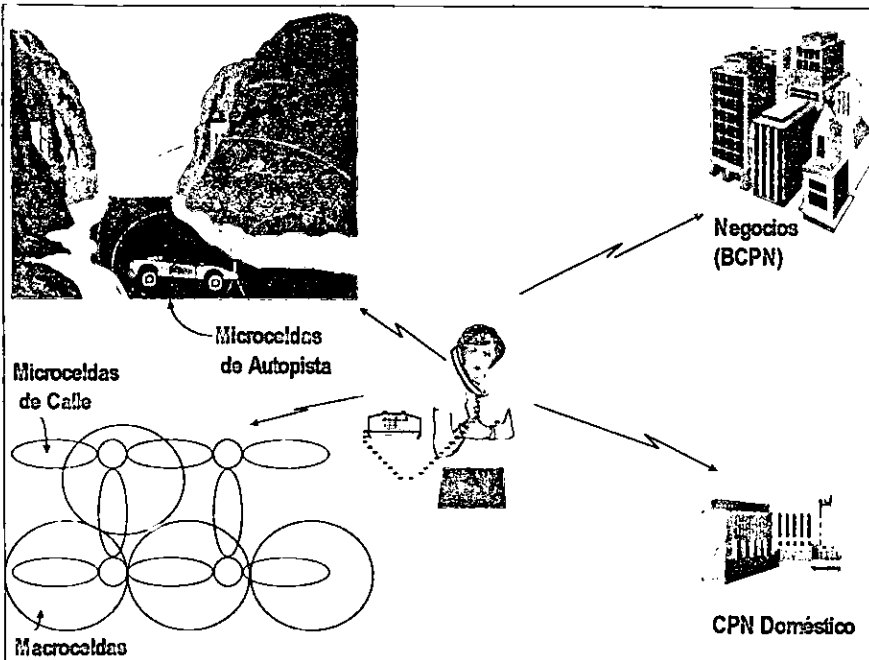


Figura 4.4.2 Esquema básico de un sistema UMTS

El concepto de microcélulas y picocélula, no contemplado de forma integrada por los sistemas de la segunda generación, permitirá al UMTS soportar mayor demanda, el concepto de células paraguas, nuevo para sistemas celulares, permitirá mantener los costos de la infraestructura en niveles adecuados. El requisito de cobertura universal obligaría en caso de no existir células paraguas, a un

despliegue generalizado de microcélulas, incluso en áreas en las que la demanda esperada no las justificara. Esta es una diferencia importante con respecto a GSM donde no existe tal tipo de célula.

Para implantar tal sistema se ha definido una arquitectura de red así como el conjunto de funciones que tal red, y sus elementos, deben realizar. A continuación se describen algunos de los aspectos más importantes.

Arquitectura Funcional

En la figura 4.4.3. aparece un diagrama de la arquitectura funcional de la red. En dicho diagrama aparecen definidos los entornos en los que el UMTS deberá trabajar así como los elementos de red que cada uno de ellos necesitará para poder dar el servicio esperado. Los elementos son tanto los propios de una red fija con la que el UMTS se conectará, así como los específicos del UMTS.

Los elementos son *TE (Terminal Equipment)*, *NT2/NT1 (Networks Terminations)*, *LT (Line Termination)* en la central de conmutación y *SW (Switch)* función de conmutación en la misma central, *TEM (Mobile Terminal Equipment)*, *BS (Base Station)*, *MCN (Mobile Control Node)* e *ISN (Information Storage Node)*.

Como se puede observar en la figura 4.4.3, no todos los elementos deben aparecer en todos los entornos. Así, en los entornos privados de negocios y dependiendo del tamaño del mismo, es posible que se requieran distintas estrategias de distribución en la inteligencia de que el sistema UMTS básicamente dependerá del tráfico interior y del tipo de servicios requeridos. Por ejemplo, es posible que se desee controlar todo el tráfico cursado dentro del propio entorno privado, esto requerirá

que en dicho entorno se incluya un nodo de control local de almacenamiento de la información, *ISN*.

Las funciones de conmutación deberán, en este caso, estar incluidas en el elemento *NT2*. En los entornos privados residenciales, la situación es semejante a la esperada en una red fija excepto que las interfases deben ahora soportar la interfase aire debido a la movilidad.

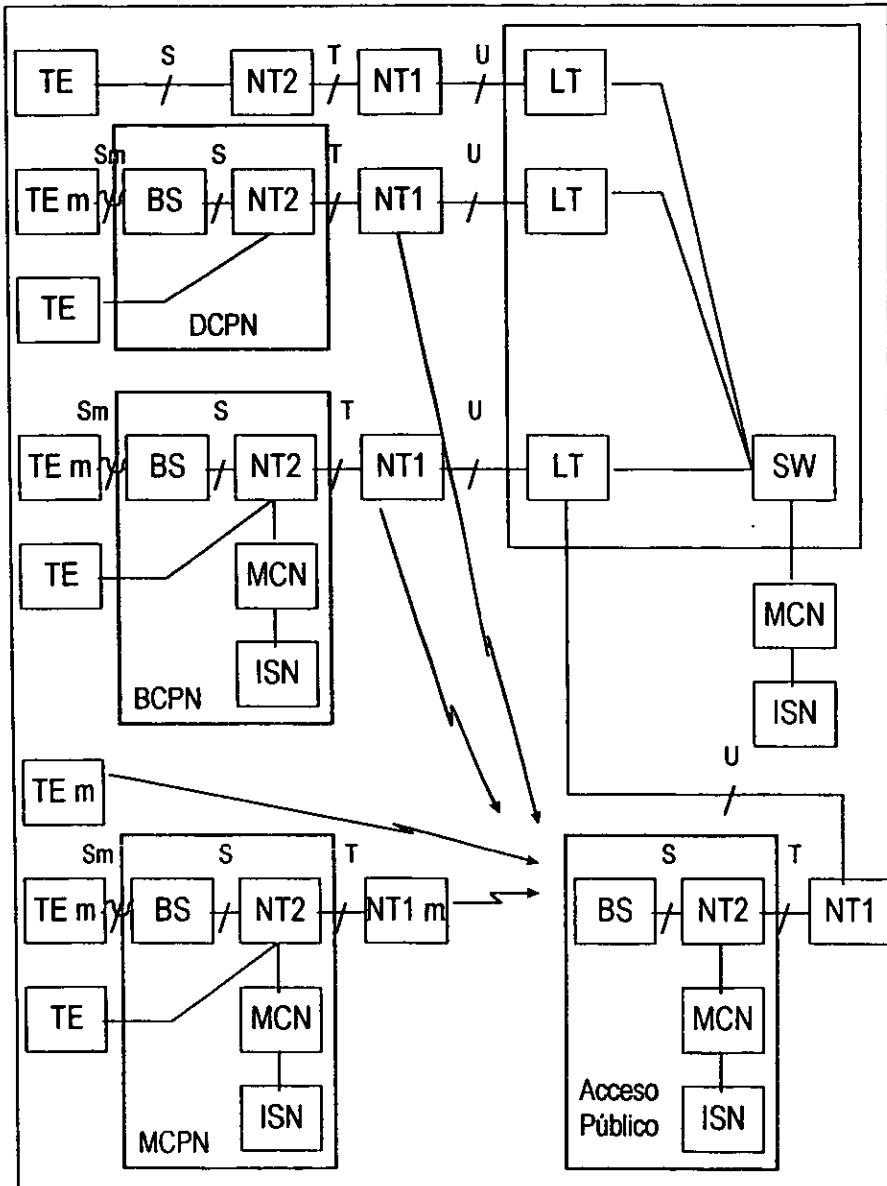


Figura 4.4.3 Diagrama de la arquitectura funcional de la red.

Es posible ver dentro de este entorno a negocios en una escala reducida en el que las funciones de movilidad están soportadas por la central de conmutación a la que dicho entorno está conectado. Esto permitirá disponer de equipos residenciales muy económicos. La parte pública de la red *UMTS* contiene estaciones base, con capacidad de conmutación y gestión de movilidad para los usuarios que pueden desplazarse de entorno a entorno dentro del área pública (*roaming*), el dominio público mantiene información de localización de los usuarios que pueden moverse entre redes y puede ofrecer servicios de señalización a usuarios de entornos privados residenciales. En este momento, aparece otro tipo de entorno en el *UMTS* que puede ser de interés en el futuro, las redes privadas móviles. Mediante estas redes el *UMTS* podrá dar servicio a usuarios de trenes, barcos, autobuses, etc. Las redes móviles contienen terminaciones de radio tanto en la terminación de red como en la interfase con las terminales. Las redes móviles plantean problemas importantes en lo relativo a traspaso, handovers múltiples y simultáneos, así como localizaciones múltiples. Todos estos aspectos influirán en los requisitos que deberá soportar la red *UMTS*.

Aplicaciones y Servicios

El sistema *UMTS* soportará voz, datos, fotografías fijas, gráficos, imagen en movimiento, etc. Estos incluirán todos los ofrecidos por *GSM* y *DCT*, así como servicios compatibles con un subconjunto de los ofrecidos por *RDSI* banda angosta y banda ancha. En la figura 4.4.4 se muestra un cuadro donde se observan las aplicaciones soportadas por *UMTS* con la estimación de la velocidad binaria y retardo máximo permitidos. En este conjunto de aplicaciones que el *UMTS* soportará junto con la universalidad en la provisión de las mismas; lo que hace del *UMTS* un sistema verdaderamente avanzado con respecto a los de la segunda generación.

Tipo de Servicio	Ejemplo de Aplicación	Velocidad Binaria Requerida	Retraso Permisible	Entorno de Empleo
Voz	Conversación (Telefonía)	8 - 32 Kbps	Tiempo real	Todos
Voz/Sonido	Grabación (Correo de voz)	8 - 32 Kbps	1 seg	Todos
Datos	Archivos informáticos	1.6 Mb por archivo	< 10 seg <120	Interiores Microcélulas Todos
Texto	Teletexto	40 Kb por página	<2 seg por página	Todos
Foto Fija	Imagen de Alta calidad, color	1.6 Mb por página	<10 seg por página <180 Seg por página	Interiores Microcélulas Todos
	Fax (G4)	500 Kb por página	<10 seg por página	Todos
Imagen en Movimiento	Videoteléfono	64 Kbps	Tiempo real	Interiores Microcélulas y algunas Macrocélulas interiores
	Videoconferencia de alta calidad	< 384 Kbps		
Portadores RDSI	2B + D	144 Kbps	Tiempo real deseable	Interiores Microcélulas Todos
	1B + D	80 Kbps		

Figura 4.4.4 Diagrama de la arquitectura funcional de la red.

WARC-92

En febrero de 1992 tuvo lugar en Torremolinos la reunión mundial de WARC-92 para definir la asignación de frecuencias a nivel mundial, como resultado del congreso se asignaron las bandas de frecuencias entre 1.885 a 2.025 GHz y de 2.110 a 2.220 GHz para el servicio móvil terrestre con la necesidad de compartir las bandas entre 1.98 a 2.010 GHz y de 2.170 a 2.200 GHz con el servicio móvil vía satélite. Esta asignación de frecuencias cubre perfectamente las necesidades de

banda de UMTS según un número de estudios previos realizados por diferentes institutos y dentro del propio RACE.

Glosario

AAL	Capa de adaptación de ATM (ATM adaptation layer)
AL	Capa de adaptación (Adaptation layer)
AL1,2,3	Capas de adaptación 1,2 y 3 de H.223 (H.223 adaptation layers 1, 2 and 3)
AL1-AL3	Capa de adaptación 1-3 (Adaptation layer 1-3)
ASN	Notación de sintaxis abstracta (Abstrac syntax notation)
ASN.1	Notación de sintaxis abstracta 1 (Abstrac syntax notation 1)
ATM	Modo de transferencia asincrónico (Asynchounous transfer mode)
B-LCSE	Entidad de señalización de canal lógico bidireccional (Bi-directional logical channel signalling entity)
CESE	Entidad de señalización de intercambio de capacidad (Capability exchange signallin entity)
CIF	Formato intermedio común (de una señal de vídeo) (Common intermediate Format)
CLCSE	Entidad de señalización de cierre de canal lógico (Close logical channel signalling entity)
CPCS	Subcapa de convergencia de parte común. (Common part convergence sublayer) de la capa 5 de adaptación de ATM
CRC	Código de redundancia ciclica (Cyclic redundance code)
DRTX	Retransmisión rechazada (decline retansmission)
DSM-CC	Medio de almacenamiento digital-instrucción y control (Digital storage media-command and control)
DTMF	Multifrecuencia de dos tonos (Dual tone multy-freccuency)
EI	Indicación de error (error indication)
GOB	Grupo de bloques (Group of blocks) de una imagen de vídeo.
GSTN	Red telefónica pública conmutada (General switched telephone network)
HDLC	Control de enlace de datos de alto nivel (High level data link control)

HEC	Control de error de encabezamiento (<i>header error control</i>)
HRD	Decodificador ficticio de referencia (<i>Hypotetical refernce decoder</i>)
IV	Vector de inicialización (<i>Initialisation vector</i>)
IWA	Adaptador de interfuncionamiento (<i>interworking adapter</i>)
LAPM	Procedimiento de acceso de enlace para módem (<i>link access procedure for modems</i>)
LAPM	Protocolo de acceso al enlace para módems (<i>link access protocol for modems</i>)
LCSE	Entidad de señalización de canal lógico (<i>Logical channel signalling entity</i>)
LNC	Número de canal lógico (<i>logical channel number</i>)
MC	Código múltiplex (<i>multiplex code</i>)
MCU	Unidad de control multipunto (<i>Multipoint control unit</i>).
MPI	Intervalo de imagen mínimo (<i>Minimun picture interval</i>)
MRSE	Entidad de señalización de petición de modo (<i>Mode request signalling entity</i>)
MSDSE	Entidad de señalización de determinación
MTSE	Entidad de señalización de cuadro múltiplex (<i>Multiplex table signalling entity</i>)
MUX	Multiplexor (<i>multiplex</i>)
OBCSE	Entidad de señalización de apertura de canal bidireccional (<i>Open bi-direccional channel signalling entity</i>)
PCR	Referencia de reloj de programa (<i>Program clock reference</i>)
PDU	Unidad de datos de protocolo (<i>Protocol data unit</i>)
PID	Identificador de paquete (<i>Packet identifier</i>)
PM	Marcador de paquete (<i>packet marker</i>)
PT	Tipo de paquete (<i>packet type</i>)
QCIF	Cuarta parte del CIF (<i>Quarter CIF</i>)
QoS	Calidad de servicio (<i>Quality service</i>)
RTDSE	Entidad de señalización de retardo de ida y vuelta (<i>Round trip delay signalling entity</i>)
SDL	Lenguaje de especificaciones y descripción (<i>Specification and description language</i>)
SDU	Unidad de datos de servicio (<i>Service data unit</i>)
SE	Mensaje de intercambio de sesión (<i>Session exchange message</i>)
SN	Número de secuencia (<i>sequence number</i>)

SQCIF	<i>Sub-QCIF</i>
SREJ	<i>Rechazo selectivo (selective reject)</i>
STD	<i>Decodificador objetivo de sistema (System target decoder)</i>
VC	<i>Canal virtual de ATM (ATM virtual channel)</i>
XID	<i>Identificación de intercambio (trama) (Exchange Identification (frame))</i>

Definiciones

Canal bidireccional	Está constituido por una pareja de canales lógicos asociados, uno para cada sentido de la transmisión.
Capacidad	La capacidad de una terminal está determinada cuando es capaz de codificar y transmitir o recibir y decodificar esa señal correcta.
Canal Instrucción	Enlace unidireccional entre dos puntos extremos. Mensaje que requiere una acción pero no una respuesta explícita.
Flujo elemental	Término genérico para describir una señal de vídeo codificada, una señal de audio codificada o cualquier otro tren de bits codificado.
En banda	Los mensajes en banda se transportan en dentro del canal lógico al que se refieren.
Indicación	Mensaje que contiene información pero no requiere una acción o respuesta.
Canal lógico	Trayectoria unidireccional para la transmisión de un tren elemental único.
Número de canal lógico	Número que identifica un canal lógico único.
Terminal principal	Terminal señalada como principal por el procedimiento de determinación principal-subordinada definido en la presente recomendación o por cualquier otro procedimiento. Una terminal principal puede iniciar la apertura de canales bidireccionales.
Tipo de medio	Forma única de información presentada a un usuario o datos que representan esa información. Por ejemplo el vídeo, el audio y el texto.
Modo	Conjunto de trenes elementales que transmite, pretenden transmitir o espera recibir una terminal.
Comunicación multimedia	Transmisión y/o recepción de señales de dos o más tipos de medios de manera simultánea
No normalizado	No conforme a ninguna norma nacional o internacional referenciada.

Multipunto	<i>Interconexión simultánea de tres o más terminales para permitir la comunicación entre varias ubicaciones mediante el empleo de unidades de control de multipunto (puentes) que dirigen el flujo de información de forma centralizada.</i>
Petición	<i>Mensaje que provoca una actuación por parte de la terminal distante y requiere una respuesta inmediata del mismo.</i>
Respuesta	<i>Mensaje emitido como contestación a una petición.</i>
Sesión	<i>Periodo de comunicación entre dos terminales que puede ser una conversación o una transmisión de datos.</i>
Terminal subordinada	<i>Terminal señalada como subordinada mediante el procedimiento principal-subordinada o mediante cualquier otro procedimiento. Una terminal subordinada no puede iniciar la apertura de canales bidireccionales, pero puede solicitar a la terminal principal que efectúe esa apertura.</i>
Sustentación	<i>Aptitud para el funcionamiento en un modo determinado. Sin embargo la exigencia de sustentación de un modo no implica que el modo deba utilizarse realmente en todo momento: a menos que esté prohibido, pueden utilizarse otros modos mediante negociación mutua.</i>
Terminal	<i>Cualquier punto extremo, pudiendo ser una terminal de usuario o cualquier otro sistema de comunicación como un MCU o un servidor de información.</i>

Bibliografía.

1. ABL
Manual de usuario para VT140A
ABL, 1996
2. UIT, H.223
Transmisión en Línea de Señales no Telefónicas. Protocolo de Multiplexación para Comunicación Multimedios Baja Velocidad Binaria.
UIT, 8 de Septiembre de 1995
3. UIT, H.245
Transmisión en Línea de Señales no Telefónicas. Protocolo de Control para Comunicación Multimedios
UIT, 6 de Julio de 1995
4. UIT, H.324
Transmisión en Línea de Señales no Telefónicas. Terminal para Comunicación Multimedios a Baja Velocidad Binaria
UIT, 5 de Julio de 1995
5. UIT, H.324 M
Transmisión en Línea de Señales no Telefónicas. Terminal para Comunicación Multimedios a Baja Velocidad Binaria
UIT, 11 DE Septiembre de 1995
6. SanJuan, Flores Tomas, Rey Eugenio
Telecomunicaciones Móviles
Ed. Alfaomega Marcombo, México, 1995, Pags. 163-171

CAPÍTULO 5

TRANSMISIÓN MULTIMEDIA VÍA B-ISDN Y LAN

5.1 RECOMENDACIÓN H.321

(Adaptación de las terminales videotelefónicas H.320 a los entornos RDSI-BE)

Esta Recomendación describe las especificaciones técnicas para la adaptación de las terminales videotelefónicas de banda estrecha definidas en la Recomendación H.320 a los entornos RDSI de banda ancha. La terminal definida en esta Recomendación puede funcionar con el mismo tipo de terminales es decir, con otras terminales H.321 de la RDSI-BA, así como con las terminales H.320 de la RDSI-BE.

También es posible encontrar algunas de las funcionalidades de las terminales H.321 en las terminales audiovisuales de banda ancha definidas en la Recomendación H.310. Un requisito obligatorio es el interfuncionamiento entre terminales H.310, H.321 y H.320. El interfuncionamiento entre terminales H.320 y H.321 se obtiene por el hecho de que los distintos tipos de terminales H.321 incluyen las

mismas funciones de los tipos correspondientes a la terminales H.320. El interfuncionamiento entre las terminales H.320/H.321 y H.310 se obtiene por medio de un grupo común de funciones H.320/H.321, las cuales se definen en las Recomendación H.310. Por ejemplo, además de soportar la Recomendación H.262 de vídeo (MPEG-2), las terminales H.310 soportan la Recomendación H.261 que forma parte de la Recomendaciones H.320 y H.321.

En las terminales H.321, la adaptación de las funciones H.320 en la RDSI-BA se logran mediante la capa 1 de adaptación del ATM (AAL 1). En esta Recomendación se consideran las funciones de segmentación y reensamblado (SAR) y de subcapa de convergencia (CS) definidas en la Recomendación I.363¹.

Las terminales H.321 tienen las mismas funcionalidades en la banda que las terminales H.320. Las funciones adicionales de señalización relacionadas con la banda ancha, tales como las de negociación para la utilización del método de recuperación de reloj adaptativo (modo asíncrono), se pueden lograr con los elementos de información Q.2931².

