



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
"CUAUTITLÁN"**

REDES DE COMPUTADORAS

**"TÉCNICAS DE COMPRESIÓN EN EQUIPOS
DE VIDEOCONFERENCIA"**

**TRABAJO DE SEMINARIO
QUE PARA OBTENER EL TÍTULO DE:
LICENCIADA EN INFORMÁTICA
P R E S E N T A:
GRACIELA ROSAS CALVO ALVEAR**

ASESOR: ING. MIGUEL ÁLVAREZ PASAYE

Cuautitlán Izcalli, Edo. de México

1997

**TESIS CON
FALLA DE ORIGEN**



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.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN
UNIDAD DE LA ADMINISTRACIÓN ESCOLAR
DEPARTAMENTO DE EXÁMENES PROFESIONALES

U. N. A. M.
FACULTAD DE ESTUDIOS
SUPERIORES-CUAUTITLÁN



DEPARTAMENTO DE
EXÁMENES PROFESIONALES

DR. JAIME KELLER TORRES
DIRECTOR DE LA FES-CUAUTITLÁN
PRESENTE.

AT'N: ING. RAFAEL RODRIGUEZ CEBALLOS

Jefe del Departamento de Exámenes
Profesionales de la FES-C.

Con base en el art. 51 del Reglamento de Exámenes Profesionales de la FES-Cuautitlán, nos permitimos comunicar a usted que revisamos el Trabajo de Seminario:

Redes de Computadoras. Técnicas de compresión en equipos de
videoconferencia.

que presenta la pasante: Graciela Rosas-Calvo Alvear
con número de cuenta: 8842153-7 para obtener el Título de:
Licenciada en Informática

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el
EXÁMEN PROFESIONAL correspondiente, otorgamos nuestro VISTO BUENO.

ATENTAMENTE.

"POR MI RAZA HABLARA EL ESPIRITU"

Cuautitlán Izcalli, Edo. de México, a 15 de Octubre de 19 97

MODULO:

1

II

III

PROFESOR:

L. C. Carlos Pineda Muñoz

Ina. Miguel Alvarez Pasayo

Ina. J. Moisés Hdez: Duarte

FIRMA:

DEP/VOBOSEM

OBJETIVOS

Objetivo General

Describir las principales técnicas de compresión de datos aplicadas a un Sistema de Videoconferencia

Objetivo Particular

Conocer las técnicas de compresión existentes, que harán posible la reducción de la velocidad de transmisión de señales de video, para que su aplicación a un Sistema de Videoconferencia logre un nivel óptimo y mejor funcionamiento al menor costo y sin que la calidad del video se pierda.

ÍNDICE

INTRODUCCIÓN

CAPÍTULO I CARACTERÍSTICAS DE UN SISTEMA DE VIDEOCONFERENCIA

1.1.	Antecedentes	2
1.2.	Tipos de Sistemas de Videoconferencia	7
1.3.	Aplicación y Beneficios	10
1.4.	Estándares de los sistemas de Videoconferencia	13

CAPÍTULO II ELEMENTOS QUE INTEGRAN UN SISTEMA DE VIDEOCONFERENCIA

2.1.	Red de Comunicaciones	21
2.2.	Sala de Videoconferencias	25
2.3.	CODEC	28

CAPÍTULO III PROCESAMIENTO DIGITAL DE SEÑALES DE VIDEO

3.1.	Conceptos Básicos de Comunicación de Datos	31
3.1.1.	Comunicación Analógica y Digital	31
3.1.2.	Modulación Digital	33
3.1.2.1.	Tipos de Modulación	35
3.1.2.1.1.	Modulación por Cambio de Amplitud (AM)	37
3.1.2.1.2.	Modulación por Cambio de Frecuencia (FM)	38
3.1.2.1.3.	Modulación por Cambio de Fase (PM)	40
3.1.2.1.4.	Modulación por Codificación de Pulsos (PCM)	45
3.1.2.1.5.	Modulación Delta (DM)	47
3.2.	Conceptos de Video	
3.3.	Medios Físicos de Transmisión	56
3.3.1.	Enlaces Terrestres	57
3.3.2.	Enlaces Aéreos	57
3.4.	Fases del Procesamiento de Imágenes	63
3.4.1.	Preprocesamiento	64
3.4.2.	Postprocesamiento	64

CAPÍTULO IV MÉTODOS DE COMPRESIÓN DE VIDEO

INTRODUCCIÓN	68
4.1. Codificación de Intracuadros	71
4.1.1. Codificación por Predicción	71
4.1.1.1. Codificación DPCM	73
4.1.1.2. Aspectos de Diseño	77
4.1.2. Codificación por Transformada	79
4.1.2.1. Transformadas	83
4.1.2.2. Cuantización	86
4.1.3. Codificación Híbrida (Transformada/DPCM)	87
4.1.4. Codificación por Interpolación	88
4.1.5. Codificación de la Subbanda	90
4.2. Codificación de Intercuadros	92
4.2.1. Estimación del Desplazamiento del Movimiento	92
4.3. Compresión con Pérdidas	94
4.4. Compresión sin Pérdidas	94

CAPÍTULO V EL MODELO DE COMPRESIÓN H.261

INTRODUCCIÓN	96
5.1. Estimación y Compensación del Movimiento	97
5.2. Etapa de Transformación	99
5.3. Cuantificación	100
5.4. Codificación de los Elementos Cuantificados	101
5.5. Codificación de Huffman	103
5.6. Ejemplo de Codificación de Bloques	108

CONCLUSIONES

BIBLIOGRAFÍA

INTRODUCCIÓN

Actualmente nuestra sociedad pasa por una etapa en que la computación y las redes informáticas han tenido en ésta, un impacto de enormes consecuencias. Estas herramientas han permitido que empresas y usuarios en particular multipliquen su productividad y eficiencia en el trabajo.