Descripción del Sistema

Esta terminal se puede utilizar para diversas aplicaciones tales como las de servicios convencionales, servicios distributivos, servicios de recuperación y servicios de mensajería, de igual forma que las terminales H.320. Esta Recomendación no especifica ningún servicio en particular.

¹ Especificación de la capa de adaptación del modo de transferencia asíncrono de la Red Digital de Servicios Integrados de banda ancha (revisada en Helsinki, 1993)

² Red Digital de Servicios Integrados de banda ancha - Señalización digital de abonado N° 2 (DSS 2) - Especificación de la capa 3 de la interfase usuario-red para el control de la llamada básica/conexión, documento de trabajo UIT-T CE 11/GT, 2 de julio de 1994, Edimburgo)

Configuración del Sistema

La terminal H.321 corresponde a la configuración de referencia RDSI-BA que se describe en la Recomendación I.413, tal como se representa en la figura 5.1.1.

Existen dos implementaciones posibles de la terminal H.321. La primera es un diseño integrado que incluye las funciones H.320, la capa de adaptación de ATM y las ATM en una sola unidad terminal (es decir, la B-TE1 de la figura 5.1.1).

La segunda implementación consiste en un equipo terminal H.320 (TE2) y un adaptador de terminal de banda ancha (B-TA). Es este caso, se transmite una señal H.320 en la interfase entre el equipo terminal H.320 y el adaptador de terminal. Además, en el adaptador de terminal se efectúa la señalización terminal-red con posibilidad de interacción con el equipo terminal H.320

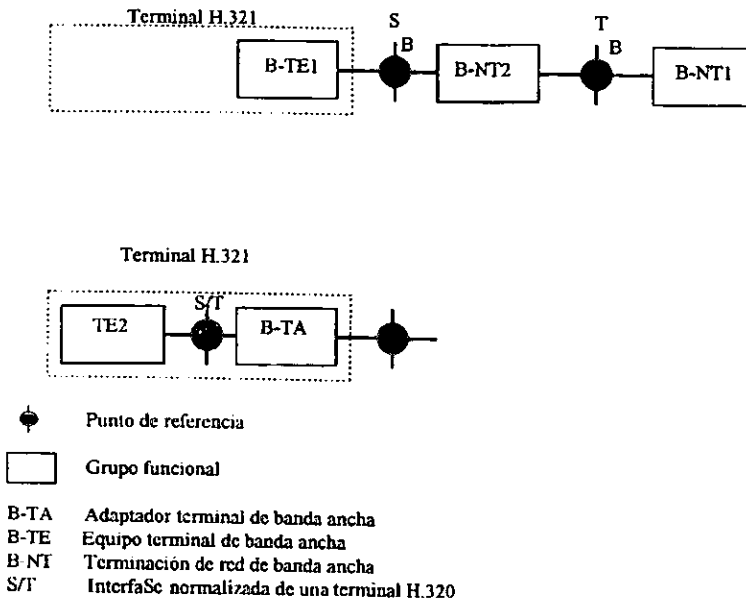


Figura 5.1.1 – Configuración de referencia

Arquitectura de la Terminal Genérica

La figura 5.1.2 muestra la arquitectura genérica de una terminal H.321 que corresponde a la configuración de la figura 5.1.1, en la que se indican los elementos constitutivos y las Recomendaciones correspondientes. La figura incluye las siguientes unidades funcionales: un equipo de entrada/salida de vídeo y uno de audio, un equipo de telemática, una unidad del control del sistema, códecs de vídeo y audio, una unidad de retardo de audio y una de multiplexación. Estas unidades se definen en la Recomendación H.320.

La AAL, el ATM y las unidades físicas realizan las funciones de adaptación y de interfase necesarias para dar cabida a un terminal H.321 en una red de banda ancha.

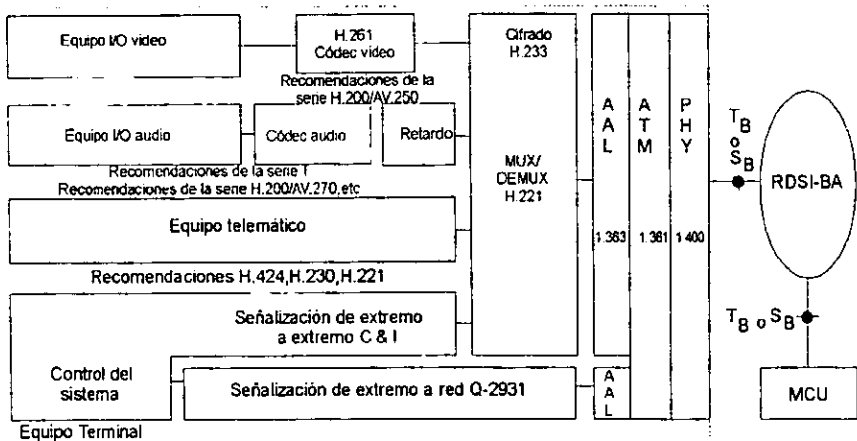


Figura 5.1.2 – Arquitectura de la terminal H.321

Tipo de Terminal

Existen diversos tipos de terminal H.321, los cuales dependen de la capacidad de acceso al canal, las clases de velocidad binaria y los esquemas de codificación de audio. Todas estas características son las que definen los tipos de terminales H.321 que se pueden utilizar y el siguiente cuadro describe los distintos modos de comunicación y de codificación de audio que soporta la H.321.

Es indispensable prestar atención especial a las capacidades de canal virtual ATM y a las funciones AAL que aplican las distintas terminales H.321.

Una terminal H.321 puede tener diversas capacidades. En las comunicaciones punto a punto, se determina un grupo común de ellas mediante los procedimientos de comunicación definidos en la H.242.

También un equipo terminal H.321 puede participar en otras terminales H.321 o H.320 en las comunicaciones multipunto a través de las MCU³ de una RDSI-BA o RDSI-BE. La figura 5.1.3 muestra un ejemplo de configuración multipunto. Los procedimientos de comunicación necesarios figuran en la Recomendación H.243⁴.

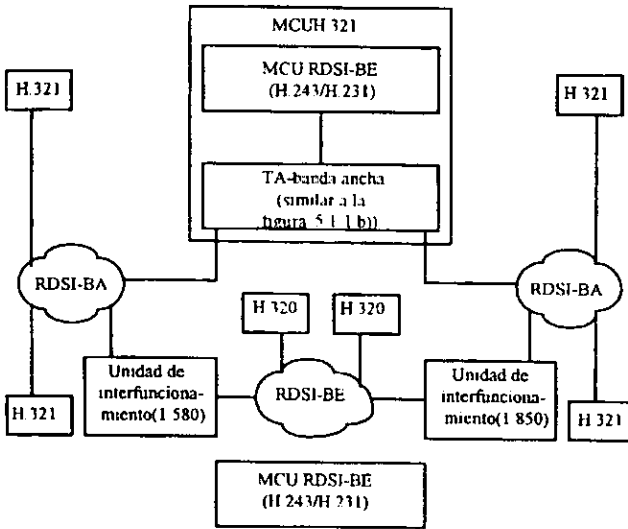


Figura 5.1.3 – Configuración multipunto para terminales H.321/H.320

Intercomunicaciones

Intercomunicaciones entre terminales H.321

Para la intercomunicación entre terminales H.321 se tiene que determinar un modo común de funcionamiento (H.320) entre terminales H.321, tal y como se describe en la Recomendación H.320. Además, cuando una terminal H.321 se comunica con otra terminal H.321, hay varias posibilidades respecto al número de canales

³ Unidades de Control Multipunto

⁴ Procedimientos para el establecimiento de comunicación entre tres o más terminales audiovisuales por canales digitales de hasta 2 Mbit/s (Helsinki, 1993)

virtuales necesarios entre las dos terminales. La figura 5.1.4 muestra los casos de VC⁵ simple y múltiple.

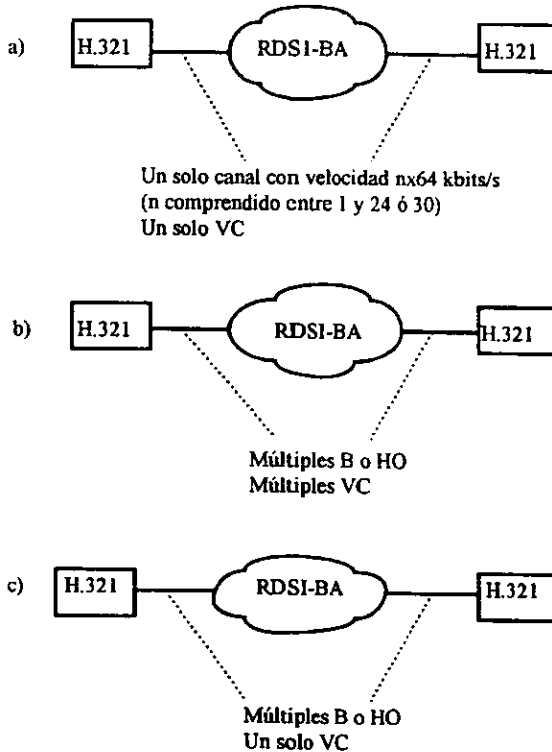


Figura 5.1.4 – Posibles interconexiones entre terminales H.321

Intercomunicación con terminales RDSI-BE

Las terminales H.321 tienen la capacidad de interfuncionamiento con terminales H.320 mediante el soporte de red de interconexión entre la RDSI-BA y la RDSI-BE que se especifica en la Recomendación I.580⁶.

⁵ Canal Virtual (del inglés Virtual Channel)

⁶ Disposiciones generales para el interfuncionamiento entre la RDSI de banda ancha u la RDSI basada en la velocidad de 64 kbit/s (Ginebra, noviembre de 1993)

Es indispensable determinar un modo común de funcionamiento entre las terminales H.321 y H.322 como se describe en la Recomendación H.320. Además, existen diversos modos de comunicación posibles entre terminales H.321 y H.321, dependiendo de, por ejemplo, el número de canales utilizados. La figura 5.1.5 muestra ejemplos de estos modos de comunicación.

Intercomunicación con la telefonía

Las terminales H.321 deben de poder funcionar con los teléfonos de la RDSI-BE y la RTPC utilizando la Recomendación G.711 de audio.

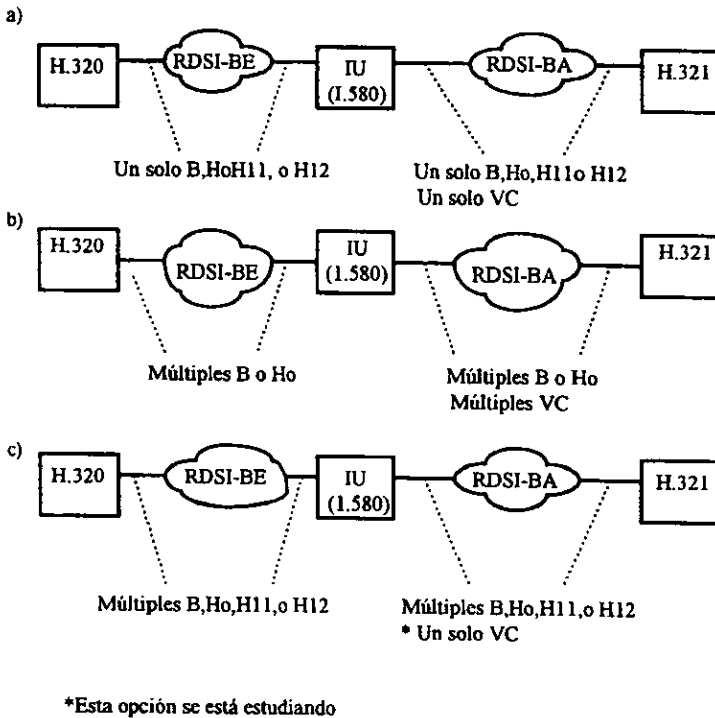


Figura 5.1.5 – Posibles interconexiones entre terminales H.320 y H.321

Intercomunicación con terminales audiovisuales conectadas a otras redes

Para poder establecer intercomunicación con terminales audiovisuales conectadas a otras redes, se tiene que determinar un modo común de funcionamiento entre terminales H.322/H.323 y H.321, tal y como se describe en la Recomendación H.320. Las terminales H.321 funcionarán con los H.322 y los H.323 de la forma como se representa en la figura 5.1.6.

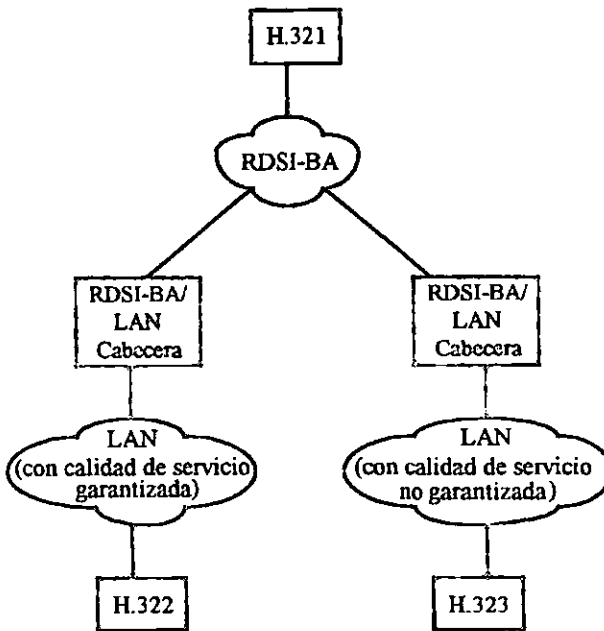


Figura 5.1.6 – Interfuncionamiento entre terminales H.322/H.323 y H.321

5.2 RECOMENDACION H.310

(Sistemas y terminales para comunicaciones audiovisuales de banda ancha)

Esta Recomendación H.310⁷ se refiere a los requisitos técnicos de los sistemas y terminales de los servicios de comunicación audiovisual de banda ancha definidos en las Recomendaciones de las series H.200/AV.100.

⁷ Esta Recomendación fue aprobada en 8 de noviembre de 1996

La H.310 define las terminales audiovisuales de banda ancha unidireccionales y bidireccionales. La clasificación de las terminales H.310 en distintos tipos de terminal se basa en las capacidades audiovisuales y de capa de adaptación ATM. Existen dos clases de terminales unidireccionales: terminal de sólo recepción (ROT) y terminal de sólo emisión (SOT).

A los tipos de terminales bidireccionales se les denomina terminales de recepción y emisión (RAST). La definición de las terminales H.310 RAST se basa en los siguientes principios de interoperabilidad:

- 1) Es obligatorio el interfuncionamiento entre los tipos de terminal H.310 RAST y otros terminales audiovisuales RDSI-BE/RDSI-BA (H.320/H.321).
- 2) También es obligatorio el interfuncionamiento entre los distintos tipos de terminal H.310 RAST.

Se definen tres tipos de terminal RAST: RAST-1, RAST-5 y RAST-1&5.

Las terminales RAST-1 y RAST-1&5 pueden conectarse a las redes públicas y a las redes en las instalaciones de usuario (redes privadas), mientras que los terminales RAST-5 sólo pueden conectarse a las redes en las instalaciones de usuario (redes privadas).

Para el interfuncionamiento con las terminales H.320/H.321, los tres tipos de terminal RAST soportan modos comunes audiovisuales H.320. Para el interfuncionamiento entre terminales RAST-5 y terminales RAST-1 y H.320/H.321, a fin de proporcionar las funciones de interoperabilidad, es necesario contar con una cabecera que no se encuentre dentro de la red pública sino en las instalaciones de usuario, entre una RDSI-BA y una red ATM en las instalaciones de usuario.

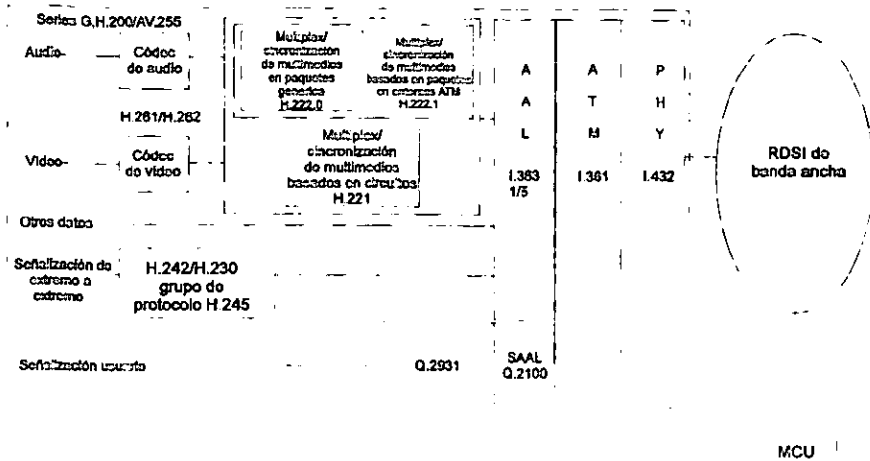


Figura 5.2.1 – Sistema de comunicación audiovisual y configuración de la terminal

La codificación de vídeo y audio y otros aspectos técnicos aplicables a más de un servicio se establecen en las Recomendaciones de las series H.200/AV.200.

La figura 5.2.1 muestra un sistema de comunicación audiovisual de banda ancha genérico. Consta de un equipo terminal, la red, la unidad de control multipunto (MCU) y los elementos que componen el equipo terminal.

Es necesario que todas las terminales H.310 soporten H.245⁸ como protocolo de control de comunicación de manera que puedan sustentar los servicios que se pretenden ofrecer e interfuncionar entre sí. Por esto, las terminales H.310 deberán utilizar los procedimientos de acuse de recibo H.222.1 para la señalización de subcanal.

⁸ Protocolo de control para comunicaciones multimedia, 1996

Es importante observar que la terminal genérica H.310 mostrada en la figura 5.2.1 puede representar cualquiera de los tipos terminales unidireccionales o bidireccionales definidos en la recomendación H.310.

La definición de los tipos de terminal H.310 está dirigida al soporte de las siguientes aplicaciones:

- Servicios conversacionales (por ejemplo, servicios de videoconferencia y de videotelefonía).
- Servicios de consulta.
- Servicios de mensajería.
- Servicios de distribución con presentación individual por el destinatario (por ejemplo, servicios de vídeo por demanda).
- Servicios de distribución sin presentación individual por el destinatario (por ejemplo, servicios de radiodifusión de televisión).
- Transmisión de vídeo.
- Vigilancia.

Descripción del Sistema

Configuración del Sistema

La interacción entre las capacidades de la terminal H.310 se basa en el modelo de referencia de protocolo mostrado en la figura 5.2.2, que representa las pilas de protocolo para las señales audiovisuales, de datos, de gestión de llamada (DSS 2 y H.245) y otras señales de control e indicación que pueden ser soportadas por los distintos tipos de terminal H.310.

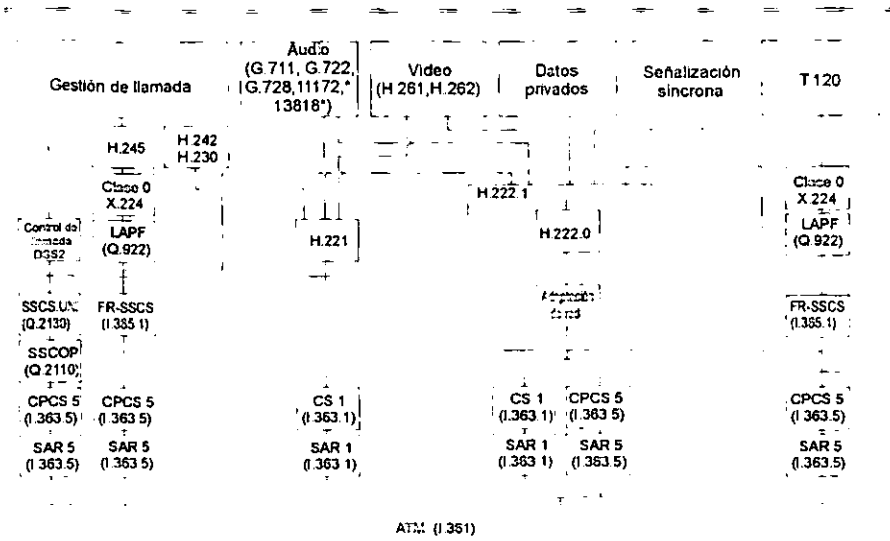


Figura 5.2.2 – Modelo de referencia de protocolo H.310

Tipos de Terminal

La clasificación de las terminales H.310 en los distintos tipos de terminales se basa en las capacidades audiovisuales y AAL que se resumen en el siguiente cuadro.

		AAL			
		AAL 1	AAL 5	AAL 1&5	
Transporte Audiovisual	Unidireccional	ROT	ROT-1	ROT-5	ROT-1&5
		SOT	SOT-1	SOT-5	SOT-1&5
	Bidireccional	RAST	RAST-1	RAST-5	RAST-1&5

Tipos de terminal unidireccional (ROT y SOT)

Se definen dos clases de terminales unidireccionales: terminal de sólo emisión (SOT) y terminal de sólo recepción (ROT).

Se definen tres tipos de terminales unidireccionales H.310 basándose

en sus AAL soportadas para cada una de las dos clases. Los tipos de terminal unidireccional definidos en H.310 son los siguientes:

- H.310 ROT-1 y SOT-1 que soportan AAL 1
- H.310 ROT-5 y SOT-5 que soportan AAL 5
- H.310 ROT-1&5 y SOT-1&5 que es una terminal compuesta que soporta tanto AAL 1 como AAL 5

Cada uno de estos tipos de terminal deberá soportar el modo de comunicación nativo H.310. Este modo consiste en los protocolos de audio, video y control de H.222.1, con ISO/CEI 11172-3 capa II, H.262 y H.245, respectivamente.

Cada una de estas terminales puede estar conectada a la RDSI-BA pública y a las redes en las instalaciones de usuario (redes privadas). Algunos pares de terminales unidireccionales no interfundionarán entre sí. Ello puede deberse a que son de clases incompatibles, tales como ROT-1 conectado con ROT-1&5, o a que son de tipos incompatibles, tales como ROT-1 conectado a SOT-5.

Tipos de terminal bidireccional (RAST)

Existen tres tipos de terminales de recepción y emisión bidireccional (RAST) H.310 basándose en sus modos de comunicación y AAL soportadas. Los tipos de terminales definidos H.310 son:

- H.310 RAST-1, que soporta AAL 1;
- H.310 RAST-5, que soporta AAL 5;
- H.310 RAST-1 & 5, una terminal compuesta que soporta AAL 1 y AAL 5.

Cada uno de estos tipos de terminal deberá soportar un modo de comunicación nativo H.310 así como un modo de interoperación

H.320/H.321. La figura 5.2.3 representa los grupos de protocolos para estos dos modos en cada uno de los tipos de terminal.

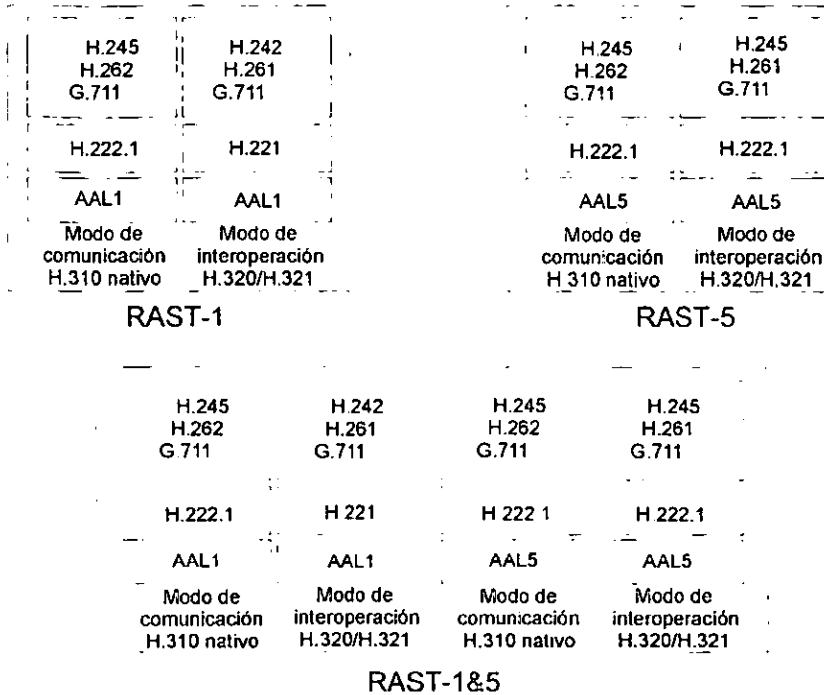


Figura 5.2.3 – Grupos de protocolos para comunicación e interoperación entre H.310 y H.320/H.321 nativos

Capacidades de la Terminal

La definición y clasificación de los tipos de terminal H.310 y sus modos de comunicación se basan en las siguientes capacidades:

- Audiovisual y datos.
- Adaptación de red.
- Señalización (usuario a usuario y usuario a red).

Un modo de comunicación se define como una combinación de ciertos parámetros con las capacidades antes indicadas. Basándose en las distintas capacidades de las terminales H.310 se especifican dos clases de modos de comunicación.

- Modos de interoperación H.320/H.321.
- Modos de comunicación H.310 nativos.

Las terminales unidireccionales H.310 sólo deben soportar modos de funcionamiento de comunicación H.310 nativos; es decir, las terminales unidireccionales pueden soportar opcionalmente los modos de interoperación H.320/H.321.

Al principio de la llamada, las terminales H.310 deberán identificar el tipo de terminal distante (H.320/H.321, H.310 bidireccional, etc.) a través del intercambio en los elementos de información Q.2931 y deberá utilizar H.245 o H.242 para llevar a cabo el intercambio de la capacidad y otros procedimientos.

Capacidades de vídeo

Todas las terminales H.310 deberán soportar la Recomendación H.262 en el punto de conformidad de perfil principal de nivel principal (MP@ML). De manera opcional, se pueden soportar otros perfiles y niveles H.262.

Las terminales H.310 deberán satisfacer la relación jerárquica definida en la cláusula 8/H.622, entre los distintos perfiles y niveles H.262. Por consiguiente, cabe señalar que todas las terminales H.310 deberán soportar el perfil sencillo del nivel principal (SP@ML) H.262 y el perfil principal de bajo nivel (MP@LL) H.262.

Cuando se satisfaga el punto de conformidad H.262 (es decir, se soporte un perfil determinado a un nivel determinado) las terminales H.310 deberán satisfacer las limitaciones del parámetro vídeo especificadas en la cláusula 8/H.262 para dicho punto de conformidad. No obstante, las terminales pueden indicar una capacidad para la funcionalidad de un perfil determinado con parámetros de nivel reducido utilizando H.245 y ese modo de comunicación puede establecerse empleando los procedimientos pertinentes de H.245 cuando ambas terminales tienen la capacidad necesaria.

En el cuadro 5.2.a se resumen las capacidades de vídeo y audio obligatorias y opcionales.

Capacidades de datos

El soporte de protocolos de datos es opcional en las terminales H.310. Los protocolos de datos pueden multiplexarse con otra información audiovisual utilizando las funciones de multiplexación H.222.1 o de capa ATM.

Capacidades de adaptación de red (NAC)

Las capacidades de adaptación de red de las terminales H.310 incluyen las capacidades del mecanismo de multiplexación y sincronización de multimedios, de capa de adaptación ATM, de velocidad de transferencia y de VC ATM.

Tipo de terminal		Capacidades audiovisuales				
		Video			Audio	
		Obligatoria	Opcional		Obligatoria	Opcional
ROT-1 SOT-1 ROT-5 SOT-5 ROT-1&5 SOT-1&5		H.262 MP@ML	H.262	MP@H14L MP@HL SNR@LL SNR@ML Espacial@H14L HP@ML HP@H14L HP@HL	11172-3 Capa II	11172-3 Capa III 13818-3 Capa I 13818-3 Capa II 13818-3 Capa III G.711 G.722 G.728 G.723.1 G.729 Anexo A a G.729
		H.262 MP@ML	H.262	MP@H14L MP@HL SNR@LL SNR@ML Espacial@H14L HP@ML HP@H14L HP@HL	11172-3 Capa II	11172-3 Capa III 13818-3 Capa I 13818-3 Capa II 13818-3 Capa III G.711 G.722 G.728 G.723.1 G.729 Anexo A a G.729
		H.261 QCIF H.261 CIF	H.262	MP@ML MP@H14L MP@HL SNR@LL SNR@ML Espacial@H14L HP@ML HP@H14L HP@HL H.263 SQCIF H.263 QCIF H.263 CIF H.263 4CIF H.263 16CIF	G.711	11172-3 Capa I 11172-3 Capa II 11172-3 Capa III 13818-3 Capa I 13818-3 Capa II 13818-3 Capa III G.722 G.728 G.723.1 G.729 Anexo A a G.729
	Modo H.320/1					

Cuadro 5.2.a - Capacidades audiovisuales de las terminales H.310

Multiplexación y sincronización de multimedios

La multiplexación de las señales de audio, vídeo, datos y control de los terminales H.310 se logra utilizando los protocolos y procedimientos de multiplexación de multimedios definidos en las Recomendaciones H.221 y H.222.1. La Recomendación H.222.1 especifica las funciones concretas H.222.0 y H.222.1 (por ejemplo, protección contra errores, supresión de la fluctuación de fase, identificadores y descriptores de trenes definidos por el UIT-T).

Todas las terminales H.310 deberán soportar las Recomendaciones H.222.0/H.222.1 para la multiplexación y sincronización de las señales de vídeo, audio, datos y control. El soporte de la multiplexación del tren de transporte H.222.1 es obligatorio para todas las terminales H.310. El soporte de la multiplexación del tren de programa H.222.1 es opcional para las terminales H.310.

Las terminales H.310 bidireccionales, salvo las terminales RAST-5, deberán soportar la Recomendación H.221 para el interfuncionamiento con las terminales H.320 y H.321. En el caso de las terminales RAST-5 y las terminales RAST-1&5 funcionando con AAL 5, deberá utilizarse un *gateway* en la red situado en las instalaciones de usuario para generar el múltiplex H.221 y los mensajes de control H.242.

Capa de adaptación ATM

Los distintos tipos de señales soportados por las terminales H.310 exigen diferentes funciones de capa de adaptación ATM.

Las funciones segmentación y reensamblado (SAR) y sus capas de convergencia (CS) AAL son soportadas por las terminales H.310. Los tipos 1 y 5 de AAL se utilizan para la transferencia de las señales de velocidad binaria constante multiplexadas por las redes ATM.

Velocidad de transferencia

Las terminales H.310 son capaces de soportar una amplia gama de posibles velocidades de transferencia. Las capacidades de velocidad de transferencia de las terminales H.310 se definen en AAL-SAP.

Intercomunicaciones

El interfuncionamiento entre terminales H.310 bidireccionales (RAST) y H.320/H.321 es obligatorio. Además, también lo es el interfuncionamiento entre los distintos tipos de terminales H.310 bidireccionales. La figura 5.2.4 representa los casos de interfuncionamiento para las terminales H.310.

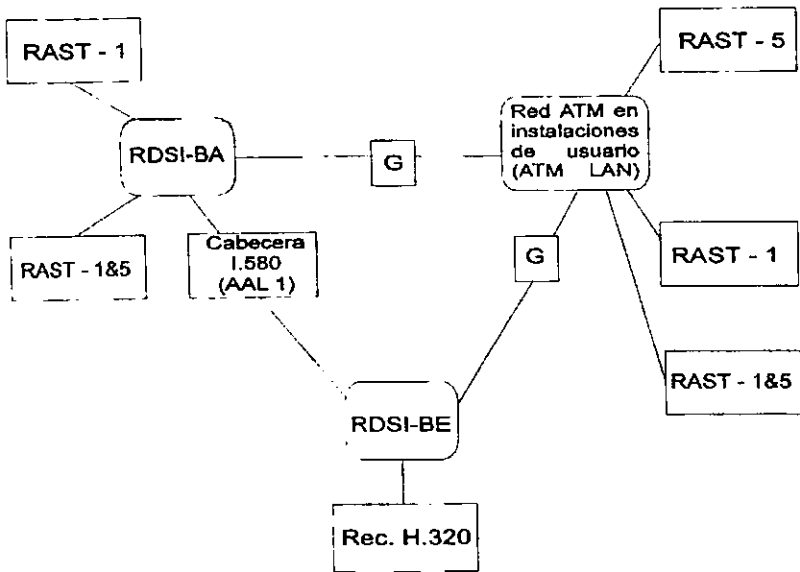


Figura 5.2.4 – Interfuncionamiento con terminales H.310

Como el soporte de AAL tipo 1 (para la transferencia de las señales audiovisuales H.222.1 y H.221 por la RDSI-BA) es obligatorio para las terminales H.310 RAST-1 y RAST-5, estas terminales pueden

interfuncionar entre si y con H.321 sin una función de *gateway*.

Las terminales RAST-5 y RAST-1&5 que funcionan en modo AAL 5 deberán interfuncionar con las terminales RAST-1 y las terminales RAST-1&5 que funcionan en modo AAL 1 y con H.321 con el *gateway* situado en la red ATM de las instalaciones de usuario. Este *gateway* es necesario para proporcionar funciones de interfuncionamiento entre:

- a) RAST-5 y RAST-1;
- b) RAST-5 y RAST-1&5 funcionando en modo AAL 1; y
- c) RAST-5 o RAST-1&5 funcionando en modo AAL 5 y terminales H.321.

De forma similar, se necesita de un *gateway* entre una RDSI-BE y una red ATM en las instalaciones de usuario para proporcionar las funciones de interfuncionamiento entre terminales RAST-5 y terminales RAST-1&5 funcionando en modo AAL 5 y terminales H.320.

Es importante observar que las terminales H.310 RAST-1 y las terminales H.310 RAST-1&5 pueden instalarse tanto en la RDSI-BA como en las redes ATM en las instalaciones de usuario, o establecer interfase con las mismas. Sin embargo, las terminales H.310 RAST-5 pueden establecer interfase únicamente con las redes en las instalaciones de usuario.

Intercomunicación con terminales RDSI-BE*

Para el interfuncionamiento con terminales H.320/H.321, todos los tipos de terminal RAST deberán soportar los siguientes modos:

- a) H.261 CIF/QCIF.
- b) G.711.

* La sincronización del reloj de las terminales conectadas a diferentes tipos de redes esta en estudio.

Adicionalmente, para el interfuncionamiento con terminales H.320/H.321, todos los tipos de terminales RAST deberán soportar los siguientes modos. Las terminales RAST-5 deberán soportar estas funciones en un *gateway* situado en la red de las instalaciones de usuario.

- c) H.221/H.242-H.230.
- d) Modos de transferencia 1B, 2B y H0.
- e) Dos VC de ATM (para soportar el modo de comunicaciones 2B con H.320).

5.3 RECOMENDACION H.322

(Sistemas y equipos terminales videotelefónicos para las redes de área local que dan un servicio con calidad garantizada)

La Recomendación H.322, define los requisitos técnicos de los servicios videotelefónicos de banda estrecha, en las situaciones en las que el trayecto de la transmisión incluye una o más redes de área local (LAN), la cual o las cuales están configuradas para ofrecer una calidad de servicio (QoS) equivalente a la calidad de servicio que ofrece la RDSI-BE, de tal forma que no sea necesario que las terminales ofrezcan una protección adicional o mecanismos de recuperación distintos a los que se estipulan en la Recomendación H.320¹⁰.

Debido a la adopción generalizada y a la creciente utilización de los servicios de videotelefonía y videoconferencia por redes que se ajustan a la característica de RDSI-BE y debido a la conveniencia y las

¹⁰ Sistemas y equipos terminales videotelefónicos de banda estrecha (revisada en Helsinki, 1993)

ventajas de permitir que los servicios que estas ofrecen se transmitan, total o parcialmente por redes de área local, manteniendo al mismo tiempo la capacidad de interfuncionamiento con las terminales H.320, y tomando en cuenta las características de muchos tipos de red de área local, es entonces que se genera la necesidad de especificar normas para que sean cumplidas por los equipos que transmitan a través de este tipo de redes, por lo que se generan las Recomendaciones H.322 o H.323.

Como se mencionó previamente la Recomendación H.322 abarca los requisitos técnicos de los servicios videotelefónicos de banda estrecha en aquellas situaciones en las que el trayecto de transmisión incluye una o más redes de área local, estando estas configuradas y gestionadas para dar una calidad de servicio (QoS) garantizada equivalente a la de la RDSI-BE, de tal forma que no sea necesario que las terminales ofrezcan una protección adicional o mecanismos de recuperación distintos de los que define la Recomendación H.320. Los parámetros de interés son los errores de datos y las propiedades de pérdidas, así como la variación del retardo de tránsito. La recomendación H.322 exige también que el reloj de la RDSI esté disponible en las terminales¹¹.

La Recomendación H.322 solo se enfoca a aquellas redes que son capaces de dar una calidad de servicio ~~garantizada~~, mientras que aquellas LAN's que no ofrecen esta calidad, se describen en la Recomendación H.323.

Aún cuando algunas LAN ofrecen una QoS satisfactoria para trabajar con la Recomendación H.322, estas no incorporan medios normalizados de identificación de terminal a la que se pretende llamar,

¹¹ Una LAN adecuada para la utilización de la Recomendación H.322 sería una LAN de servicios integrados (IS): servicios isócronos IEEE 802.9a de acceso múltiple con detección de portadora y detección de colisiones (CSMA/CD) y servicios de control de acceso a los medios (MAC)

por lo que para dichas LAN se deberá aplicar la Recomendación H.323.

En la Recomendación H.322, no se tocan las características de las LAN ATM, ya que estas entran en el ámbito de la Recomendación H.321.

Los sistemas y el equipo terminal de la Recomendación H.322 pueden funcionar entre sí y con aquellos que se definen por la Recomendaciones H.320, H.321 y H.323. La figura 5.3.1 nos muestra el concepto, en el cual, cualquier terminal puede conectarse con cualquiera otra. La unidad de cabecera H.322 ofrece una interconexión entre la LAN y una red de área extensa (WAN) que puede ser una RDSI-BE o la RDSI-BA, o ambas. Una terminal H.322 se comunica con otra terminal H.322 directamente por la misma LAN. La cabecera puede conectarse a través de la RDSI-BE o la RDSI-BA, con otras cabeceras y otras redes LAN para establecer comunicación entre terminales H.322 o H.323 que no estén en la misma LAN.

Aún cuando esta Recomendación aborda específicamente los sistemas videotelefónicos, los métodos utilizados no dependen del contenido de las señales transmitidas. Es por esto, que esta Recomendación tiene una aplicación mucho más general que la conexión de terminales concebidas originalmente para la RDSI-BE, para las LAN híbridas y las redes RDSI.

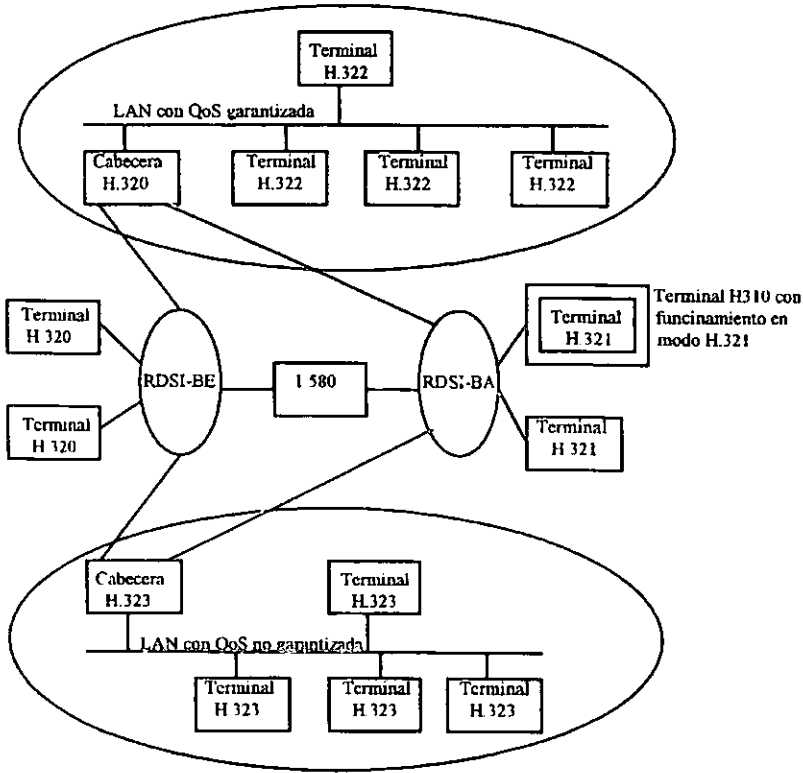


Figura 5.3.1 – Interfuncionamiento de la H.322

Equipo Terminal

La figura 5.3.2 representa un diagrama de bloques de la terminal especificada en esta Recomendación. Esta terminal es idéntica en todos sus elementos a los especificados en la Recomendación H.320, excepto en lo referente a la interfase, la cual en este caso es una interfase LAN.