La evolución de las videocomunicaciones ha traído al video al escritorio y finalmente hasta la casa. La combinación de video y computadoras ha sido llamada de maneras diferentes: multimedia, producción de video, telecomputadora o videoconferencia de escritorio. Todas involucran, en varios niveles, la conversión de video a datos, su manipulación en una forma digital y su conversión de vuelta a video para su despliegue. Las videocomunicaciones se están desplazando desde la sala especial hacia el escritorio y el vehículo que acelera este desplazamiento es la microcomputadora.

Al sistema que nos permite llevar a cabo el encuentro de varias personas ubicadas en sitios distantes, y establecer una conversación como lo harían si todas estuviesen reunidas en una sala de juntas se le llama "Videoconferencia". En los Sistemas de videoconferencia, video, audio y datos son compartidos de forma interactiva entre las diferentes locaciones participantes. Debido a que el video es una señal muy grande, requiere de gran energía para su transmisión de un lugar a otro. Las compañías de teléfonos proveen estas conexiones (conexión a una línea ISDN), y entre más grande la energía o más amplio el ancho de banda, es mayor el costo.

Los primeros sistemas experimentales de Videoconferencia realizados por Bell Laboratories y por British Telecom Research empleaban canales analógicos de TV para la transmisión, actualmente debido a la digitalización de las redes de telecomunicación y desarrollo de las técnicas de transmisión ha sido posible reducir los costos empleando tecnología digital.

Así que, debido a los costos de control y a que los Sistemas de Videoconferencia actuales requieren reducir la velocidad de transmisión, las señales de video necesitan ser comprimidas y es precisamente este aspecto del que se ocupara el presente trabajo, la compresión disminuye la calidad del video. A medida que haya más compresión, es menor la calidad y es menor el costo. La unidad que realiza la compresión y descompresión es el "CODEC" (CODificador, DEcodificador), que actúa como interfaz en todo el equipo y la red, video, audio e información son conectados en el "CODEC", el cual transmite una sola señal digital a través de la red hacia las locaciones remotas.

Por otro lado la codificación eficiente de señales de video, ha sido objeto de estudio y desarrollo por muchos años y el número de algoritmos de compresión que han sido estudiados son extremadamente largos. En términos generales existen cuatro técnicas de compresión aplicadas a los Sistemas de Videoconferencia: Técnica predictiva, Técnica por transformada, Técnica híbrida y por Interpolación, que son de las que nos ocuparemos como punto central, en los capítulos subsecuentes de este trabajo.

CAPITULO I
CARACTERÍSTICAS DE UN SISTEMA DE
VIDEOCONFERENCIA

1.1. ANTECEDENTES

En 1964 AT&T presentó en la feria del comercio mundial de Nueva York un prototipo de videoteléfono el cual necesitaba líneas de comunicación bastante costosas para transmitir video en movimiento. El dilema fue la cantidad y tipo de información requerida para desplegar imágenes de video. Las señales de video incluyen frecuencias mucho más altas que las que la red telefónica podía soportar (sobre todo en los años 60's). Para ese entonces el único método posible que se utilizó para transmitir la señal de video a través de largas distancias fue por medio de satélites. La industria del satélite estaba en sus inicios entonces, y el costo del equipo terrestre combinado con la renta de tiempo de satélite excedía con mucho los beneficios que podrían obtenerse al tener a un grupo pequeño de personas comunicados utilizando este medio.

A través de los años 70's se realizaron progresos substanciales en lo que se refiere a comunicaciones y los diferentes proveedores de redes telefónicas entraron en una etapa de transición hacia métodos de transmisión digital. La industria computacional avanzó también, en el poder y velocidad del procesamiento de datos mejoraron, asimismo se descubrieron métodos de muestreo y conversión de señales analógicas (audio y video) en bits digitales.

El procesamiento de señales digitales también ofreció ciertas ventajas como son en las áreas de calidad y análisis de la señal; el almacenamiento y transmisión todavía presenta obstáculos significativos para este entonces. Una representación digital de una señal analógica requiere de mayor capacidad de almacenamiento y transmisión que la original, los métodos de video digital comunes de fines de los años 70's y principios de los 80's, requirieron de relaciones de transferencia de 90 megabits por segundo. La señal estándar de video era digitalizada empleando el método común

PCM (Modulación por Codificación de Pulsos) de 8 bits, con 780 píxeles por línea, 480 líneas activas por cuadro de las 525 para NTSC¹ y con 30 cuadros por segundo.