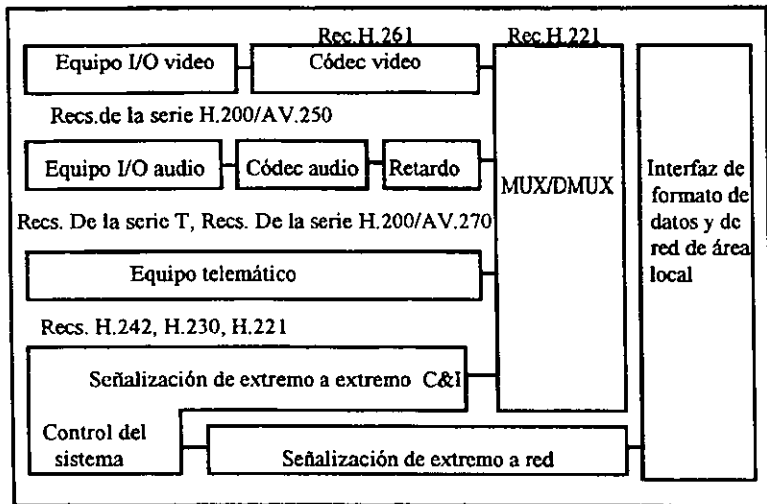


Figura 5.3.2 – Diagrama de bloques de la terminal H.322

5.4 RECOMENDACIÓN H.323¹²

(Sistemas y equipos videotelefónicos para redes de área local que proporcionan una calidad de servicio no garantizada)

Esta recomendación describe terminales, equipos y servicios para comunicaciones multimedia por redes de área local (LAN, *local area networks*) que proporcionan una calidad de servicio **no garantizada**. Las terminales y equipos H.323 pueden transportar voz en tiempo real, datos y vídeo, o cualquier combinación de los mismos incluyendo la videotelefonía.

¹² La Recomendación UIT-T H.323 fue aprobada el 8 de noviembre de 1996

Las LAN a través de las cuales se comunican las terminales H.323 pueden ser un solo segmento o un anillo o múltiples segmentos con topologías complejas. La explotación de terminales H.323 por los múltiples segmentos de la LAN (incluido el Internet) puede dar lugar a una calidad de funcionamiento deficiente. Las posibles maneras de garantizar una calidad de servicio en esos tipos de LAN/interredes quedan fuera del alcance de esta Recomendación.

Las terminales H.323 pueden estar integrados en computadoras personales o implementadas en dispositivos independientes, tales como los videoteléfonos. El soporte de voz es obligatorio mientras que el de datos y video es opcional, pero si se soportan, es preciso poder utilizar un modo común especificado, de tal manera que las terminales que soportan ese tipo de medios puedan interfuncionar. La Recomendación H.323 permite que estén en uso más de un canal de cada tipo. Las Recomendaciones relacionadas con la serie H.323 son: la H.225.0¹³ relativa a empaquetado y sincronización, la H.245¹⁴, la H.261¹⁵ y H.263¹⁶ relativas a códecs de video, las G.711¹⁷, G.722¹⁸, G.728¹⁹, G.729²⁰ y G.723²¹ sobre códecs de audio y las Recomendaciones de la serie T.120²² sobre protocolos de comunicaciones multimedia.

Esta Recomendación utiliza los procedimientos de señalización de

¹³ Paquetización y sincronización de trenes de medios para sistemas de videotelefonía en redes de área local con calidad de servicio no garantizada.

¹⁴ Protocolo de control para comunicaciones multimedia.

¹⁵ Códec de video para servicios audiovisuales a p x 64kbit/s.

¹⁶ Codificación de video para comunicación a baja velocidad binaria.

¹⁷ Modulación por impulsos codificados (MIC) de frecuencias vocales.

¹⁸ Codificación de audio de 7kHz dentro de 64 kbit/s.

¹⁹ Codificación de señales locales a 16 kbit/s utilizando predicción lineal con excitación por código de bajo retardo.

²⁰ Codificación de la voz a 8 kbit/s mediante predicción lineal con excitación por código algebraico de estructura conjugada.

²¹ Codificadores vocales. Codificador de voz de doble velocidad para transmisión en comunicaciones multimedia de 5.3 y 6.3 kb/s.

²² Protocolos de datos para conferencias multimedia.

canal lógico de la Recomendación H.245, en la que se describe el contenido de cada canal lógico cuando se abre el canal. Puesto que los procedimientos de la Recomendación H.245 también son utilizados por la Recomendación H.310 para las redes ATM, la Recomendación H.324 para la RTGC²³ y la Recomendación V.70, el interfuncionamiento con estos sistemas no debería necesitar conversión de H.242 a H.245, como sería el caso para los sistemas de la Recomendación H.320.

Las terminales H.323 se pueden utilizar en configuraciones multipunto y pueden interfuncionar con terminales H.310 en la RDSI-BA, terminales H.320 en la RDSI-BE, terminales H.321 en la RDSI-BA, terminales H.322 en las LAN de calidad de servicio garantizada, terminales H.324 en la RTGC y redes inalámbricas y terminales V.70 en la RTGC (figura 5.4.1).

Alcances

La Recomendación H.323 abarca los requisitos técnicos de los servicios de videotelefonía en banda estrecha definidos en las Recomendaciones de las series H.200/AV.120, en aquellas situaciones en las que el trayecto de transmisión incluye una o más redes de área local (LAN), que quizás no proporcionen una calidad de servicio (QoS, *quality of service*) garantizada equivalente a la de la RDSI-BE. Ejemplos de este tipo LAN son:

- Ethernet (IEEE 802.3).
- Fast Ethernet (IEEE 802.10).
- FDDI (modo calidad de servicio no garantizada).
- Token Ring (IEEE 802.5).

²³ Red Telefónica General Conmutada.

La Recomendación H.322 trata el caso de los servicios de videotelefonía en aquellas situaciones en las que el trayecto de transmisión incluye una o más redes de área local (LAN) que se configuran y gestionan para proporcionar una calidad de servicio (QoS) equivalente a la de la RDSI-BE, con lo que no es preciso proporcionar en las terminales otros mecanismos de protección o recuperación que los especificados por la Recomendación H.320. Los parámetros relevantes son las características de errores en los datos y atenuación y la variación del retardo de tránsito. Una LAN adecuada sería la LAN de servicios integrados: Servicios isócronos IEEE 802.9A²⁴.

El ámbito de aplicación de la H.323 no incluye la propia LAN o la capa de transporte que puede ser utilizada para conectar diversas LAN, sino únicamente los elementos necesarios para la interacción con la red con conmutación de circuitos (RCC). La combinación del *gateway* (GW)²⁵ H.323, la terminal H.323 y la LAN fuera de alcance aparece en la RCC como una terminal H.320, H.310, H.324 o V.70.

En la Recomendación H.323 se describen los componentes de un sistema H.323, lo que incluye terminales, *gateways*, *gatekeepers* (GK)²⁶, controladores multipunto (MC)²⁷, procesadores multipunto (MP)²⁸ y unidades de control multipunto (MCU)²⁹. Los mensajes de control y los procedimientos de la H.323 definen la manera de comunicar de esos componentes.

²⁴ Isochronous services with Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Media Access Control (MAC) service

²⁵ El GATEWAY es un punto extremo en la red de área local que proporciona comunicaciones en ambos sentidos en tiempo real entre terminales H.323 de la LAN y otras terminales en una red de área amplia.

²⁶ El GATEKEEPER es una entidad H.323 (cualquier componente H.323) de la LAN que facilita la traducción de direcciones y controla el acceso a la red de área local para otras entidades H.323.

²⁷ Es una entidad H.323 de la LAN que permite el control de tres o más terminales que participan en una conferencia multipunto.

²⁸ Es una entidad H.323 de la LAN que permite el procesamiento centralizado de los trenes de audio, vídeo y/o datos en una conferencia multipunto.

²⁹ Es un punto extremo de la LAN que permite que tres o más terminales y *gateways* participen en una conferencia multipunto.

Los componentes que se describen están formados por puntos extremos H.323 y entidades H.323. Los puntos extremos pueden llamar y son llamables de acuerdo con los procedimientos de establecimiento de la llamada. Las entidades no son llamables, pero pueden ser direccionadas para efectuar sus funciones específicas. Por ejemplo, una terminal no puede efectuar una llamada a un *gatekeeper*, pero el *gatekeeper* es direccionado como parte de los procedimientos de establecimiento de la llamada.

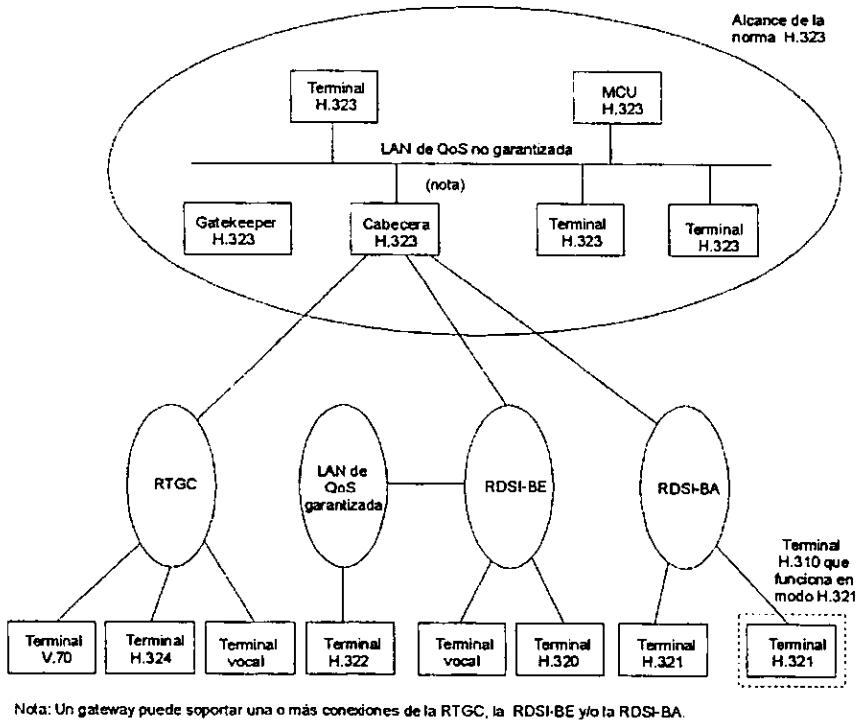


Figura 5.4.1 – Interfuncionamiento de la terminal H.323

Los componentes videotelefónicos se comunican mediante la transmisión de trenes de información. Los trenes de información se

clasifican en trenes de vídeo, audio, datos, control de las comunicaciones y control de la llamada de la siguiente manera.

- Señales de audio, que contienen señales vocales digitalizadas y codificadas. Para reducir la velocidad binaria media de las señales de audio se puede proporcionar activación por la voz. La señal de audio va acompañada por una señal de control de audio.
- Señales de vídeo, que contienen vídeo en movimiento digitalizado y codificado. El vídeo se transmite a una velocidad no superior a la seleccionada como resultado del intercambio de capacidad. La señal de vídeo va acompañada por una señal de control de vídeo.
- Señales de datos, que incluyen imágenes fijas, facsímil, documentos, archivos de computadora y otros trenes de datos.
- Señales de control de las comunicaciones, que transfieren datos de control entre elementos funcionales como si estuvieran distantes y se utilizan para el intercambio de capacidad, apertura y cierre de canales lógicos, control de modo y otras funciones que forman parte del control de las comunicaciones.
- Señales de control de la llamada, que se utilizan para el establecimiento de llamadas, la desconexión de las mismas y otras funciones del control de la llamada.

Los trenes de información descritos anteriormente son formateados y enviados a la interfase de red como se describe en la Recomendación H.225.0.

Características de las Terminales

En la figura 5.4.2, se muestra un ejemplo de terminal H.323. El diagrama muestra las interfaces del equipo de usuario, el códec de vídeo, el códec de audio, la capa H.225.0, las funciones de control del sistema y la interfase con la LAN. Todas las terminales H.323 tendrán

una unidad de control del sistema, capa H.225.0, interfase de red y unidad códec de audio. La unidad códec de vídeo y las aplicaciones de datos de usuario son opcionales.

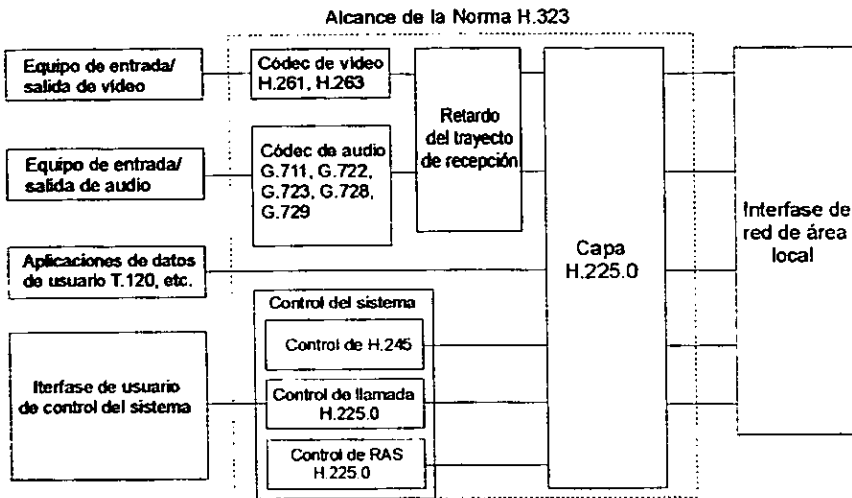


Figura 5.4.2 – Terminal H.323

Elementos de la terminal fuera del alcance de la H.323

Los siguientes elementos quedan fuera del alcance de la H.323 y, por consiguiente, no se definen:

- Los dispositivos de audio asociados, que proporcionan detección de activación por la voz, micrófono y altavoz, instrumento telefónico o equivalente, mezcladores de micrófonos múltiples y compensación del eco acústico.
- El equipo de vídeo asociado, que proporciona cámaras y monitores y su control y selección, así como el procesamiento de vídeo para mejorar la compresión o proporcionar funciones de división de la pantalla.
- Las aplicaciones de datos e interfaces de usuario asociadas,

que emplean la Recomendación T.120 u otros servicios de datos por el canal de datos.

- La interfase de red asociada, que proporciona la interfase con la LAN, soportando la señalización apropiada y los niveles de voltaje de acuerdo con las normas nacionales e internacionales.
- El control del sistema por el usuario humano, la interfase de usuario y su funcionamiento.

Elementos de la terminal dentro del alcance de la H.323

Los siguientes elementos quedan dentro del alcance de la H.323 y, por consiguiente, son objeto de normalización:

- El códec de vídeo (Recomendación H.261, etc.), que codifica el vídeo a partir de la fuente de vídeo (es decir, una cámara) para transmisión y decodifica el código de vídeo recibido, que es la salida hacia una presentación visual del vídeo.
- El códec de audio (Recomendación G.711, etc.), que codifica la señal de audio del micrófono para transmisión y decodifica el código de audio recibido que es la salida hacia el altavoz.
- El canal de datos, que soporta aplicaciones de comunicación tales como pizarrones electrónicos, transferencia de imágenes fijas, intercambio de archivos, acceso a bases de datos, etc.
- La unidad de control del sistema (Recomendaciones H.245 y H.225.0), que proporciona la señalización para un funcionamiento adecuado de la terminal H.323. Permite el control de la llamada, el intercambio de capacidad, la señalización de instrucciones e indicaciones y facilita mensajes de apertura y descripción completa del contenido de los canales lógicos.
- La capa H.225.0 (Recomendación H.225.0), que establece el formato de los trenes de vídeo, audio, datos y control

transmitidos en mensajes de salida hacia la interfase de la red y recupera los trenes de vídeo, audio, datos y control recibidos de los mensajes que llegan desde la interfase de la red. Además, lleva a cabo la detección de errores y la corrección de los mismos según conviene a cada tipo de medio.

Características del Gateway

El *gateway* proporcionará la conversión adecuada entre formatos de transmisión (por ejemplo, H.225.0 a/de H.221) y entre procedimientos de comunicaciones (por ejemplo, H.245 a/de H.242). El *gateway* llevará a cabo además el establecimiento y la liberación de la llamada en el lado LAN y en el lado RCC. La conversión entre formatos de vídeo, audio y datos también puede efectuarse en el *gateway*. Por lo general, la finalidad del *gateway* (cuando no funciona como una MCU), consiste en reflejar las características de un punto extremo de LAN a un punto extremo de RCC, y a la inversa, de manera transparente.

Un punto extremo H.323 puede comunicar con otro punto extremo H.323 de la misma LAN directamente y sin que participe en ello un *gateway*. Se puede prescindir del *gateway* si no se requieren comunicaciones con terminales RCC (terminales no en la LAN). También es posible que una terminal de un segmento de la LAN llame al exterior a través de un *gateway* y de nuevo a la LAN a través de otro *gateway* para evitar un direccionador o un enlace de bajo ancho de banda.

El *gateway* tiene las características de una terminal o una MCU H.323 de la LAN y de la terminal o MCU de RCC de la RCC. El *gateway* proporciona la conversión necesaria entre los diferentes tipos de terminal. El *gateway* puede funcionar al principio como una terminal, pero utilizando más tarde la señalización H.245 empieza a funcionar como una MCU para la misma llamada que inicialmente era punto a

punto. Los *gatekeepers* saben qué terminales son *gateways* ya que esto es algo que se indica cuando la terminal y/o el *gateway* se registra en el *gatekeeper*.

En la figura 5.4.3 se representan cuatro ejemplos de *gateways* H.323. Los diagramas muestran la función de terminal o MCU H.323, la función de terminal o MCU RCC y la función de conversión. El *gateway* tiene, para las otras terminales H.323 de la LAN la apariencia de una o más terminales H.323 o una MCU H.323.

Los *gateways* que sustentan el interfuncionamiento con terminales sólo vocales de la RTGC o la RDSI deben generar y detectar señales de multifrecuencia bitono (DTMF) correspondientes a las indicaciones de entrada de usuario de la Recomendación H.245 para 0-9, *, y #.

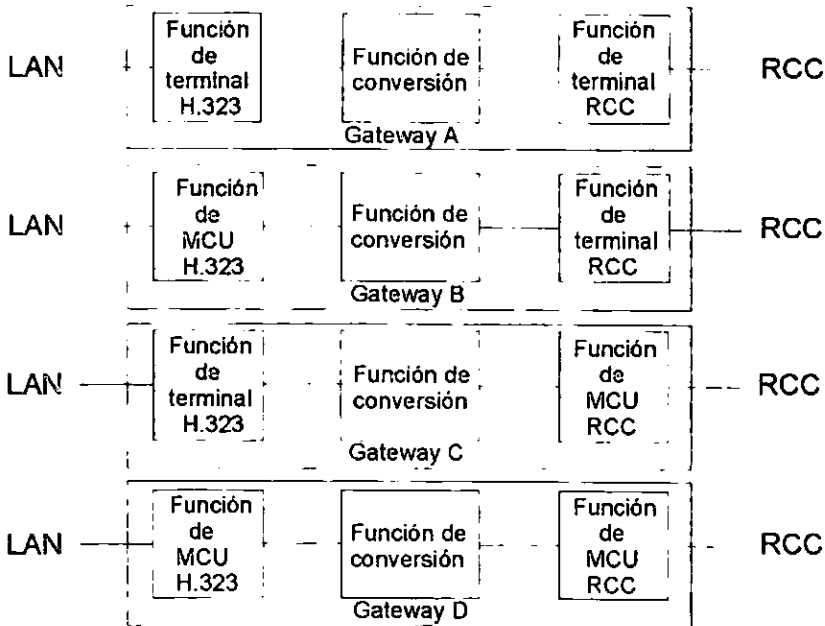


Figura 5.4.3 – Configuraciones del gateway H.323

La función de conversión proporciona la conversión necesaria de formato de transmisión y trenes de control, audio, vídeo y/o datos entre las diferentes Recomendaciones de terminales.

Características del Gatekeeper

El *gatekeeper* (que es opcional en un sistema H.323) presta servicios de control de llamada a los puntos extremos H.323. Puede existir más de un *gatekeeper* y comunicarse con cada uno de los demás de una manera no especificada. El *gatekeeper* está separado lógicamente de los puntos extremos. Sin embargo, su implementación física puede coexistir con una terminal, MCU, *gateway*, MC u otro dispositivo de LAN no H.323.

Cuando existe en un sistema, el *gatekeeper* presta los siguientes servicios:

- Conversión de dirección – El *gatekeeper* efectuará la conversión de dirección de alias a dirección de transporte. Esto se debe hacer utilizando un cuadro de conversión.
- Control de admisiones – El *gatekeeper* autorizará el acceso a la LAN utilizando mensajes ARQ³⁰/ACF³¹/ARJ³² de la Recomendación H.225.0. La autorización del acceso puede basarse en la autorización de la llamada o en el ancho de banda.
- Control de ancho de banda – El *gatekeeper* admitirá mensajes BRQ³³/BRJ³⁴/BCF³⁵. Esto puede basarse en la gestión del ancho de banda.

³⁰ Petición de admisión (admission request)

³¹ Confirmación de admisión (admission confirmation)

³² Rechazo de admisión (admission reject)

³³ Petición de cambio de ancho de banda (bandwidth change request)

³⁴ Rechazo de cambio de ancho de banda (bandwidth change reject)

³⁵ Confirmación de cambio de ancho de banda (bandwidth change confirmation)

- Gestión de zona – El *gatekeeper* proporcionará las funciones anteriores para terminales, MCU y *gateways* que se hayan registrado en él.

El *gatekeeper* también puede efectuar otras funciones opcionales, tales como:

- Señalización de control de llamada – El *gatekeeper* puede optar por completar la señalización de la llamada con los puntos extremos y puede procesar él mismo la señalización de la llamada.
- Autorización de llamada – Utilizando la señalización H.225.0, el *gatekeeper* puede rechazar llamadas procedentes de una terminal por ausencia de autorización. Pueden ser motivos de rechazo, entre otros, el acceso restringido hacia/desde terminales o *gateways* particulares y el acceso restringido durante determinados periodos de tiempo.
- Gestión de ancho de banda – Control del número de terminales H.323 a las que se permite el acceso simultáneo a la LAN. Utilizando la señalización H.225.0, el *gatekeeper* puede rechazar llamadas procedentes de una terminal debido a limitaciones de ancho de banda. Tal cosa puede ocurrir si el *gatekeeper* determina que no hay suficiente ancho de banda disponible en la red para soportar la llamada.
- Gestión de llamada – Por ejemplo, el *gatekeeper* puede mantener una lista de llamadas H.323 en curso. Esta información puede ser necesaria para indicar que una terminal llamada está ocupada.

Características del Controlador Multipunto

El MC proporciona funciones de control para sustentar conferencias entre tres o más puntos extremos de una conferencia multipunto. El

MC lleva a cabo el intercambio de capacidades con cada uno de los puntos extremos de una conferencia multipunto y envía un conjunto de capacidades a los puntos extremos de la conferencia indicando los modos de funcionamiento en los que pueden transmitir. El MC puede revisar el conjunto de capacidades que envía a las terminales como consecuencia de la incorporación de terminales a la conferencia o el abandono de terminales de la misma, o por otros motivos.

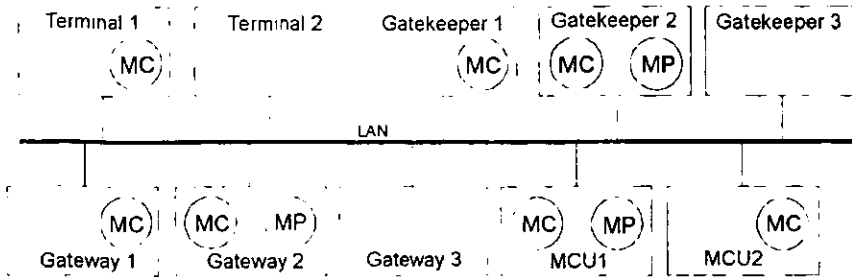
De esta manera, el MC determina el modo de comunicación seleccionado (SCM) para la conferencia. El SCM puede ser común para todos los puntos extremos de la conferencia o algunos de ellos pueden tener un SCM distinto del de los otros puntos extremos.

Como parte del establecimiento de una conferencia multipunto, un punto extremo quedará conectado a un MC en su canal de control H.245. La conexión puede producirse:

- vía una conexión explícita con una MCU;
- vía una conexión implícita al MC dentro de un *gatekeeper*;
- vía una conexión implícita al MC dentro de otra terminal o *gateway* de la conferencia multipunto;
- vía una conexión implícita a través de un *gatekeeper* a una MCU.

La elección del modo de conferencia (por ejemplo, descentralizada o centralizada) se produce después de la conexión con el MC utilizando la señalización H.245. Dicha elección puede verse limitada por la capacidad de los puntos extremos o del MC.

El MC puede estar situado dentro de un *gatekeeper*, un *gateway*, una terminal, o una MCU, como lo muestra la figura 5.4.4.



NOTA. El gateway, el gatekeeper y la MCU pueden ser un solo dispositivo

Figura 5.4.4 – Posibles ubicaciones del MC y el MP en un sistema H.323

Un MC situado dentro de una terminal no es "llamable". Puede ser incluido en la llamada para procesar la señalización H.245 de soporte de las conferencias multipunto ad hoc³⁶.

Un MC ubicado con el *gatekeeper* no es "llamable", sin embargo, una MCU ubicada con un *gatekeeper* sí lo es. Una MCU ubicada con un *gatekeeper* funciona como una MCU independiente. Un MC ubicado con un *gatekeeper* puede ser utilizado para sustentar conferencias multipunto ad hoc cuando el *gatekeeper* reciba los canales de control H.245 desde los puntos extremos. De esta manera, el *gatekeeper* puede encaminar los canales de control H.245 al MC al comienzo de la llamada o cuando la conferencia pase a ser conferencia multipunto.

Características del Procesador Multipunto

El MP recibe trenes de audio, video y/o datos de los puntos extremos que participan en una conferencia multipunto centralizada o híbrida. El MP procesa estos trenes de medios y los devuelve a los puntos extremos.

³⁶ Este tipo de conferencias son aquellas que comenzaron siendo punto a punto y que, en algún momento durante la llamada, se amplió a conferencia multipunto.

El MP puede procesar uno o más tipos trenes de medios.

Un MP que procese video deberá proporcionar conmutación o mezcla de video. La conmutación de video es el proceso de selección del video que el MP envía como salida hacia las terminales de una fuente a otra. La mezcla de video es el proceso de formateo de más de una fuente de video en el tren de video que el MP envía como salida hacia las terminales. Un ejemplo de mezcla de video es la combinación de cuatro imágenes fuente en una matriz de dos por dos en la imagen de salida de video.

Un MP que procese audio deberá preparar N salidas de audio a partir de M entradas de audio conmutando, mezclando o combinando ambas cosas. La mezcla de audio requiere la decodificación del audio de entrada en señales lineales (MIC o analógicas), efectuando una combinación lineal de las señales y recodificando el resultado en el formato de audio apropiado.

El MP no es "llamable"; la MCU que forma parte de él sí lo es. El MP termina y origina los canales de medios.

Características de la Unidad de Control Multipunto

La MCU es un punto extremo que da soporte a conferencias multipunto y deberá estar formada por un MC y puede tener MP's.

Una MCU típica, que ofrece conferencias multipunto centralizadas, consta de un MC y de un MP de audio, video y datos. Una MCU típica, que ofrece conferencias multipunto descentralizadas, consta de un MC y de un MP de datos que admite la Recomendación T.120. Se basa en el procesamiento descentralizado de audio y video.

El lado LAN de un *gateway* puede ser una MCU. Un *gatekeeper* puede

incluir también una MCU. En uno y otro caso, se trata de funciones independientes que casualmente están ubicadas en el mismo dispositivo (cubicación).

Señalización de llamada

La señalización de la llamada consiste en los mensajes y procedimientos utilizados para establecer una llamada, pedir cambios de ancho de banda de la llamada, obtener la situación de los puntos extremos de la llamada y desconectar la llamada. En la señalización de llamadas se utilizan mensajes definidos en la Recomendación H.225.0 y los procedimientos descritos en el inciso de Procedimientos de señalización de llamadas.

Direcciones

- Dirección LAN

Cada una de las entidades H.323 deberá tener por lo menos una dirección LAN. Dicha dirección identifica de manera exclusiva a la entidad H.323 en la LAN. Algunas entidades pueden compartir una dirección LAN (por ejemplo, una terminal y un MC cubicado). Esta dirección es específica del entorno LAN en el que está situado el punto extremo. Un punto extremo puede utilizar direcciones LAN diferentes para canales diferentes dentro de la misma llamada.

- Identificador TSAP

Por cada dirección LAN, cada una de las entidades H.323 puede tener varios identificadores TSAP. Los identificadores TSAP permiten la multiplexación de varios canales que comparten la misma dirección LAN.

- Dirección de alias

Un punto extremo puede tener también una o más direcciones de alias asociadas con él. Las direcciones de alias proporcionan un método alternativo de direccionamiento del punto extremo. Dichas direcciones

incluyen direcciones E.164 (número de acceso a la red, número telefónico, etc.), ID (identificadores) H.323 (nombre, dirección similar a la del correo electrónico, etc.) y cualesquiera otras direcciones definidas en la Recomendación H.225.0. Las direcciones de alias deberán ser únicas dentro de una zona.

La dirección E.164 del punto extremo llamado puede estar formada por un código de acceso opcional, seguido de la dirección E.164. El código de acceso consta de n cifras de 0 a 9, * y #.

El ID (identificador) H.323 consta de una cadena de caracteres de ISO/CEI 10646-1 definida en la Recomendación H.225.0. Puede ser un nombre de usuario, un nombre de correo electrónico u otro identificador.

Un punto extremo puede tener más de una dirección de alias (entre ellas, más de una del mismo tipo) que son convertidas a la misma dirección de transporte. Las direcciones de alias de punto extremo serán únicas dentro de una zona.

Descubrimiento del Gatekeeper

El descubrimiento del *gatekeeper* es el proceso que utiliza un punto extremo para determinar en qué *gatekeeper* se tiene que registrar. El proceso puede ser manual o automático. El punto extremo se configura con la dirección de transporte del *gatekeeper* asociado. Por ejemplo, se puede introducir en la configuración del punto extremo o en un archivo de inicialización. De esta manera, el punto extremo conoce *a priori* con qué *gatekeeper* está asociado. El punto extremo puede registrarse a continuación en ese *gatekeeper*.

El método automático permite que la asociación punto extremo-*gatekeeper* cambie con el tiempo. Es posible que el punto extremo no

conozca quién es su *gatekeeper*, o quizás necesite identificar otro *gatekeeper* debido a una falla, lo que puede hacerse mediante el descubrimiento automático. El descubrimiento automático permite una tarea administrativa menor al configurar puntos extremos individuales y permite además reemplazar un *gatekeeper* existente sin reconfigurar manualmente todos los puntos extremos afectados.

El punto extremo puede multidifundir un mensaje de petición de *gatekeeper* (GRQ³⁷) preguntando "¿Quién es mi *gatekeeper*?". El mensaje se envía a las direcciones multidifusión de descubrimiento conocidas. Uno o más *gatekeepers* puede responder con el mensaje de confirmación del *gatekeeper* (GCF³⁸) indicando "Yo puedo ser su *gatekeeper*" y devolver la dirección de transporte del canal de RAS del *gatekeeper*. Si un *gatekeeper* no desea que un punto extremo se registre en él, deberá devolver un rechazo de *gatekeeper* (GRJ³⁹) como se muestra en la figura 5.4.5. Si responde más de un *gatekeeper*, el punto extremo puede elegir el *gatekeeper* que desea utilizar. En este momento, el punto extremo conoce en qué *gatekeeper* se tiene que registrar. El punto extremo se puede ahora registrar en él.

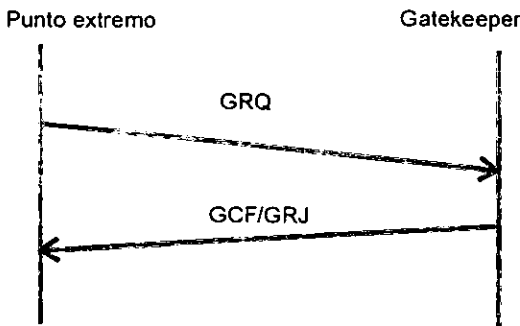


Figura 5.4.5 – Descubrimiento automático

³⁷ Del inglés Gatekeeper request

³⁸ Del inglés Gatekeeper confirmation

³⁹ Del inglés Gatekeeper reject

Si no responde ningún *gatekeeper* dentro de un plazo determinado, el punto extremo puede intentar de nuevo la GRQ. Un punto extremo no deberá enviar una GRQ durante los cinco segundos siguientes al envío de la petición previa.

Procedimientos de señalización de la llamada

La provisión de la comunicación se efectúa siguiendo los pasos que a continuación se indican:

- Fase A: Establecimiento de la llamada.
- Fase B: Comunicación inicial e intercambio de capacidad.
- Fase C: Establecimiento de comunicación audiovisual.
- Fase D: Servicios de la llamada.
- Fase E: Terminación de la llamada.

Fase A – Establecimiento de la llamada

El establecimiento de la llamada se efectúa utilizando los mensajes de control de llamada definidos en la Recomendación H.225.0, de acuerdo con los procedimientos de control de llamada definidos más abajo. Las peticiones de reserva de ancho de banda deberán efectuarse lo antes posible.

- Establecimiento de llamada básica – Ninguno de los puntos extremos está registrado

En la figura 5.4.6 se muestra un ejemplo en el que ninguno de los puntos extremos está registrado en un *gatekeeper*. Los dos puntos extremos se comunican directamente. El punto extremo 1 (punto extremo "llamante") envía el mensaje Establecimiento(1) al identificador TSAP⁴⁰ de canal de señalización de llamada conocido del punto extremo 2. El punto extremo 2 responde con el mensaje

⁴⁰ Punto de acceso al servicio de capa de transporte (Del inglés Transport layer service access point)

Conexión(4) que contiene una dirección de transporte de canal de control H.245 para su utilización en la señalización H.245.

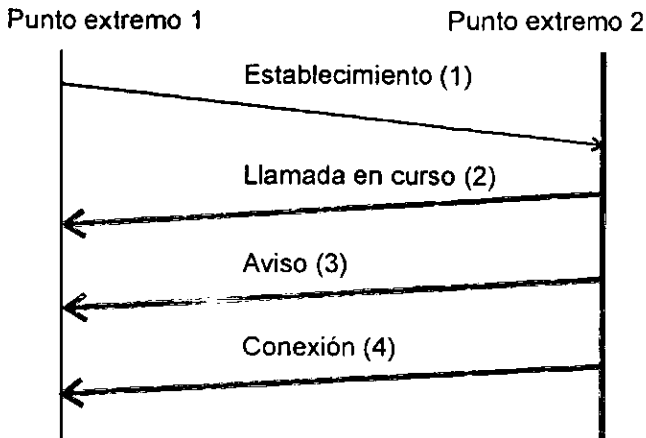


Figura 5.4.6 – Establecimiento de llamada básica

- Ambos puntos extremos registrados en el mismo *gatekeeper*
En el ejemplo mostrado en la figura 5.4.7, ambos puntos extremos están registrados en el mismo *gatekeeper* y el *gatekeeper* ha optado por señalización de llamada directa. El punto extremo 1 (punto extremo "llamante") inicia el intercambio ARQ(1)/ACF(2) con ese *gatekeeper*. El *gatekeeper* devolverá la dirección de transporte de canal de señalización de llamada del punto extremo 2 (punto extremo llamado) en el ACF. El punto extremo 1 envía entonces el mensaje Establecimiento (3) al punto extremo 2 utilizando esa dirección de transporte. Si el punto extremo 2 desea aceptar la llamada, inicia un intercambio ARQ(5)/ACF(6) con el *gatekeeper*. Es posible que el punto extremo 2 reciba un ARJ (6) en cuyo caso envía el mensaje Liberación completa al punto extremo 1. El punto extremo 2 responde con el mensaje Conexión (4) que contiene una dirección de transporte de canal de control H.245 para su utilización en la señalización H.245.

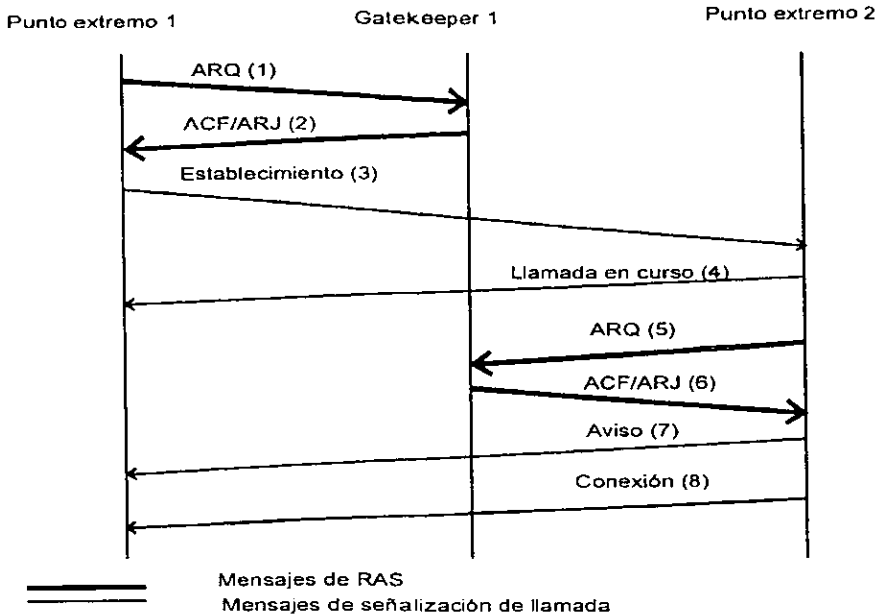


Figura 5.4.7 – Señalización de llamada directa

En la figura 5.4.8, ambos puntos extremos están registrados en el mismo *gatekeeper* y el *gatekeeper* ha optado por encaminar la señalización de la llamada. El punto extremo 1 (punto extremo "llamante") inicia el intercambio ARQ(1)/ACF(2) con el *gatekeeper*. El *gatekeeper* devolverá una dirección de transporte de canal de señalización de llamada de él mismo en el ACF. El punto extremo 1 envía entonces el mensaje Establecimiento(3) utilizando esa dirección de transporte. El *gatekeeper* envía a continuación el mensaje Establecimiento(4) al punto extremo 2. Si el punto extremo 2 desea aceptar la llamada, inicia un intercambio ARQ(6)/ACF(7) con el *gatekeeper*. Es posible que el punto extremo 2 reciba un ARJ(7), en cuyo caso envía el mensaje Liberación completa al *gatekeeper*. El punto extremo 2 responde con el mensaje Conexión(9) que contiene una dirección de transporte de canal de

control H.245 para su utilización en la señalización H.245. El *gatekeeper* envía al punto extremo 1 el mensaje Conexión(10) que puede contener la dirección de transporte de canal de control H.245 del punto extremo 2 o una dirección de transporte de canal de control H.245 de *gatekeeper* (MC), dependiendo de si el *gatekeeper* decide encaminar o no el canal de control H.245.

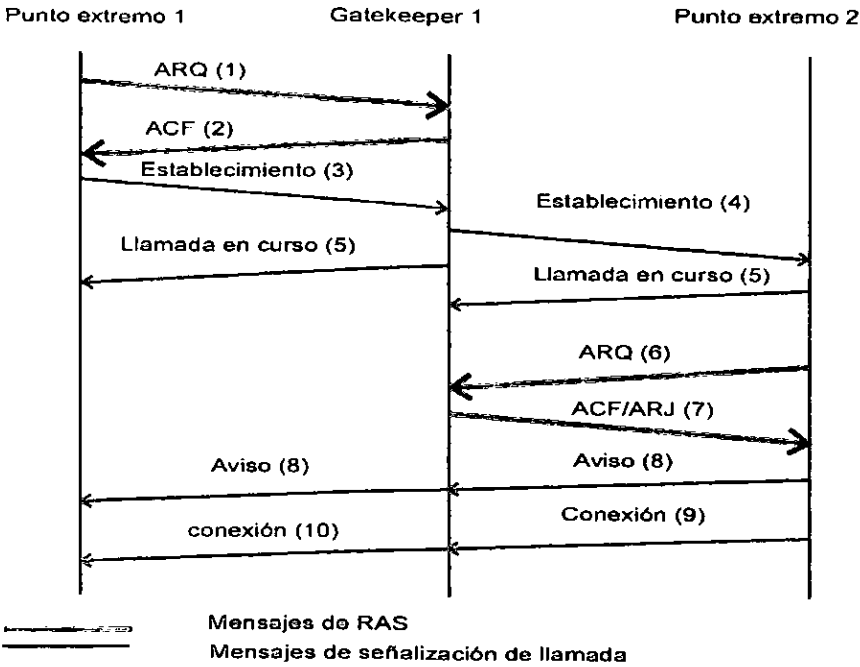


Figura 5.4.8 – Señalización de llamada encaminada por el gatekeeper

Existen aún más combinaciones que no se verán a detalle, pero que son las siguientes:

- Sólo el punto extremo "llamante" tiene *gatekeeper*
- Sólo el punto extremo llamado tiene *gatekeeper*
- Ambos puntos extremos registrados en *gatekeepers* diferentes

Fase B – Comunicación inicial e intercambio de capacidad

Una vez que ambos lados han intercambiado los mensajes de establecimiento de llamada de la fase A, los puntos extremos establecerán el canal de control H.245. Se utilizan los procedimientos de la Recomendación H.245 en el canal de control H.245 para el intercambio de capacidad y la apertura de canales de medios.

El canal de control H.245 puede ser establecido por el punto extremo llamado al recibir el mensaje Establecimiento, y por el punto extremo "llamante" al recibir Aviso o Llamada en curso. En el caso de que no llegue un mensaje Conexión, o un punto extremo envíe Liberación completa, el canal de control H.245 será cerrado.

Fase C – Establecimiento de comunicación audiovisual

Después del intercambio de capacidades y la determinación de principal/subordinado, se utilizarán los procedimientos de la Recomendación H.245 para abrir canales lógicos para los diversos trenes de información. Los trenes de audio y vídeo, que se transmiten por los canales lógicos establecidos según la Recomendación H.245, se transportan en identificadores TSAP dinámicos utilizando un protocolo no fiable⁴¹. Las comunicaciones de datos, que se transmiten por los canales lógicos establecidos según la Recomendación H.245,

⁴¹ véase la Recomendación H.225.0

se transportan utilizando un protocolo fiable⁴².