Los datos de video digital son un candidato natural para comprimir ya que existen muchas redundancias inherentes en la señal analógica original. La necesidad de una compresión confiable de datos digitales fue crítica para ese entonces.

Ciertos métodos de compresión de datos fueron descubiertos, los cuales eliminaron la información redundante en la señal, con lo cual se consiguió una reducción de la cantidad de datos utilizados en un 50% aproximadamente, es decir 45 mbps, una razón de compresión de 2:1.

Las redes telefónicas en su transición a digitales han utilizado diferentes relaciones de transferencia, la primera fue de 56 Kbps necesaria para una llamada telefónica (utilizando métodos de muestreo actuales), en seguida grupos de canales de 56 Kbps fueron reunidos para formar un canal de información más grande el cual corría a 1.5 mbps (comúnmente llamado canal T1), de esta forma varios grupos de canales T1 fueron reunidos para conformar un canal que corría a 45 mbps ("T3"). Así usando video comprimido a 45 mbps para transmitir video en movimiento a través de la red telefónica pública fue finalmente posible, pero todavía extremadamente caro.

Era necesario comprimir aún más el video digital para hacer uso de un canal T1 (con una razón de compresión de 60:1), el cual se requería para iniciar en el mercado.

¹ Es un sistema de color, adoptado en los EUA, fue desarrollado a finales de 1940, y obedece a las siglas de la organización por la que fue creado, "The National Television Systems Committee". Este es un estándar conocido también como EIA RS-170*, la señal NTSC está compuesta por la combinación de color y brillo en una sola señal

Fue entonces, que a principios de los 80's cuando algunos métodos de compresión aparecieron en escena, eliminaban la temporización y sincronización de la señal y además realizaban un análisis del contenido de la imagen para eliminar redundancias. La razón de imágenes presentadas en el video en Norte América es de 30 cuadros por segundo, sin embargo, esto excede los requerimientos del sistema visual humano para percibir movimiento, la mayoría de las películas cinematográficas muestran una secuencia de 24 cuadros por segundo. La percepción del movimiento continuo puede ser obtenida entre 15 y 20 cuadros por segundo, por tanto una reducción de 30 cuadros a 15 cuadros por segundo por sí misma logra un porcentaje de compresión del 50%. Una relación 4:1 se logra obtener de esta forma, pero aún sin lograr el objetivo de compresión 60:1.

El *codec* (COdificador/DEcodificador), del cual se hablará posteriormente, de principios de los 80's utilizó una tecnología conocida como codificación de la Transformada Discreta del Coseno (DCT), con esta tecnología las imágenes de video pueden ser analizadas para encontrar redundancia espacial y temporal, la *redundancia espacial* es aquella que puede ser encontrada dentro de un cuadro sencillo de video —áreas de la imagen que se parecen bastante y que pueden ser representadas con una misma secuencia, la *redundancia temporal* es aquella que puede ser encontrada de un cuadro de la imagen a otro —áreas de la imagen que no cambian en cuadros sucesivos. Combinando los métodos antes mencionados, se logró obtener una razón de compresión de 60:1.

La compañía Compresión Labs Inc. (CLI) introdujo al mercado el primer CODEC, conocido como el VTS 1.5, el VTS significa (*Video Teleconference System*) y 1.5 (*1.5 mbps o T1*). En menos de un año CLI mejoró el VTS 1.5 para obtener una razón de compresión de 117:1 (768 Kbps), y tomó el nombre de VTS 1.5E. La corporación británica GEC y la corporación japonesa NEC entraron al mercado lanzando codecs que operaban con un T1. Ninguno de estos codecs fueron baratos, el VTS 1.5E era vendido en un promedio de \$180,000 dólares, sin incluir el equipo de video y audio necesarios para completar el sistema de conferencia, el cual era adquirido por un

costo aproximado de \$70,000 dólares, tampoco incluía el costo de acceso a la red de transmisión, el costo de utilización de un T1 era de \$1000 dólares la hora aproximadamente.

A mediados de los 80's se observó un mejoramiento dramático en la tecnología empleada en los codecs de manera similar, hubo una baja sustancial en los costos de los medios de transmisión. CLI introdujo el sistema de video denominado Rembrandt el cual utilizó una razón de compresión de 235:1 (384 Kbps). Fue después cuando una compañía, Picture Tel, introdujo un nuevo codec que utilizaba una relación de compresión de 1600:1 (56 Kbps). Picture Tel utilizó un nuevo método de codificación denominado Cuantificación Jerárquica de Vectores (HVQ). CLI lanzó poco después el codec denominado Rembrandt 56 el cual operó a 56 Kbps utilizando una técnica nueva denominada compensación del movimiento.

Al mismo tiempo que los proveedores de redes de comunicación empleaban nuevas tecnologías para abaratar el costo de acceso a las redes de comunicaciones, el precio de los codecs fue disminuyendo tan rápido como el aumento de los porcentajes de compresión.

En 1990 los codecs existentes redujeron su costo en más del 80% y su tamaño, el VTS 1.5E media cerca de 5 pies de alto y cubría un área de 2 y medio pies cuadrados. El Rembrandt 56 media cerca de 19 pulgadas cuadradas por 25 pulgadas de fondo.