El acuse de apertura de canal lógico⁴³ devuelve la dirección de transporte que el punto extremo receptor ha asignado a ese canal lógico. El canal transmisor deberá enviar a continuación el tren de información asociado con el canal lógico a esa dirección de transporte.

Después de la apertura de canales lógicos para audio y vídeo, se enviará (por el transmisor), un mensaje indicación de asimetría máxima H.225.0⁴⁴ para cada par de audio y vídeo asociado.

Intercambio de vídeo por mutuo acuerdo

La indicación de vídeo preparado para activación se define en la Recomendación H.245. Su empleo es opcional, pero cuando se utilice, deberá seguirse el siguiente procedimiento.

Se fija el punto extremo 1 de tal manera que no se transmita vídeo a menos que, y hasta que, el punto extremo 2 indique también que está preparado para transmitir vídeo. El punto extremo 1 enviará la indicación *videoIndicateReadyToActivate* cuando se haya completado el intercambio de capacidad inicial, pero no transmitirá una señal de vídeo hasta que haya recibido *videoIndicateReadyToActivate* o vídeo entrante procedente del punto extremo 2.

Fase D – Servicios de la llamada

- Cambios de ancho de banda

El ancho de banda de la llamada la establece y aprueba inicialmente el *gatekeeper*, durante el intercambio de admisiones. Un punto

⁴² véase la Recomendación H.225.0

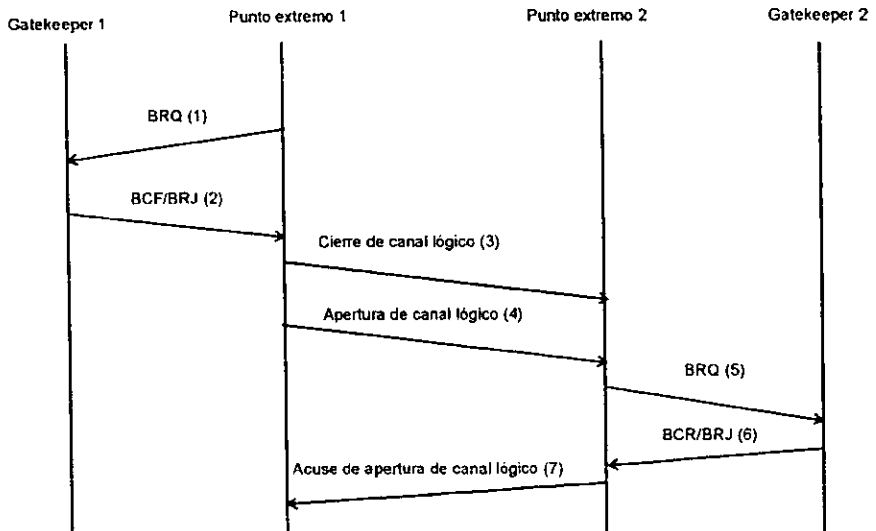
⁴³ *openLogicalChannelAck*

⁴⁴ *h2250MaximumSkewIndication*

extremo deberá asegurar que la suma correspondiente a todos los canales transmitidos y recibidos de audio y de vídeo se halla dentro de ese ancho de banda. Los canales de datos y de control no se incluyen en ese límite.

En cualquier momento durante una conferencia, los puntos extremos o el *gatekeeper* pueden solicitar un aumento o una disminución del ancho de banda de la llamada. Se recomienda pedir un cambio de ancho de banda cuando un punto extremo utilice un ancho de banda reducido durante un período de tiempo prolongado, liberando así ancho de banda para otras llamadas.

El procedimiento de petición de cambio de ancho de banda se muestra en la figura 5.4.9.



NOTA - El gatekeeper 1 y el gatekeeper 2 pueden ser el mismo gatekeeper.

Figura 5.4.9 – Petición de cambio de ancho de banda

Fase E – Terminación de la llamada

Cualquiera de los puntos extremos puede terminar una llamada mediante el procedimiento que se indica a continuación:

- 1) Interrumpir la transmisión de vídeo al final de una imagen completa y a continuación cerrar todos los canales lógicos de vídeo.
- 2) Interrumpir la transmisión de datos y a continuación cerrar todos los canales lógicos de datos.
- 3) Interrumpir la transmisión de audio y a continuación cerrar todos los canales lógicos de audio.
- 4) Transmitir el mensaje instrucción finalizar sesión de la Recomendación H.245 por el canal de control H.245, indicando al extremo distante que desea desconectarse de la llamada.
- 5) Esperar recibir el mensaje instrucción finalizar sesión del otro punto extremo y cerrar el canal de control H.245.
- 6) Si el canal de señalización de llamada está abierto, enviar un mensaje Liberación completa y se cerrará el canal.

La terminación de una llamada puede no terminar una conferencia; una conferencia puede ser terminada explícitamente utilizando un mensaje H.245 (abandono de conferencia). En este caso, los puntos extremos esperarán que el MC termine la llamada.

- Liberación de la llamada sin un gatekeeper

En las redes que no contienen un *gatekeeper*, después de los pasos 1) a 6) anteriores se termina la llamada. No se requiere ninguna acción posterior.

- Liberación de la llamada con un *gatekeeper*

En las redes que contienen un *gatekeeper*, no es preciso que el *gatekeeper* esté al corriente de la liberación de ancho de banda. La figura 5.4.10 muestra el modelo de llamada directa. Después de ejecutar los pasos 1) a 6), cada punto extremo transmitirá un mensaje de petición de desligamiento DRQ⁴⁵(3) de la Recomendación H.225.0 a su *gatekeeper*. El *gatekeeper* responderá con un mensaje de confirmación de desligamiento DCF⁴⁶(4). Después de enviar el mensaje DRQ, los puntos extremos no enviarán más mensajes IRR⁴⁷ no solicitados al *gatekeeper*. En este punto la llamada está terminada.

⁴⁵ Del inglés Disengage request

⁴⁶ Del inglés Disengage confirmation.

⁴⁷ Respuesta a petición de información (Del inglés Information request response)

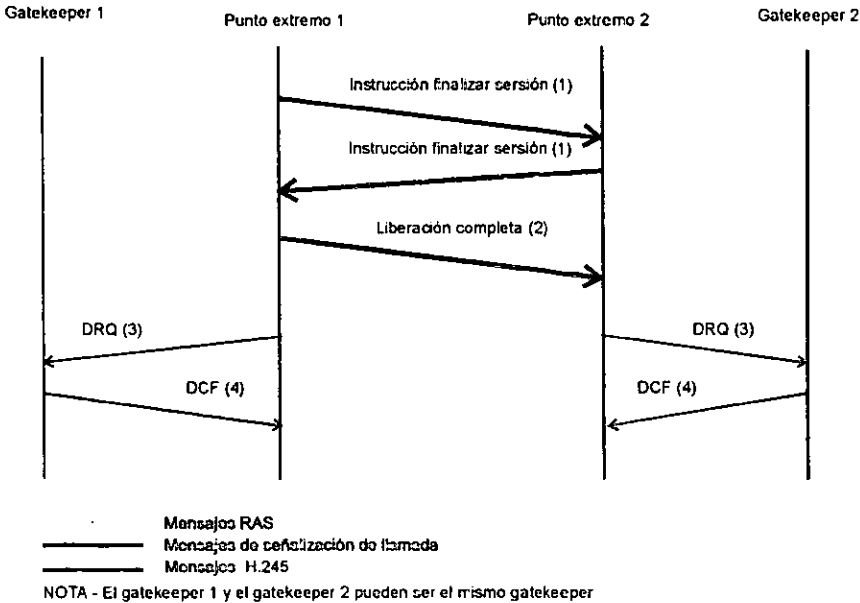


Figura 5.4.10 – Liberación de llamada iniciada por punto extremo

- Liberación de la llamada por el *gatekeeper*

El *gatekeeper* puede terminar cualquier conferencia enviando un DRQ a un punto extremo (figura 5.4.11). El punto extremo seguirá inmediatamente los pasos 1) a 6) anteriores y responderá a continuación al *gatekeeper* con un DCF.

Si la conferencia es una conferencia multipunto, el *gatekeeper* debe enviar un DRQ a cada punto extremo de la conferencia, para cerrar toda la conferencia.

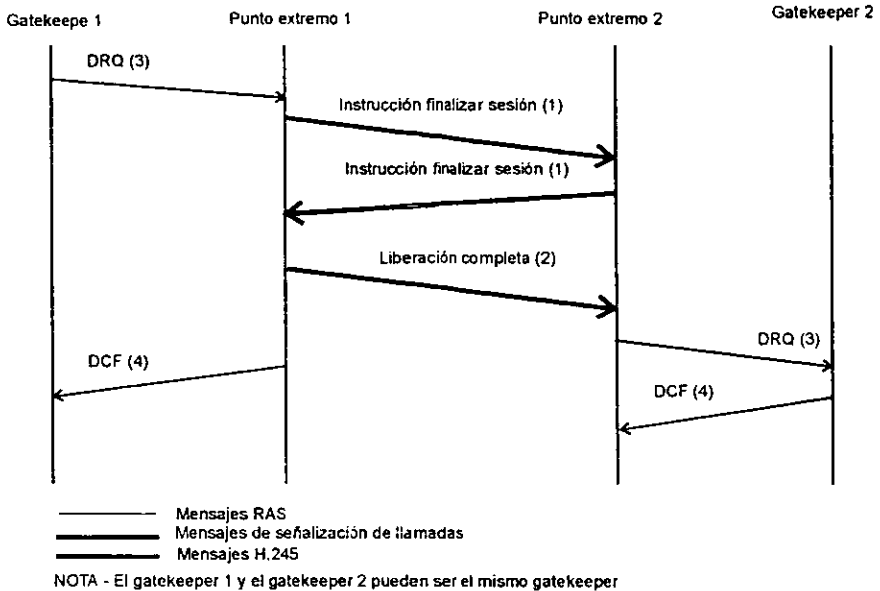


Figura 5.4.11 – Liberación de llamada iniciada por el gatekeeper

Interfuncionamiento con terminales de otros tipos

El interfuncionamiento con otras terminales se llevará a cabo a través del gateway.

- Terminales sólo vocales

El interfuncionamiento con terminales sólo vocales (telefonía) por la RDSI o la RTGC se puede efectuar de las dos maneras siguientes:

- 1) utilizando un gateway telefónico RDSI-H.323;
- 2) utilizando un gateway telefónico RTGC-H.323.

- Terminales de videotelefonía en la RDSI (Recomendación H.320)

El interfuncionamiento con terminales videotelefónicas en la RDSI (Recomendación H.320) se puede efectuar utilizando un gateway

H.323-H.320.

- Terminales videotelefónicas en la RTGC (Recomendación H.324)

El interfuncionamiento con terminales videotelefónicas en la RTGC (Recomendación H.324) se puede efectuar de las dos maneras siguientes:

- 1) Utilizando un *gateway* H.323-H.324.
- 2) Utilizando un *gateway* H.323-H.320 en el supuesto de que exista una unidad de interfuncionamiento RDSI/RTGC en la red.

- Terminales videotelefónicas en redes ATM (Recomendación H.321)

El interfuncionamiento con terminales videtelefónicas en redes ATM (Recomendación H.321) se puede efectuar de las dos maneras siguientes:

- 1) Utilizando un *gateway* H.323-H.321.
- 2) Utilizando un *gateway* H.323-H.320 en el supuesto de que exista una unidad de interfuncionamiento RDSI/ATM de la Recomendación I.580 en la red.

- Terminales videotelefónicas en las LAN con calidad de servicio garantizada (Recomendación H.322)

El interfuncionamiento con terminales videotelefónicas en LAN con calidad de servicio garantizada (Recomendación H.322) se puede efectuar utilizando un *gateway* H.323-H.320 en el supuesto de que exista un *gateway* LAN-RDSI con calidad de servicio garantizada (GQoS) en la red.

- Terminales de señales vocales y datos simultáneos en la RTGC (Recomendación V.70)

El interfuncionamiento con terminales de señales vocales y datos simultáneos en la RTGC (Recomendación V.70) se puede utilizando un gateway H.323-V.70.

- Terminales de la Recomendación T.120 en la LAN

Una terminal H.323 que tenga la capacidad de la Recomendación T.120 deberá poder ser configurada como una terminal sólo Recomendación T.120 que escucha y transmite en el identificador TSAP conocido T.120 normalizado. De esta manera, la terminal H.323 con capacidad de T.120 podrá participar en conferencias solamente conformes a la Recomendación T.120.

Sólo una terminal conforme con la Recomendación T.120 en la LAN podrá participar en la porción T.120 de las conferencias.

Bibliografía:

1. <http://www.itu.int/itudoc/itu-t/rec/h>.
2. Recomendaciones de la Unión Internacional de Telecomunicaciones. Sector de Normalización de las telecomunicaciones de la UIT (UIT-T) Serie H: Transmisión de señales no telefónicas.

CAPÍTULO 6

CODIFICACIÓN DE VIDEO CON VLSI (*Very Large Scale Integration*) [Pirsch, 1995]

La gran cantidad de aplicaciones consideradas de los servicios de vídeo comunicación; explorados de manera general en el presente trabajo, tienden hacia un equipo de codificación con bajo costo de manufactura y tamaño reducido. Los esquemas de codificación recomendados por los grupos internacionales de estandarización son muy sofisticados para alcanzar una alta reducción de la tasa de bits, bajo la restricción de la más alta calidad de imagen posible. Tales esquemas de codificación pueden resultar en sistemas de hardware de alta complejidad. Así, el costo efectivo de la implementación de dichos sistemas de alta complejidad dependen de la VLSI (*muy alta escala de integración*). Los recientes avances en la tecnología VLSI auxilian a los problemas de *hardware*, pero las grandes demandas de vídeo codificación requieren aproximaciones arquitectónicas especialmente adaptadas a los esquemas de procesamiento de la señal de vídeo.

6.1 DE ALGORITMOS A ARQUITECTURAS VLSI

Los costos de manufactura son dominados por el número de circuitos integrados (chips), el empaque del chip y el área de silicio por chip. De modo que la producción de chips de áreas muy grandes no es muy económica. Los sistemas con alta complejidad requieren en general de varios chips para su implementación. El costo por empaque del chip está relacionado con el número de capas; por esta razón, la división del sistema en chips debe considerar las interconexiones. Mediante la aplicación de tecnologías avanzadas de semiconductores los sistemas de alta complejidad pueden ser implementados con un moderado número de chips, siendo la finalidad lograr una implementación *VLSI* con la más pequeña área de silicio. Como se lista en la tabla 6.1 la tasa de transmisión de las señales de vídeo tienen un amplio rango de acuerdo al tipo de aplicación. Las estructuras de hardware para fuentes a tasas bajas como el formato QCIF con 9 Mbps, son diferentes de aquellas para tasas muy altas tales como HDTV con 1.9 Gbps.

<i>Formato de Video</i>	<i>Tamaño de Trama de Luminancia</i>	<i>Trama (Tasa)</i>	<i>Fuente de video (Tasa)</i>	<i>Almacenamiento de trama</i>
HDTV	1920 * 1250	50 Hz	1.9 Gbps	19 Mb
CCIR	720 * 576	25 Hz	166 Mbps	3.3 Mb
CIF	352 * 288	30 Hz	36 Mbps	0.8 Mb
QCIF	172 * 144	30 Hz	9 Mbps	0.2 Mb

Tabla 6.1 Formatos de video digital.

El área de silicio requerida para la implementación *VLSI* de los algoritmos está relacionada a los requerimientos de las fuentes; tales como las compuertas lógicas, la memoria y las interconexiones entre

los módulos. El total, por lógica depende de la concurrencia de operaciones. La *figura de mérito* para la concurrencia de operaciones requerida se puede obtener como

$$N_{con,op} = R_s \cdot N_{op,pel} \cdot T_{op} \quad (1)$$

donde:

R_s	tasa de la fuente en pels por unidad de tiempo.
$N_{op,pel}$	número de operaciones por pel.
T_{op}	promedio de tiempo por ejecución de operación.

El número de operaciones por pel es un valor promedio derivado de la cuenta total de operaciones requeridas para la ejecución del esquema de codificación. En el caso de esquemas de codificación híbridos (como los aplicados en MPEG) este intervalo es un *macro block*. Casi todas las tareas del esquema de codificación son definidas para un *macro block* de 16 x 16 pels de luminancia y los pels asociados de crominancia. Por esta razón la cuenta debe ser realizada sobre un *macro block* (como se refiere en el capítulo 1 parte 2, referente a las técnicas de codificación de vídeo). El número de operaciones depende específicamente del algoritmo. Por ejemplo block matching basado en búsqueda exhaustiva requiere de muchas más operaciones que en un esquema de búsqueda estratégica. El número de operaciones de un esquema completo de codificación (codificador y decodificador) es del orden de 200 cuando se aplica block matching con búsqueda 2D log y un algoritmo rápido DCT (transformada coseno discreta). Con las tecnologías actuales, es posible que T_{op} esté en el orden de 20 ns. De aquí se desprende que el número de operaciones concurrentes van de 5 a 1000, dependiendo del formato de vídeo.

Las interconexiones requeridas entre la parte operativa y la memoria dependen de la tasa de acceso. Considerando una gran memoria externa para el almacenamiento de datos de video y los resultados inmediatos, el *número de líneas en paralelo del bus* para conectar la parte operativa se convierten en aproximadamente

$$N_{\text{bus}} = R_s \cdot N_{\text{acc, pel}} \cdot T_{\text{acc}} \quad (2)$$

donde:

$N_{\text{acc/pel}}$ número medio de accesos por pel.
 T_{acc} tiempo de acceso a la memoria.

Como en las expresiones anteriores 1 y 2 $N_{\text{acc/pel}}$ es más grande que $N_{\text{op/pel}}$. Y por simplicidad se asume que $N_{\text{acc, pel}} \cdot T_{\text{acc}}$ está en el mismo orden que $N_{\text{op, pel}} \cdot T_{\text{op}}$. De aquí se establece cómo a la figura de mérito la determinó la concurrencia de operaciones; y así el número de las líneas de bus necesarias se vuelve muy grande. Tomando en consideración que la tasa de acceso está influenciada principalmente por múltiples accesos desde la fuente de datos de la imagen y los resultados inmediatos, la tasa de acceso a una memoria externa puede ser reducida esencialmente asignando una memoria local a la parte operativa. El tamaño de la memoria local dependerá de la estructura de acceso específica del algoritmo de codificación.

La concurrencia de operaciones se puede lograr mediante arquitecturas en paralelo o *pipelining* (canalizado). Dado que los algoritmos se pueden especificar de manera jerárquica, también el procesamiento en paralelo o *pipelining* pueden estar definidos jerárquicamente. Los algoritmos pueden ser especificados por tareas como se muestra en la figura 6.1; por consiguiente, las tareas requeridas son block matching, la determinación de predicción de

error, DCT (transformada coseno discreta), cuantización, codificación en longitud variable y las operaciones inversas asociadas¹.

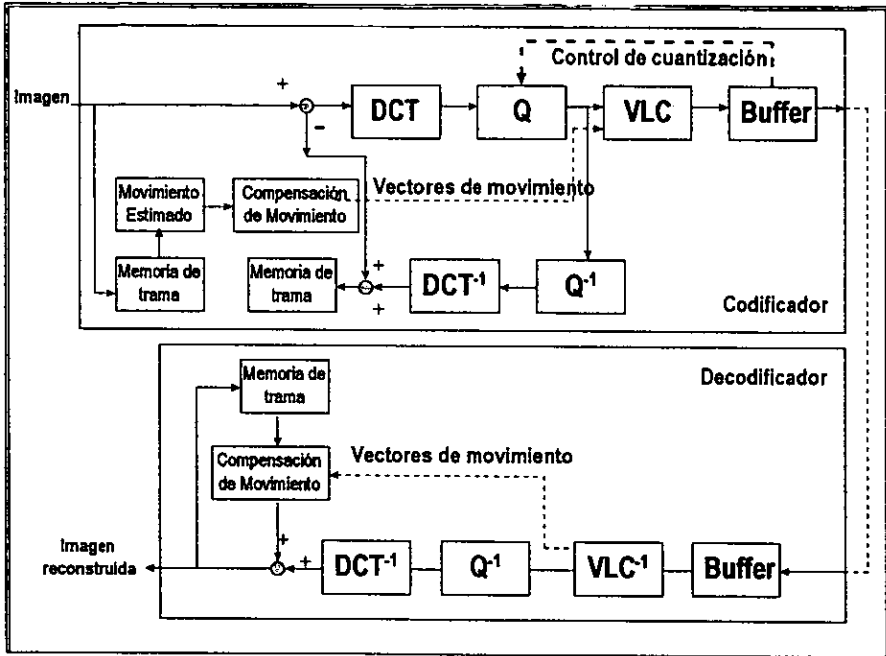


Figura 6.1 Codificador-decodificador MPEG.

Las tareas tienen que ser ejecutadas en una secuencia específica. Asignando cada tarea directamente a un procesador especial y considerando una secuencia específica de procesamiento, así la cascada de procesadores estarán derivando al que pueda operar fácilmente en *pipeline* colocando entre los procesadores una memoria del tamaño apropiado. Este tipo de mapeo de tareas hacia los

¹ La Figura 6.1 especifica la estructura básica del H.261, codificador-decodificador MPEG-1 y MPEG-2.

procesadores pueden referirse como una *distribución de tareas*, asignando la arquitectura de un procesador individualmente a una tarea especial, y derivando otras funciones a otras arquitecturas específicas. En este caso el hardware del procesador puede ser utilizado muy eficientemente adaptando la arquitectura del circuito al algoritmo específico de la tarea. En particular, tareas como la DCT, la DCT inversa y block matching tienen altos requerimientos computacionales.

Esos requerimientos computacionales pueden estar definidos en términos de concurrencia de operaciones por la expresión (1). Esas tareas serán muy completas con una secuencia bien predefinida de operaciones y accesos de datos. Ahora bien, mediante refinamientos al algoritmo bajo las operaciones básicas, las técnicas bien conocidas en la literatura se pueden aplicar para mapear algoritmos completos a las arquitecturas, con uso extensivo de procesamiento paralelo y dedicado.

Por la planeación de las tareas en la dirección de procesamiento (ver figura 6.2), cada procesador debe llevar a cabo todas las tareas definidas en la secuencia especificada, y de acuerdo a esta secuencia, el procesador debe realizar diferentes operaciones.

Debido a la gran variedad de operaciones y accesos de datos el procesador programable debe ser el apropiado. Como los resultados de una tarea son requeridos para los procedimientos de las demás tareas, se requiere una memoria local para el almacenamiento de los resultados inmediatos. La concurrencia requerida puede lograrse mediante la operación de procesadores en paralelo con lo cual una subsección de la imagen es asignada a cada procesador. Debido a

ésto, la finalidad es explotar los segmentos de imagen que puedan ser procesados casi independientemente.

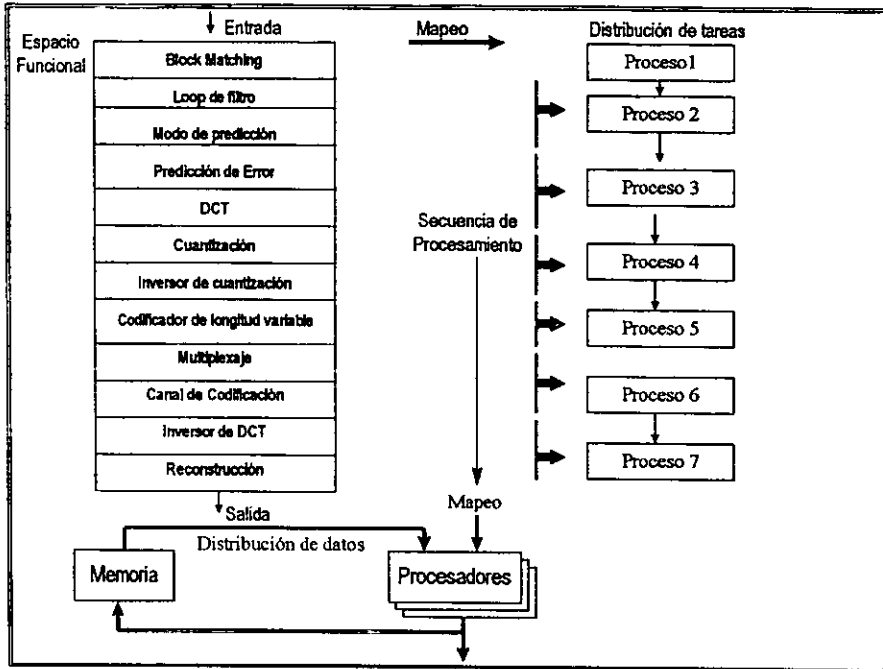


Figura 6.2 Espacio funcional del decodificador híbrido y mapeo a sistemas multiprocesadores, explotando la distribución de datos y distribución de tareas.

El segmento más pequeño para procesamiento independiente es, en el caso de la codificación MPEG un *macro block*. Este tipo de mapeo de tareas hacia los procesadores es descrito como una distribución de datos y es la base de los microprocesadores programables homogéneos. Para que aumente la eficiencia del silicio de los microprocesadores programables, la parte operacional debe agruparse en tareas. Esto puede lograrse mediante partes operacionales especiales o agregando coprocesadores para funciones específicas, lo

cual puede resultar finalmente en multiprocesadores programables heterogéneos.

6.2 ESTIMACIÓN DE ARQUITECTURAS ALTERNATIVAS

Aplicando las estrategias de mapeo discutidas, éstas llevan a una amplia variedad de soluciones arquitectónicas para la implementación de esquemas de compresión de video. Para hacer una comparación de esas alternativas de arquitecturas, debe de introducirse una estimación de medición considerando la eficiencia arquitectónica, la cual, en general puede ser definida como una razón de *rendimiento* y costo. Típicamente, el rendimiento de una arquitectura puede ser igualado con el recíproco del tiempo de procesamiento efectivo alcanzado para una muestra ($1/T_P$). La determinación del costo para una arquitectura específica es más problemática, puesto que esta determinación del costo se halla influenciada por un amplio rango de parámetros interactuando, como son el área de silicio, estilo de diseño, complejidad arquitectónica, proceso del semiconductor, total de pines, etc. Para una primera aproximación simplificada, el costo puede ser expresado por el área de silicio requerida para la implementación de una arquitectura específica. Esto lleva al producto AT para la estimación de la arquitectura. En este, caso la *eficiencia* "E" está definida por

$$E = \frac{1}{A_{Si} \cdot T_P} \quad (3)$$

La *tasa de transferencia* R_T de una arquitectura es proporcional al inverso del tiempo de procesamiento efectivo para una muestra T_P

$$R_T \sim \frac{1}{T_P} \quad (4)$$

Para una arquitectura con una eficiencia especificada fija, de la expresión (3) se obtiene que el *área de silicio* es proporcional a la tasa de transferencia.

$$A_{Si} = \alpha_T \cdot RT \quad (5)$$

La relación anterior es para partes operacionales donde el incremento de transferencia requiere un incremento de concurrencia, mediante la implementación en paralelo de unidades de procesamiento básico.

Dado que el área de silicio requerida y el tiempo de procesamiento para la implementación de una aplicación específica depende de la tecnología del semiconductor aplicada, una estimación arquitectónica realista debe considerar la ganancia proporcionada por los progresos en la tecnología del semiconductor. Una manera factible de lograr una estimación realista, es la normalización de los parámetros de la arquitectura con alguna referencia tecnológica.

6.3 FUNCIONES ESPECÍFICAS DE LAS ARQUITECTURAS.

Considerando las tecnologías disponibles y los requerimientos computacionales necesarios para los vídeo codificadores y decodificadores, el uso de implementaciones dedicadas (o funciones específicas), es frecuentemente obligatorio por los altos volúmenes de consumo de productos.

La optimización del área de silicio que contienen las arquitecturas dedicadas, comparadas con las arquitecturas programables llevan a un más bajo costo de producción.

A continuación se describen brevemente dos ejemplos de arquitecturas específicas: el cálculo de la (DCT) transformada coseno discreta y la estimación de movimiento.

A. La Transformada Coseno Discreta (DCT).

(Descrita en el capítulo 2 parte 2.1.2) Una importante propiedad para la implementación VLSI de la transformada 2D-DCT es su separabilidad, ya que es posible calcularla ejecutando varias DCT's unidimensionales sucesivamente.

Como un ejemplo, la implementación de una DCT se muestra en la figura 6.3. Esta arquitectura se basa en dos arreglos de procesamiento unidimensionales.

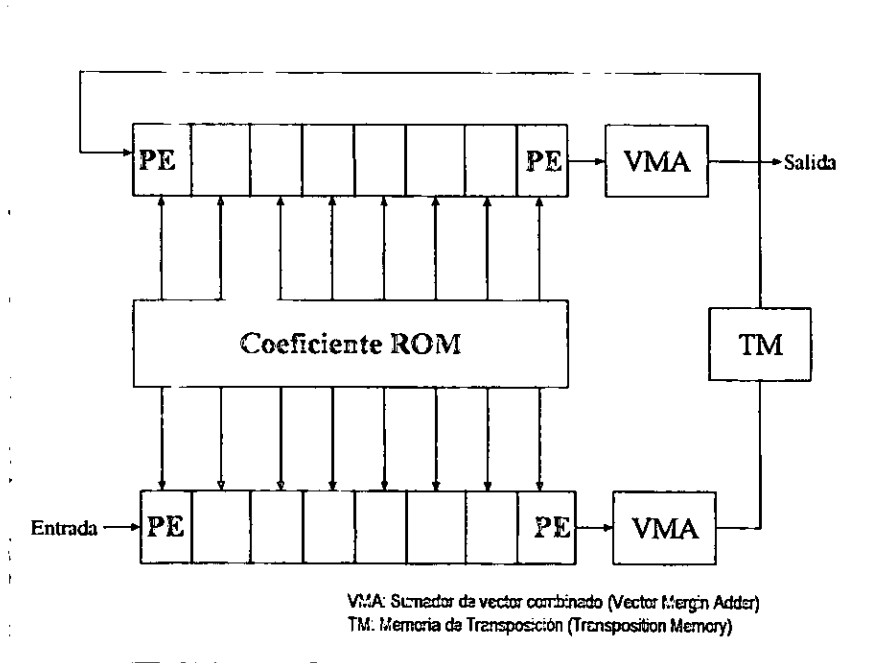


Figura 6.3 Implementación de la DCT

Dado que esta arquitectura está basada en una implementación directa ó dedicada multiplicador/acumulador con técnica de transportar y guardar, los vectores convergentes son colocados en la salida a cada lado del arreglo y los resultados de la DCT-1D tienen que ser reordenados para el segundo estado de la DCT-1D. Para este propósito se utiliza una memoria de transposición, y como ambos arreglos de procesadores de ID requieren idénticos coeficientes DCT, esos coeficientes son almacenados en una ROM (memoria de sólo lectura) común a ambos arreglos.

Se han propuesto muchos otros algoritmos para calcular la DCT como los algoritmos de vector radial, Winograd-DCT, rotaciones, algoritmos de factor primo, estructuras convolucionales, transformadas polinomiales, etc., saliendo todos éstos del alcance del presente trabajo.

Comparación de Arquitecturas

Los algoritmos mencionados anteriormente difieren en varios aspectos: los requerimientos de cálculo (número de sumas y multiplicaciones) y los requerimientos de almacenamiento (tamaño de la ROM para la aritmética distribuida, tamaño de la RAM para transposición de memoria). Desafortunadamente, existen algunos factores que empañan el panorama:

- El esfuerzo arquitectónico: para cada implementación de esos algoritmos, el diseñador puede usar un número diferente de aproximaciones arquitectónicas (*pipelining*, cálculo de bit serial, *hardware* de multiplexaje, compartición de bus, etc.).
- La restricción de transferencia: dependiendo de la fuente de transferencia, algunas arquitecturas pueden ser más interesantes

que otras. Por ejemplo citaremos que es más difícil acelerar los diseños basados en ROM pues la transferencia de la aritmética distribuida está limitada por el tiempo de acceso a la ROM.

- La regularidad del algoritmo: las arquitecturas irregulares tienden a generar una sobrecarga de cableado. A éste respecto, la aritmética distribuida y las implementaciones directas de producto escalar son generalmente más regulares que las arquitecturas basadas en flujo gráfico.
- La tecnología: dependiendo del costo relativo del almacenamiento (ROM) y las operaciones aritméticas, la selección entre arquitecturas basadas en flujo gráfico o producto escalar, y las arquitecturas basadas en aritmética se pueden modificar.

Debido a la complejidad del espacio de diseño, no pueden hacerse estimaciones generales.

B. Block Matching.

(descrita en el capítulo 2 parte 2.3) La estimación de movimiento se ha presentado como una técnica para el mejoramiento de la eficiencia de la codificación. Varias técnicas para la estimación de movimiento se han presentado en el pasado. Actualmente, la más importante técnica para estimación de movimiento es block matching, la cual está basada en una comparación de bloques entre el cuadro actual y uno de referencia. Esto puede realizarse por medio de una búsqueda completa (exhaustiva) dentro de una ventana de búsqueda; aún cuando se hayan reportado algunas otras aproximaciones, con el fin de reducir los requerimientos de cálculo empleando una búsqueda "inteligente" o "direccionada".

C. Implementaciones Dedicadas del Esquema de Codificación Híbrido Completo.

Anteriormente se mencionaron diferentes arquitecturas e implementaciones de funciones básicas para esquemas de vídeo codificación y decodificación (DCT, estimación de movimiento). Aún cuando se ha propuesto la implementación de otras funciones básicas, tales como la codificación/decodificación o como los códigos de longitud variable. En la primera generación de circuitos, cada una de esas funciones básicas se implementaron en un solo chip y así, fuer necesario un grupo de chips para crear un sistema de codificación y decodificación MPEG.

Un ejemplo de una implementación dedicada a un sistema híbrido completo se presenta en la figura 6.4. La arquitectura apunta hacia un solo chip vídeo codificador MPEG-1.

El chip consiste de cuatro módulos principales: una unidad decodificadora para decodificar el código de longitud variable, unidad IDCT, unidad de compensación de movimiento, y una unidad de instrucción la cual controla las unidades funcionales y la trama de memoria externa DRAM. Para intercalar un código de interfase al bus, fueron integradas en el procesador una interfase de vídeo al bus y una interfase al bus de memoria.

Un solo chip proporciona suficiente potencia de procesamiento para la decodificación en tiempo real de señales SIF (352 x 240 pels, a una tasa de cuadro de 30 Hz), en un canal de tasa de datos máxima hasta de 6 Mbps. La resolución CCIR se puede lograr utilizando una unidad de interpolación en chip.

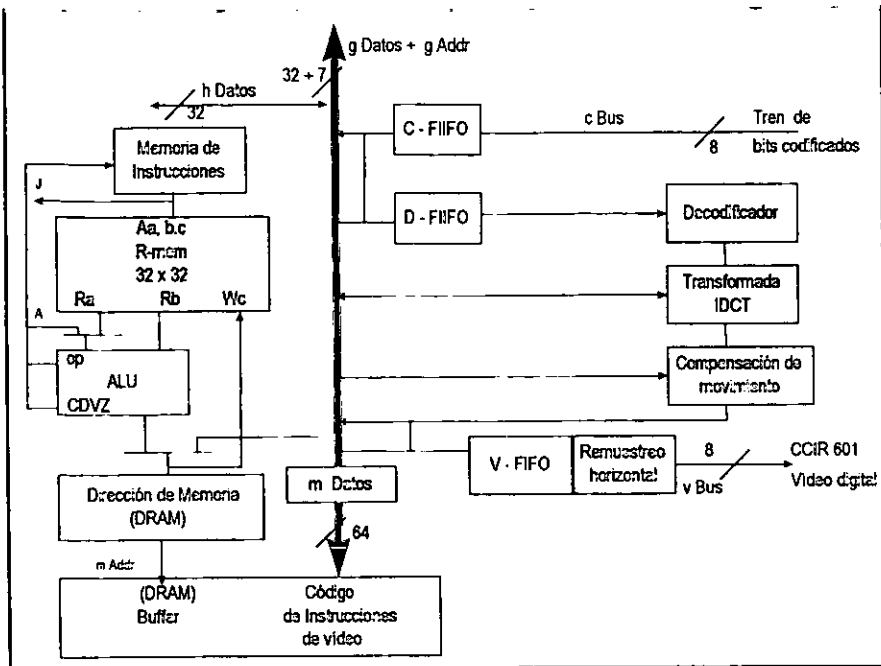


Figura 6.4 Decodificador MPEG.

6.4 SISTEMA MULTIPROCESADOR PROGRAMABLE

Las arquitecturas programables permiten el procesamiento de diferentes tareas o programas bajo el control de software. La principal ventaja de la arquitectura programable es la flexibilidad. Se pueden manejar cambios de requerimientos de arquitectura debido a cambios de algoritmos o una extensión del campo de aplicación, por cambios en el respectivo software. Como consecuencia se puede evitar un rediseño costoso de hardware. Por otra parte las arquitecturas programables cubren un amplio rango de aplicaciones, y se pueden usar para aplicaciones de bajo volumen, donde el diseño para una función específica por un chip VLSI no es una solución económica.

Pero por otra parte, las arquitecturas programables requieren un alto costo de diseño y manufactura, ya que se requiere un hardware adicional para el control de programa. Además, las arquitecturas programables requieren el desarrollo de software para afrontar la aplicación y aunque algunos fabricantes proporcionan tal herramienta, incluyendo un lenguaje compilador de alto nivel, no se puede negar el costo por software desarrollado y se debe considerar para tomar la decisión de cual tipo de arquitectura, función orientada o programable, se debe usar para la aplicación específica.

El procesamiento de imágenes, especialmente para aplicaciones de codificación de vídeo, muchas veces requiere un procesamiento en tiempo real de los datos de la imagen. Para realizar esto, se han empleado estrategias de paralelización. Las dos alternativas de paralelización son distribución de datos o información y distribución de tarea, estas han sido discutidas anteriormente en la sección 6.1.

Se han propuesto varias arquitecturas programables para aplicaciones de codificación de vídeo durante los últimos años. Algunas de esas arquitecturas están diseñadas especialmente para aplicaciones de codificación de vídeo, mientras que otras apuntan a un amplio rango de aplicaciones.

El principal problema que se tiene que solucionar con una implementación de *VLSI* es el soportar el alto poder computacional requerido para la aplicación de codificación de vídeo. En seguida se mencionan tres caminos para enfrentar este problema:

1 Incremento de la Frecuencia de Reloj:

Debido a la estrecha interacción entre la frecuencia de reloj y la potencia de computación, un camino para incrementar el poder de procesamiento es incrementar la frecuencia de reloj del procesador para un "pipelining" intensivo (una categoría de técnicas que proveen procesamiento simultáneo, o paralelo, por ejemplo, mientras una instrucción está siendo ejecutada, se está decodificando la próxima instrucción). Esto es limitado por dos efectos. Primero, el incremento de la frecuencia de reloj por pipelining incrementa el estado latente del circuito (el tiempo entre la iniciación de un pedido de datos y el comienzo de la efectiva transferencia de los datos o bien el tiempo que tarda el canal en desocuparse con objeto de transferir datos). Para algoritmos que requieren datos dependientes del control de flujo este factor puede limitar tal desempeño. Adicionalmente, incrementar el poder de procesamiento aritmético lleva a un incremento de la velocidad de acceso de datos. Segundo, generalmente la velocidad de acceso de datos requerida no puede ser proporcionada por memorias externas.

Pipelining se puede referir como una implementación específica de distribución de tareas, las cuales son procesadas en paralelo en diferentes unidades y esto incrementa la velocidad de procesamiento.

Ejemplos de procesadores de video con alta frecuencia de reloj es el VSP3, debido al intensivo pipelining se puede alcanzar una frecuencia de reloj arriba de 300 Mhz, el total VSP3 consiste de 1.27 millones de transistores, tecnología BiCMOS en un 16.5 x 17 mm. Un codificador H.261 se puede implementar basado en un VSP3 (omitiendo la codificación de longitud variable).

2 Enlace o Línea de Datos en Paralelo:

Una alternativa para la distribución de tareas, es la distribución de datos a través de unidades funcionales. La aplicación de esta estrategia nos lleva a una implementación de una vía de acceso de datos en paralelo para el aumento del poder computacional, típicamente cada ruta de datos es conectada a un chip de memoria el cual proporciona el acceso a los segmentos de imagen distribuidos.

Generalmente, se pueden distinguir dos tipos de estrategia de control para las rutas de datos en paralelo. Una *MIMD (Multiple Instruction stream Multiple Data stream*, en español flujo múltiple de instrucciones, flujo múltiple de datos), arquitectura que utiliza múltiples procesadores, cada uno procesando su propio conjunto de instrucciones simultánea e independientemente de los otros. Ésta proporciona una unidad de control privado para cada ruta de datos, mientras que el *SIMD (Single Instruction stream Multiple Data stream* en español flujo de instrucciones únicas, flujo de datos múltiples), arquitectura que realiza una operación sobre múltiples conjuntos de datos, por ejemplo en un procesador de arreglos se emplea un procesador para la lógica de control y los procesadores restantes se usan como esclavos, ejecutando cada uno la misma instrucción. Este SIMD provee un controlador común único para la trayectoria de los datos en paralelo. Comparado con el SIMD, la ventaja de MIMD es que es más flexible y tiene mejor desempeño para algoritmos complejos con datos dependientes del control de flujo, pero por otra parte MIMD requiere un significativo incremento de área de silicio.