Definición de Videoconferencia

La *videoconferencia* permite a un grupo de personas ubicadas en lugares distantes llevar a cabo reuniones como si estuvieran todas en una misma sala. Los participantes se pueden escuchar uno a otros y pueden verse en video en movimiento. Imágenes de documentos o archivos de computadora se pueden compartir también por todos los participantes.

Un sistema de Videoconferencia es una herramienta, pero además representa una arma estratégica en un mercado de información de alta competitividad.

Como sucede con todas las tecnologías nuevas, los términos que se emplean no se encuentran perfectamente definidos. La palabra "Teleconferencia" esta formada por el prefijo "tele" que significa distancia y la palabra "conferencia" que se refiere a encuentro, de tal manera que combinadas se refieren a un encuentro a distancia. En algunos lugares la palabra teleconferencia es usada como un término genérico para referirse a cualquier encuentro a distancia por medio de la tecnología de comunicaciones; de tal forma que frecuentemente es adicionada la palabra video a "teleconferencia" o a "conferencia" para especificar exactamente a que tipo de encuentro se está haciendo mención. De igual forma suele suceder con el término "audioconferencia" para referirse a una conferencia realizada mediante señales de audio.

La videoconferencia puede ser dividida en dos áreas: 1) la Videoconferencia Grupal o videoconferencia sala a sala con comunicación de video comprimido a velocidades desde 64 Kbps hasta 2.048 mbps y 2) Videotelefonía, la cual está asociada con la Red Digital de Servicios Integrados mejor conocida como ISDN operando a velocidades de 64 y 128 Kbps. Esta forma de videoconferencia esta asociada a la comunicación personal o videoconferencia escritorio a escritorio.

1.2. TIPOS DE SISTEMAS DE VIDEOCONFERENCIA

Existen diferentes tipos de sistemas de videoconferencia, y estos va en relación a los diferentes tipos de aplicaciones. Estos sistemas pueden ser de desktop (en una computadora), rollabout (sobreruedas), interconstruidos, para educación a distancia, telemedicina, etc. En todos estos sistemas, las partes que hacen funcionar el equipo son muy similares.

✓ Sistemas Desktop

Son basados usualmente en computadoras personales con procesador 486 o superior con tarjetas de expansión, una cámara, un sistema de audio y software basado en Windows. Requiere también una conexión a una línea ISDN (u otro tipo de línea digital) para realizar la transmisión. Durante una llamada se puede ver una imagen en movimiento de la persona en el otro extremo de la línea, lo puede escuchar, y todavía más útil, se pueden compartir archivos de PC y aplicaciones. La compresión realizada en este tipo de sistemas es muy severa y el resultado es que para algunas aplicaciones estos sistemas no llenan los requerimientos.

✓ Sistemas Rollabout o Sobreruedas

Son sistemas de Videoconferencia diseñados para pequeños y medianos grupos, estos sistemas rodantes son del tipo más común que se usa hoy en día. Usualmente están compuestos de uno o dos monitores que se encuentran contenidos en el mueble, junto con una cámara o más montadas en una cabeza que se puede girar e inclinar, así como el sistema de audio, el sistema de control y el CODEC.

El sistema de audio consiste en un cancelador de eco, micrófonos, bocinas y amplificadores. El sistema de control provee a los participantes de un control de las imágenes de video, orientación de la cámara, niveles de audio y otros periféricos.

La cámara en el sistema rodante captura la imagen de los participantes, ahí agrupados, y puede ser controlada vía remota para seleccionar varias tomas para switchear entre diferentes tomas comúnmente usadas. Generalmente, hay una cámara para documentos o gráficas, el cual es utilizado para intercambiar documentos, cartas, mapas, objetos y otras gráficas.

Los CODECS fueron diseñados para transmitir y recibir dos señales de video:

- Una señal de video en movimiento: Usualmente de la gente que se encuentra al frente de la cámara en el salón y,
- Una señal de video estática: Usualmente de la cámara de documentos.

Un sistema de dos monitores puede mostrar una señal en cada monitor, y un sistema de monitor individual usando una función llamada picture-in-picture (imagen sobre imagen) para mostrar ambas señales en el monitor. Cuando la imagen estática se envía la imagen de video en movimiento se congela momentáneamente hasta que la transmisión se completa.

✓ ***Sistemas Built-in (Interconstruidos)***

Los Sistemas Interconstruidos incluyen todos los equipos que un sistema rollabout, pero en vez de que todo el equipo resida en un gabinete con ruedas, estos sistemas se ubican en un lugar especialmente diseñado para ellos, pueden estar empotrados en una pared o en un rack. Esto crea una vista permanente de la sala que es conveniente para algunas aplicaciones especiales. Las capacidades de los dos tipos de sistemas son similares; aunque los sistemas interconstruidos frecuentemente tienen más periféricos conectados y se utilizan para aplicaciones más específicas.

✓ **Sistemas Especializados**

Existen otros sistemas más especializados, como los de educación a distancia y telemedicina que pueden ser fácilmente acomodados con diseño propio, y con periféricos adicionales como reproductores de 35 mm, toda clase de gráficas basadas en computadora, cámaras adicionales, monitores, otras fuentes de video, videograbadoras, pizarrones electrónicos, microscopios, otros dispositivos de imágenes médicas, etc. Existen actualmente una extensa lista de herramientas que pueden ser integradas en un sistema de videoconferencia para aplicarlas, según las necesidades de los usuarios.