Un ejemplo de un procesador de vídeo basado en la trayectoria de datos en paralelo es el procesador IDSP que integra 910 000

transistores en 15.2x15.2 mm usando una tecnología 0.8 mm BiCMOS. Para un total CIF-H.261 codificador de video se requieren cuatro IDSP's.

3 Coprocesador:

La mayor parte de las arquitecturas programables para aplicaciones de procesamiento de video consiguen un aumento del poder de procesamiento por la adaptación de las arquitecturas a los requerimientos de los algoritmos.

Típicamente, las tareas intensivas de computación como por ejemplo DCT, block matching, o códigos de longitud variable, son candidatos para una implementación adaptada o dedicada.

Un coprocesador (procesador secundario empleado para acelerar las operaciones, llevando a cabo parte de la carga de trabajo del CPU principal) se usa a menudo para tareas específicas, por ejemplo, operación del punto flotante. La idea de la adaptación a una tarea específica, incrementando el poder computacional sin requerir un aumento del área de semiconductor, se ha aplicado para muchos diseños. La principal desventaja de esta aplicación es la limitada flexibilidad por un aumento de la adaptación.

La aplicación del concepto de coprocesador abre una gran variedad de posibles arquitecturas, algunas de ellas son, la TMS320C80, el VDSP2, el AxPe640 V, etc. Debido a los altos requerimientos computacionales de la codificación de video en tiempo real, muchas de las arquitecturas aplican un coprocesador con módulos programables flexibles en combinación con módulos que son más o menos adaptados a una tarea específica de los esquemas de codificación híbridos.

La mayoría de las arquitecturas explotan más o menos las estrategias de arquitecturas mencionadas, que llevan a un amplio rango de diferentes arquitecturas, pero en general se distinguen dos clases de arquitecturas:

1).- Arquitecturas programables flexibles

Estos procesadores proporcionan de moderada a alta flexibilidad. Estas arquitecturas están basadas en el concepto de coprocesador así como en la trayectoria de datos en paralelo y diseño de procesamiento simultáneo con alta frecuencia de reloj.

2).- Arquitecturas programables adaptadas

Estas arquitecturas logran una elevada eficiencia por la adaptación de la arquitectura a los requerimientos específicos de las aplicaciones de la codificación de vídeo. Estas arquitecturas prevén módulos dedicados para varias tareas de los esquemas de codificación híbridos, por ejemplo, estimación de movimiento o códigos de longitud variable.

PARTE III

CONCLUSIONES

CONCLUSIONES

Los sistemas de transmisión de vídeo digital pueden ser muy sencillos o muy complejos dependiendo de las características de la información que se intenta transmitir; y es por ello que se analizaron, aunque no muy profundamente algunos de los más simples, se explican de manera breve las características de un sistema de vídeo digital, así como las alternativas para la corrección de las inestabilidades de los mismos. De esto, se deduce que actualmente, prácticamente toda la información que llega a nuestras manos, llega de manera digital debido a las bondades que nos ofrece el hecho de la transmisión de información de manera digital.

La calidad de reproducción de un sistema digital de vídeo depende de la calidad de los procesos de conversión y es independiente de los medios de transmisión; lo cual no es el caso de los sistemas analógicos, es decir, con la tecnología digital se superan problemas como la atenuación de la señal, la interferencia y la degradación. Se puede asegurar que el sistema digital se acerca más a lo ideal que los sistemas analógicos y a un costo menor gracias a la tecnología de los circuitos integrados.

El vidicón cuyas principales ventajas se tienen en su constitución sencilla y tamaño pequeño como equipo portátil de campo, presenta inconvenientes tales como: poca resolución, necesidad de mayor iluminación, velocidad de respuesta lenta, sombreado desigual de la imagen y quemadura de la capa fotoconductoras debido a objetos inmóviles. Sin embargo, hasta no hace mucho tiempo la necesidad de compatibilidad y simplicidad del equipo accesorio hacía necesario el uso de vidicones a pesar de tales inconvenientes.

Un detector de C.C.D. (*Charge Coupled Device*) es un arreglo de elementos que tiene propiedades capacitivas. Un electrodo metálico está separado de un semiconductor tipo p por una capa aislante de dióxido de silicio, lo que conforma un arreglo tipo MOS (*Metal Oxide Semiconductor*). Los C.C.D.'s son muy sensibles y almacenan del 20% al 40% de todos los fotones que llegan a su superficie, además que los C.C.D.'s son sensibles a la radiación roja y a la infrarroja cercana, mientras que las placas fotográficas trabajan mejor en la región azul del espectro.

Las ventajas de los C.C.D.'s sobre otros dispositivos de detección son:

- Dispositivos muy pequeños y compactos.
- Excelente sensibilidad a un amplio intervalo de longitudes de onda.
- Buena estabilidad y linealidad.
- Compatibilidad inmediata con las computadoras. No se necesita el revelado de películas, ya que la salida del C.C.D. es una secuencia de pulsos eléctricos listos para ser transferidos a una computadora.

Las principales desventajas que los C.C.D.'s presentan ante otros detectores son:

- La imagen de sesgo.
- Diferencias físicas entre los píxeles, generando vibraciones de sensibilidad de pixel a pixel.
- Efecto de corriente oscura, menor a la vidición.
- Gran tamaño y peso de la cámara de enfriamiento del C.C.D.
- Ruido en la lectura.

Como una necesidad de la transmisión de señales digitales, para ahorrar tiempo y volumen de información en la transmisión, se desarrollaron técnicas de codificación, como la predictiva, en la que se pretende que el envío de la información sea lo más eficiente posible, evitando hasta lo posible el envío de información redundante, con lo que se optimiza el envío de información. La modulación delta es, tal vez, el más sencillo de los codificadores predictivos y probablemente una de las técnicas más antiguas, pero a la fecha se sigue usando como base para las técnicas de compresión de vídeo. Muchas de las técnicas de compresión que se usan actualmente están basadas en los conceptos básicos de la modulación delta.

Una de las técnicas para la descomposición en el dominio de la frecuencia es la DCT. La señal a menudo tiene más energía concentrada en un pequeño rango de frecuencias (típicamente las bajas frecuencias), por lo que se necesitan pocos bits para describir frecuencias sin importancia. La DCT toma bloques de una imagen generalmente de 8x8 píxeles y los convierte en una serie de coeficientes, representando diferentes componentes de frecuencia. La DCT tiene excelente compactación de energía para datos altamente correlacionados. La mayoría de la energía de los datos es empaquetada en unos pocos coeficientes de la DCT, mientras que los coeficientes de pequeña magnitud son tomados como cero. La DCT es

aplicada en las normas JPEG y H.261. como un paso importante para la compresión de video.

En la transmisión de mensajes que tienen valores de muestra repetidos hay muy poco contenido de información. Una forma de superar esta situación es enviar sólo las diferencias entre valores de muestra sucesivos, codificadas en forma digital. Esto se conoce como *Modulación de Código de Pulsos Diferencial*, abreviada como *DPCM* (*differential pulse-code modulation*). El esquema de compresión *DPCM* trabaja sobre una imagen entera, pixel por pixel. El primer pixel del ángulo superior izquierdo es codificado exactamente con la brillantez original. El proceso lo mueve hacia el segundo pixel en la línea. El valor de brillantez del pixel precedente es sustraído de la corriente de brillantez del pixel, por lo que el resultado de la sustracción es el nuevo valor de codificación para el segundo pixel en la imagen. Este proceso se repite a través de la línea 0 de inicialización. De la siguiente línea el proceso inicia hacia abajo. El proceso continúa hasta que la imagen es codificada totalmente en *DPCM*. La operación *DPCM* trabaja asumiendo que los valores de los pixeles adyacentes son similares o altamente correlacionados, como resultado, estas diferencias serán usualmente valores pequeños. La operación de compresión *DPCM* trabaja mejor sobre imágenes que no tienen un número extraordinario de oscilaciones de gran brillantez entre pixeles adjuntos. Cuando se aplica a imágenes normales, el código *DPCM* puede proporcionar alrededor de 2:1 de tasa de compresión. Para imágenes de largas series de pixeles de valores constantes, las tasas de compresión se incrementan significativamente.

Entre las desventajas de la *DPCM* está el hecho de que si se comete un error, se mantiene una polaridad incorrecta hasta que se corrija. Además estos sistemas adolecen de una posible sobrecarga por tasa

de elevación debido a las operaciones de diferenciación y truncamiento.

La norma conocida como JPEG, es una de las normas que define los estándares para la compresión de imágenes fijas. Como todo método de compresión, es bastante adecuado, sobre todo si lo que se pretende es tener una imagen razonablemente adecuada al ojo humano. Esta técnica de compresión nos permite tener relaciones de compresión relativamente altas, aún y cuando se sacrifique un poco la calidad de la imagen, sin embargo, es una norma de compresión que nos permite obtener tanto buena calidad de imagen como una alta relación de compresión. Actualmente, este estándar de compresión de imágenes fijas es muy común en imágenes que se generan y se manipulan por computadora, motivo por el cual se ha hecho muy popular en tiempos recientes.

Para la compresión de imágenes en movimiento se utiliza la técnica conocida como compensación de movimiento, en ésta, el eje de interpolación es alineado con el de la trayectoria del objeto en movimiento, de modo que ya no se generan frecuencias temporales, de esta manera son posibles mayores factores de reducción de movimiento, pues solamente es necesario enviar los parámetros originales de movimiento y un pequeño número de diferencias de imágenes originales. Estimación de movimiento es literalmente un proceso, el cual analiza campos sucesivos de imágenes y determina como los objetos se mueven de un campo al siguiente.

En H.261 la compensación de movimiento se realiza basada en bloques para registrar las diferencias entre tramas, las cuales son entonces codificadas en DCT. Aquí, el dato del cuadro en la trama anterior se puede usar para predecir los bloques de imagen en la

trama actual. Como resultado solamente se transmiten las diferencias, típicamente de pequeña magnitud entre los bloques previamente desplazados y el bloque actual. Esta recomendación describe los métodos de codificación y decodificación de los servicios audiovisuales a las velocidades de $n \times 64$ kbps, donde n toma los valores de 1 a 5.

Existen consideraciones de diseño en H.261.

- Define esencialmente solo el **decodificador**. Sin embargo, el **codificador**, el cual no es completa y explícitamente especificado por el estándar, es compatible con el decodificador.
- H.261 es propuesto para comunicaciones en tiempo real, usa solamente la previa y más cercana trama como predicción para reducir el retardo de codificación.
- Trata de balancear las complejidades del *hardware* del codificador y del decodificador, dado que ambos son necesarios para aplicaciones de videoteléfono en tiempo real.
- H.261 es un compromiso entre el funcionamiento del codificador, requerimiento de tiempo real, complejidad de implementación, y robustez del sistema. El movimiento compensado codificado en DCT es un algoritmo maduro, y después de años de estudio, bastante general y robusto que puede manejar varios tipos de cuadros.
- Las estructuras de codificación finales y parámetros son orientados más hacia aplicaciones de baja tasa de bits. Esta selección es lógica, porque la estructura del codificador y parámetros de codificación es más crítica para el desempeño del códec a muy bajas tasas de bits. A tasas de bits más altas, los valores de parámetros más bajos que el óptimo no afectan mucho el rendimiento del códec.

H.261 especifica un grupo de protocolos que comprime cada ráfaga de bits, y un grupo de operaciones que cada decodificador estándar compatible debe ser capaz de realizar. La actual implementación de *hardware* del códec y la estructura del codificador puede variar drásticamente de un diseñador a otro.

La Recomendación UIT H.263, es un híbrido de la predicción entre imágenes que utiliza la redundancia temporal y la codificación con transformada de la señal restante para reducir la redundancia espacial. La configuración básica del algoritmo de codificación de fuente de vídeo se basa en la Recomendación UIT H.261. En la compensación de movimiento se utiliza la precisión de mitad de píxel, mientras que en la recomendación H.261 se utiliza la precisión de píxel entero y un filtro de bucle. Emplea también la codificación de longitud variable para la transmisión de los símbolos.

Las normas MPEG descritas someramente, son complejas y necesitarán notas adicionales en el futuro para su mejor comprensión. Sin embargo, son indicativas para los procesos que se esperan como es el caso de la transmisión digital de alta definición (HDTV) y el Grabador de Cinta de Vídeo digital o VCR digital, entre otros procesos.

La norma MPEG-1 tiene como objetivo la compresión de vídeo de 320 x 240, de movimiento total y es optimizado para aplicaciones de almacenamiento en CD-ROM, multimedia interactiva, codificado a una tasa aproximada de 1.5 Mbit/seg.

La norma MPEG-2 soporta formatos de vídeo entrelazados y características para HDTV, el decodificador MPEG-2 también debe poder decodificar MPEG-1.

La norma MPEG-3 fue pensada para aplicaciones de HDTV a una tasa de codificación de entre 20 y 40 Mbit/s. Se descubrió más tarde que con algunas mejoras, la sintaxis de MPEG-2 funcionaba bien para HDTV, de manera que ahora HDTV es parte de MPEG-2.

La MPEG4 tiene como objetivo aplicaciones que incluyen comunicación audiovisual en tiempo real, multimedia y percepción remota. El foco común de la MPEG4 es el desarrollo de los modelos de verificación (VM). Un modelo de verificación es una plataforma común con una definición precisa de algoritmos de codificación y decodificación que pueden presentarse como herramientas dirigidas a funciones específicas.

Las aplicaciones beneficiadas con las funciones de MPEG4 son: compras interactivas desde la casa, producción y edición de películas, búsqueda de información sobre Internet, realidad virtual, comunicaciones inalámbricas y búsqueda en bases de datos. Un modelo de verificación es un algoritmo de codificación y decodificación que está compuesto de herramientas y algoritmos que sirven para evaluar el desempeño de otras herramientas y algoritmos. Las herramientas ocupadas por MPEG4 son similares a las de otros estándares como MPEG1 y MPEG2 e incluye algunas herramientas nuevas. Es importante notar que en poco tiempo aparecerán nuevas versiones de MPEG4 y algunas de las herramientas descritas aquí serán diferentes. Resultados concretos muestran que MPEG4 provee nuevas funcionalidades así como contenido básico interactivo. Únicamente el futuro dirá el impacto que MPEG4 tendrá en las aplicaciones multimedia emergentes.

Actualmente MPEG7 centra sus estudios en "Descripción de interfase de contenido multimedia". Esto se extenderá a la búsqueda de

capacidades que incluyen más tipos de información. MPEG7 estandarizará la descripción de varios tipos de información multimedia incluyendo imágenes fijas, gráficos, audio, vídeo en movimiento e información de cuantos de estos elementos son combinados en una presentación multimedia.

Las ideas novedosas sobre la Televisión de Alta Definición(HDTV) de los sistemas analógicos; "Muse" y "HD-Mac", han ido perdiendo terreno. Con el avance de la tecnología y también por el hecho de que ahora en los Estados Unidos se proyecta pasar a la tecnología digital en televisión, que solo podrá considerarse "de alta definición" cuando cumpla el requisito de ofrecer una resolución vertical y horizontal de la imagen que sea el doble del sistema de televisión convencional. Este último requisito demostró ser posible gracias a las soluciones planteadas por el sistema de transmisión de la Gran Alianza HDTV que usa la norma MPEG-2 para la compresión de vídeo, en donde se proclama la posibilidad de utilizar canales de 6 Mhz para transmitir imágenes digitalizadas de alta definición. La normalización de la HDTV en un canal de ancho de banda de 6 Mhz está muy cercano y beneficiará a varios países.

Un sistema de videoconferencia debe satisfacer algunos estándares. El estándar H.323 de ITU permite interoperar entre sí a los diferentes equipos de videoconferencia. Entre otras cosas, H.323 especifica la codificación y decodificación de las señales de vídeo y audio. Mientras que T.120 es un estándar que soporta la colaboración de datos, con beneficios que incluyen la interoperabilidad entre múltiples plataformas, transporte de red, y velocidad independiente. T.120 soporta conferencia en pizarrón blanco, pantalla compartida, anotaciones, e intercambio de imágenes. De manera que resulta importante el seleccionar un producto que soporte ambos estándares H.323 y T.120.

Los recientes avances tecnológicos en computación y telecomunicaciones han despertado gran interés en sistemas de compresión de vídeo, los cuales hoy en día transmiten información vía Internet o la red telefónica, reduciendo con mucho los costos de la videoconferencia. Alentados por los bajos precios del equipo y el crecimiento del ancho de banda en la red, los proveedores de sistemas de videoconferencia ofrecen ahora sistemas que llevan videoconferencias full-duplex a través de redes LAN (*Local Area Network*) y WANs (*Wide Area Networks*).

La recomendación H.320 especifica la comunicación multimedia vía N-ISDN, y engloba otras normas de codificación de audio, vídeo, datos, señalización y técnicas de multiplexaje que deben cubrir los equipos para cumplir con la misma. Algunos de los modelos de equipos que existen en el mercado soportan algunos algoritmos de codificación (audio, vídeo) propios del fabricante; sin embargo la mayoría cuentan también con el estándar H.320.

Respondiendo a las necesidades de sistemas de comunicaciones móviles, se está desarrollando en Europa el *Universal Mobile Telecommunication System* (UMTS), que empleará técnicas digitales desarrolladas a lo largo de los noventa y tecnologías del año 2000, y cuyo paralelo es el *Future Public Land Mobile Telecommunication System* (FPMTS) que está estandarizado en la UIT, antes CCITT. El UMTS soportará voz, datos, fotografía, imágenes en movimiento, gráficos, multimedia, etc. Considerando esencial para el UMTS ofrecer mejor calidad en términos de cobertura y probabilidad de bloqueo. Y para el servicio de voz, la percepción global de calidad y fiabilidad de la llamada deberá ser comparable a la ofrecida por la red fija actual, desde luego sin olvidar la disponibilidad que deberá ofrecer cobertura universal.

En la segunda parte de la tesis, se revisaron las normas que definen tanto las terminales como los sistemas para la transmisión multimedia a través de B-ISDN así como de Redes de Área Local.

La norma H.321 define las características técnicas para la adaptación de las terminales de banda estrecha conocidas como H.320 a la RDSI de banda estrecha. La necesidad de establecer una comunicación entre todos los elementos desarrollados con el paso de los años nos lleva a la definición de estas especificaciones técnicas para adaptar las nuevas terminales existentes a las tecnologías existentes de transmisión de información como lo es la RDSI.

La misma necesidad de crecimiento así como las nuevas características de la información, llevó a la necesidad de definir una recomendación para la definición de una nueva terminal, la cual tiene como base la capacidad audiovisual, y su adaptabilidad y su capacidad de interacción tanto con otras terminales como con otros tipos de redes (llámese Red ATM, RDSI-BE, RDSI-BA, etc.).

El crecimiento de la tecnología informática con ha llevado al desarrollo de nuevos tipos de redes como las Redes de Área Local, las cuales debido a la rapidez con la que han ido progresando requieren de tener ya no únicamente el compartir información, sino también el intercomunicarse a través de vídeo. Las recomendaciones H.322 y H.323 nos definen las características que deben de tener la información y las terminales videotelefónicas que se conecten a ellas. La primera de las recomendaciones define las características en las que la LAN se ajustan a las características de la RDSI-BE (que ofrecen una calidad de servicio garantizada). La segunda recomendación (H.323) nos define las características de los equipos que deben de

utilizarse para el caso en que la LAN no ofrezca una calidad de servicio garantizada como en el caso de la H.322. Establece las características de las terminales, los equipos y los servicios de comunicación que se requieren para la integración de terminales videotelefónicas a este tipo de redes.

Para la codificación de video, se cuenta con la tecnología de VLSI que ayuda en la implementación de sistemas muy complejos, una herramienta de la tecnología de VLSI son las arquitecturas programables, las cuales proporcionan una significativa flexibilidad comparadas con las arquitecturas dedicadas a una función específica, pero este incremento de flexibilidad lleva a un decremento de la eficiencia de la arquitectura, para solucionar esto, muchas arquitecturas combinan módulos adaptados a una tarea específica y módulos programables. Por consiguiente la optimización del área de silicio que contienen las arquitecturas dedicadas, comparadas con las arquitecturas programables llevan a un más bajo costo de producción, adaptándose a requerimientos de paralelismo y sistemas multiprocesadores programables.

Dado que el crecimiento de la información y de las necesidades de comunicación, así como la inclusión de video además de voz y datos, el desarrollo tecnológico nos ha llevado al punto en que es necesario el desarrollar tanto equipos como sistemas y terminales capaces de poder adaptarse a casi cualquier tipo de medio, y ofrecer al usuario una calidad de información, lo suficientemente buena y transparente a él.

Esto se lleva a cabo gracias a la especificación de normas y estándares internacionales, los cuales fueron tratados de una manera descriptiva a lo largo de este trabajo de tesis. Consideramos importante mencionar

que cada una de las normas que sobre el vídeo se han escrito podrían ser por sí solas un tema de tesis, sin embargo confiamos en que el presente trabajo cubra de una manera teórico-descriptiva algunos de los aspectos más importantes en el área del vídeo digital.