1.3. APLICACIÓN Y BENEFICIOS

Aplicaciones

La baja registrada en los equipos de videoconferencia, así como el abaratamiento y disponibilidad de los servicios de comunicación han hecho que la industria de videoconferencia sea la de mayor crecimiento en el mercado.

Las aplicaciones de videoconferencia incluyen:

- Juntas de Dirección
- Servicio al Cliente
- Educación a Distancia
- Desarrollo de Ingeniería
- Reunión de Ejecutivos
- Estudios Financieros
- Coordinación de Proyectos entre Compañías
- Control de la Manufactura
- Diagnósticos Médicos
- Coordinación de Fusiones y Adquisiciones
- Compras
- Gestión y Apoyo de Ventas
- Contratación/Entrevistas
- Supervisión
- Adiestramiento y Capacitación

Algunos ejemplos más concretos de cómo se ha aplicado la videoconferencia en algunas de las áreas antes mencionadas es la siguiente:

Grupos de Trabajo Divididos: El departamento de Defensa de Estados Unidos y la Industria Aeroespacial han manejado el desarrollo de sistemas de armas muy complejas involucrando cooperaciones múltiples con agencias del Departamento de defensa a través de un sistema de seguridad de videoconferencia.

Viajes Internacionales en una Crisis: Algunos ejecutivos hicieron uso de sistemas de videoconferencia para manejar operaciones transnacionales durante la guerra del Golfo en 1991, cuando el viaje era difícil o peligroso.

Educación y Capacitación: Aprendizaje a distancia, el uso de videoconferencia para impartir educación y capacitación corporativa directamente en el lugar de trabajo ha sido la aplicación más exitosa y de mayor crecimiento de la videoconferencia. La Universidad de Minnesota está impartiendo un curso de Maestría en Educación a través de videoconferencia y opina que los beneficios institucionales obtenidos con el uso de la videoconferencia son entre otros: el incremento en la población estudiantil que recibe los cursos, reducción en la demanda de salones de clase, reducción de los costos de operación y organización de los cursos.

Beneficios

El beneficio potencial que representa el reunir personas situadas en diferentes lugares geográficos para que puedan compartir ideas, conocimientos, información, para solucionar problemas y para planear estrategias de negocios utilizando técnicas audiovisuales sin las inconveniencias asociadas de viajar, gastar dinero y perder tiempo, ha capturado la imaginación de las personas de negocios, líderes gubernamentales y educadores.

Probablemente, el mayor beneficio de la videoconferencia percibido por el cliente, sea el incremento en la calidad de comunicación interpersonal a distancia, que no se logra en otros servicios de comunicación tales como la audioconferencia, la

transmisión vía módem o fax, entre las más usuales. Otros de los beneficios que ofrece la videoconferencia son:

➤ **Eficiencia en tiempo y costos**

El avanzado concepto de videoconferencia permite la interacción y manejo simultáneo de señales de voz, video y datos, por lo que realizar una junta de trabajo entre personas que se encuentran distantes, puede significar un sensible ahorro en tiempo y dinero al evitar el desplazamiento a un mismo sitio.

➤ **Incremento de la productividad**

Las organizaciones modernas requieren alimentar el círculo virtuoso de los negocios, en los que la eficiencia se traduce en productividad.

Los novedosos productos que pone al alcance de las empresas la actual tecnología en telecomunicaciones, no puede sino contribuir a fortalecer esta sinergia.

➤ **Toma de decisiones oportunamente**

La oportunidad de respuesta, que consiste en reaccionar en el momento preciso y acertadamente, de una cualidad deseable en el ritmo de los negocios de hoy. Entablar comunicación con personas que se encuentran en puntos lejanos, es ya una realidad y sus múltiples ventajas harán del video enlace, sin duda, uno de los apoyos tecnológicos de uso más cotidiano en un breve lapso.

➤ **Apoyo de material gráfico**

Los recursos de la videoconferencia son ilimitados, pues al tiempo que las personas pueden interactuar en tiempo real, la cámara puede alternar la transmisión con gráficas, ilustraciones o documentos que se estén generando como fruto de la comunicación.

➤ **Video archivo**

Las videoconferencias pueden ser grabadas y guardarse como archivo documental.

1.4. ESTÁNDARES DE LOS SISTEMAS DE VIDEOCONFERENCIA

El mercado de la videoconferencia punto a punto estuvo restringido por la falta de compatibilidad, ya que los equipos (CODEC) manufacturados por diferentes vendedores no eran compatibles, es por eso que han sido desarrollados estándares que permiten que halla una compatibilidad entre los diversos proveedores de equipo.

La CCITT² es una parte de la Organización de las Naciones Unidas, y su propósito es el desarrollo formal de "recomendaciones" para asegurar que las comunicaciones mundiales sean establecidas eficiente y efectivamente. La CCITT trabaja en ciclos de 4 años, y al final de cada periodo un grupo de recomendaciones es publicado. Los libros "rojo" y "azul" que contienen estas recomendaciones fueron publicados en 1984 y 1988 respectivamente. En el libro rojo de 1984 fueron establecidas las primeras recomendaciones para codecs de videoconferencia la norma **H.120** y **H.130**. Estas recomendaciones fueron definidas específicamente para la región de Europa y para la interconexión entre Europa y otras regiones.

Debido a que no existían recomendaciones para regiones fuera de Europa, la CCITT designó a un grupo de especialistas en Telefonía Visual con el fin de desarrollar una recomendación Internacional, el resultado del trabajo de estos especialistas fue una sola recomendación (Px64) que se aplica a los rangos de velocidad desde 64 kbps hasta 2 Mbps, donde los valores claves para P son 1, 2, 6, 24 y 30.

Posterior a esto en 1989 un diverso número de organizaciones en Europa, EUA y Japón desarrollaron codecs flexibles para encontrar una especificación preliminar de la recomendación. Varios sistemas fueron interconectados en los Laboratorios y a través de largas distancias para poder validar la recomendación. Estas pruebas dieron lugar a la aparición de una versión preliminar de la recomendación H.261 "Video codec para servicios audiovisuales a Px64 Kbps" (de la cual se hablará más específicamente en un capítulo posterior) en el libro azul de CCITT. Sin embargo,

esta versión estaba incompleta y finalmente en diciembre de 1990 la versión final de la recomendación fue aprobada.

A continuación se enlistan cuales son las recomendaciones de la CCITT que definen las comunicaciones audiovisuales sobre redes digitales de banda angosta, así como la descripción de algunas de ellas.

En cuanto a servicios:

F.710 Servicios de Videoconferencia

F.721 Servicio Básico de videoteléfono en banda angosta en la ISDN

H.200 Recomendaciones para servicios audiovisuales

En cuanto a equipo terminal Audiovisual: punto a punto

H.320 Equipo terminal y sistemas de telefonía visual para banda angosta

H.261 Video codec para servicios audiovisuales a Px64 Kbps

H.221 Estructura de comunicaciones para un canal de 64 Kbps a 1920 Mbps en teleservicios audiovisuales

H.242 Sistemas para el establecimiento de las comunicaciones entre terminales audiovisuales usando canales digitales arriba de 2 Mbps

G.711 Modulación por codificación por pulsos (MPC) de frecuencias de voz

G.722 Codificación de audio de 7 khz dentro de 64 Kbps

H.120 Codecs para videoconferencia para grupos primarios de transmisiones digitales

H.130 Estructuras para la interconexión internacional de codecs digitales para videoconferencia de telefonía visual.

² Comité Consultivo Internacional de Telegrafía y Telefonía.

En cuanto a conexiones Multipunto:

- H.231** Unidades de control de multipunto (MCU) para sistemas audiovisuales usando canales digitales de más de 2 Mbps
- H.243** Procedimientos básicos para el establecimiento de las comunicaciones entre tres o más terminales audiovisuales usando canales digitales de más de 2 Mbps

En cuanto a la seguridad.

- H.233** Recomendaciones para sistemas de confiabilidad para servicios audiovisuales
- H.KEY** Recomendaciones de la CCITT de encriptación para servicios audiovisuales

Estándares ISO para almacenamiento y utilización de material audiovisual (MPEG)³

- Codificación de imágenes con movimiento y medios de almacenamiento digital para video para más de 1.5 Mbps (MPEG1)
- Codificación de imágenes con movimiento y medios de almacenamiento digital para video para más de 10 Mbps (MPEG2)
- Codificación de imágenes con movimiento y medios de almacenamiento digital para video para más de 40 Mbps (MPEG3)

³ Es un grupo de trabajo establecido por la Organización de Estándares (ISO), denominado Moving Picture Experts Group - Grupo de expertos en imágenes en movimiento

Recomendación ISO para almacenamiento y utilización de material audiovisual MPEG (Motion Picture Experts Group)

Los estándares MPEG son diseñados para la recuperación del video en movimiento. Hay dos estándares de recuperación de codificación de la ISO. MPEG1, el cual logra una buena calidad de video en 1.5 Mbps; y MPEG2 el cual pretende tener una buena calidad de video usando de 3 hasta 15 Mbps.

ESTANDAR H.320

En diciembre de 1990, la CCITT finalizó una serie de cinco recomendaciones (H.261, H.221, H.242, H.230 y H.320), las cuales definen en conjunto a una terminal audiovisual para proveer servicios de video teleconferencia y videotelefonía, sobre la ISDN⁴.

La recomendación de CCITT H.320 define la interrelación entre las cinco recomendaciones. Entre las funciones de la recomendación H.320 se encuentra la definición de las fases del establecimiento de una llamada en un teléfono visual y la definición de 16 tipos diferentes de terminales audiovisuales y sus modos de operación.

⁴ Red de Servicios Integrados.

ESTANDAR H.221

El propósito de esta recomendación es definir la estructura de la trama de comunicaciones para los teleservicios audiovisuales en un canal de 64 Kbps múltiple o sencillo ó canales de 1536 Kbps y 1920 Kbps los cuales hacen el mejor uso de las propiedades y características de los algoritmos de codificación de audio y video, de la estructura de trama de comunicaciones y de las recomendaciones CCITT existentes. Este estándar provee de la subdivisión dinámica o e un uso total de un canal de transmisión de 64 a 1920 Kbps dentro de velocidades más bajas utilizadas para audio, video, datos y propósitos telemáticos. Un canal simple de 64 Kbps está estructurado dentro de octetos transmitidos a 8 Khz. La posición de cada bit del octeto puede ser considerada como un subcanal de 8 Kbps. El octavo subcanal es denominado el canal de servicio (SC), el cual contiene dos partes:

- 1) FAS.- Este código de 8 bits se utiliza para situar los 80 octetos de información en un canal B (64 Kbps).
- 2) BAS.- Este código de 8 bits describe la habilidad de una terminal de estructurar la capacidad de un canal o canales múltiples sincronizados de varias maneras, y dirigir un receptor para demultiplexar y hacer uso de las señales constituyentes en esa estructura.

ESTANDAR H.242

La recomendación H.242 define el protocolo detallado de comunicación y los procedimientos que son empleados por las terminales H.320.

ESTANDAR H.230

Los servicios audiovisuales digitales son provistos por un sistema de transmisión en el cual, las señales relevantes son multiplexadas dentro de un patrón digital. Además de la información de audio, video, datos de usuario, etc. estas señales incluyen información utilizada para el funcionamiento adecuado del sistema. La información adicional ha sido llamada de "control e indicación" (C&I) para reflejar el hecho de que mientras algunos bits están genuinamente para el "control", causando un estado de cambio en algún otro lado en el mismo sistema, otros proveen de las indicaciones para los usuarios como para el funcionamiento del sistema.

El estándar H.230 tiene dos elementos primarios. El primero, define a los símbolos C&I relacionados al video, audio, mantenimiento y multipunto. El segundo, contiene la tabla de códigos de escape BAS, los cuales especifican las circunstancias bajo las cuales algunas funciones C&I son prioritarias y otras opcionales.

ESTÁNDAR G.711

Esta recomendación es utilizada para la voz y es muestreada a 8,000 muestras/segundo y codificada a 8 bits/muestra para una velocidad de 64 Kbps.

La recomendación G.711 describe las características de un sistema de codificación de audio (50 a 7000 Haz) el cual puede ser utilizado en una gran variedad de aplicaciones de voz de una mayor calidad.

RECOMENDACIÓN H.233

Este es un documento que describe los sistemas de confidencialidad para los servicios audiovisuales, describe la parte de confidencialidad de un sistema de privacidad apropiado para su utilización en servicios audiovisuales de banda angosta conforme a las recomendaciones de CCITT H.221, H.230 y H.242.

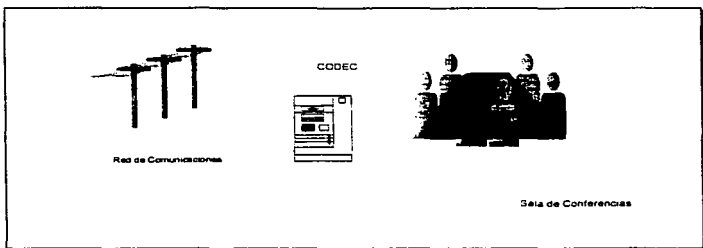
H.KEY

Documento que describe la autenticidad y los métodos de administración de las claves para un sistema de privacidad apropiado para ser utilizado en servicios audiovisuales de banda angosta que cumplan con las recomendaciones del CCITT H.221, H.230 y H.242. La privacidad se lleva a cabo mediante el uso de claves secretas, las claves son cargadas dentro de la parte de confidencialidad y controlan la manera en la cual los datos transmitidos son encriptados y descryptados.

CAPITULO II
ELEMENTOS QUE INTEGRAN UN SISTEMA
DE VIDEOCONFERENCIA

2.1. RED DE COMUNICACIONES

Para poder realizar cualquier tipo de comunicación es necesario contar primero con un medio que transporte la información del transmisor al receptor y viceversa o paralelamente, es decir en dos direcciones. En los sistemas de Videoconferencia es preciso que este medio proporcione una conexión digital bidireccional y de alta velocidad entre los puntos a conectar.



Elementos Básicos de un Sistema de Videoconferencia

La conexión para transmitir datos entre dos lugares puede hacerse en la mayoría de los países de tres formas: a) por medio del sistema telefónico (conmutado), b) A través de canales telefónicos privados (no conmutados) y c) por redes digitales de servicios integrados.

Las conexiones de los Sistemas de Videoconferencia son generalmente RDSI (Redes Digitales de Servicios Integrados). Las líneas RDSI son instaladas de la

misma manera que una línea telefónica. Después de instalada, hay un cargo mensual por tener la línea y cargos por llamadas individuales, basados en la duración y distancia de la llamada. Una llamada RDSI (de 28 Kbps*) cuesta más o menos el doble de una llamada de teléfono regular y esto se aplica internacionalmente.

Redes digitales de servicios integrados

Las redes de servicios integrados brindan a los usuarios el beneficio de poder acceder a múltiples servicios a través de un único punto de interconexión integrado y estandarizado ya que permiten la transmisión de cualquier tipo de señales: voz, datos y video. Desde el punto de vista del usuario, se puede imaginar que tan sólo a través de un enchufe se puedan acceder a varios servicios.

RDI (Red Digital Integrada)

TELMEX, inmerso en la corriente modernizadora que caracteriza al México actual, ofrece al mercado una forma de comunicación interactiva a distancia, la Videoconferencia ha dado un gran paso en la evolución de su sistema telefónico al proporcionar un elevado nivel de calidad con los sistemas más avanzados de conmutación y transmisión.

La RDI (Red Digital Integrada) es una red completamente digital capaz de transmitir cualquier tipo de señal e información mediante el uso de las tecnologías más avanzadas a nivel mundial.

* Kilobytes por segundo.

La RDI proporciona un medio de transporte de señales digitales conmutadas y de punto a punto con todas las modalidades de transmisión de información como voz, datos, textos y video en un solo sistema.

Así mismo, incorpora una red multiusuario de satélite para la interconexión de localidades remotas o aisladas que requieren ser integradas a los servicios de la Red Digital Terrestre.

Uno de los principales servicios de la RDI es el servicio de Videoconferencia. Este tipo de comunicación había sido manejado desde hace mucho tiempo en películas y libros de ciencia-ficción, considerando imposible su realización en el futuro.

El 18 de octubre de 1991 TELMEX, inaugura el servicio público de Videoconferencia en nuestro país, con el nombre de **VIDEO ENLACE DIGITAL TELMEX**, a través de un enlace entre la Ciudad de México y la Ciudad de Resto, Virginia, EUA.

Este servicio cuenta con 17 salas especialmente equipadas con tecnología de punta en red digital. Adicionalmente, por medio de su alianza con Sprint, la red de mayor cobertura del mundo, brinda la mayor oferta de salas públicas de videoconferencia.

MODALIDADES DEL SERVICIO:

Público:

En sus instalaciones de la Cd. de México, TELMEX cuenta con salas totalmente equipadas con la más avanzada tecnología y conectadas a la Red Digital Integrada, disponibles para ser rentadas por el público en general.

Privado:

Para aquellas organizaciones que tengan una sala propia, acondicionada con los requerimientos básicos, TELMEX se encarga de conectar el enlace con la Red Digital Integrada.

2.2. SALA DE VIDEOCONFERENCIAS

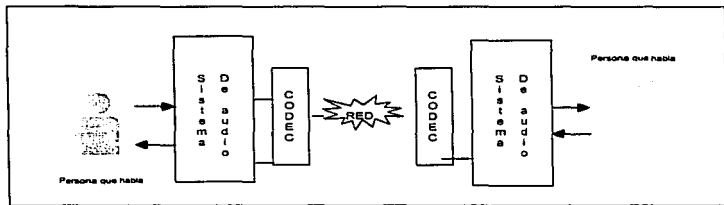
La sala de videoconferencia es el área especialmente acondicionada en la cual se alojarán los participantes de la videoconferencia, esta sala está dividida en cuatro componentes esenciales: el ambiente físico, el sistema de video, el sistema de audio y el sistema de control.

El ambiente físico se constituye por el aire acondicionado, la iluminación, la acústica del lugar y el amueblado.

El sistema de video esta conformado por los dispositivos que generan video, los dispositivos que reciben video, y los dispositivos que mueven el video de un extremo a otro. Las fuentes de video incluyen cámaras, proyectores en video de diapositivas, salidas de videograbadoras para reproducción, las salidas del video del codec. El destino del video incluye monitores de video, entradas de videograbadoras para grabación, entradas del codec para transmisión, impresoras de video, etc. cualquier videocámara (u otra fuente de video) puede ser enrutada a través del sistema de distribución al codec para su transmisión al otro extremo.

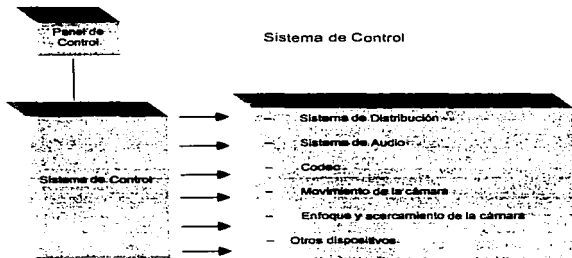
Los sistemas de videoconferencia normalmente incluyen una cámara simple localizada al frente de la sala de conferencia y cerca del monitor principal de video. El sistema de distribución es un cable el cual conecta directamente la salida de la cámara a la entrada del codec.

El sistema de audio tiene como propósito fundamental permitir a los participantes de ambos extremos de la junta escuchar y ser escuchados y está compuesto por micrófonos, bocinas, canceladores de eco y supresores de eco.



Sistema de Audio

El sistema de control es el corazón y el alma de la videoconferencia. El sistema de control tiene dos componentes principales: el panel de control y el sistema de control central



Es a través del control que los participantes trasladan sus deseos hacia acciones. Los participantes seleccionan cual fuente de video será vista en el extremo distante, como serán puestas las cámaras, la reproducción de un material, etc.

La sala de videoconferencia es un cuarto que se deberá sentir tan agradable como una sala de conferencias normal.

Una sala de videoconferencia típica está cerca de los 7.5 metros de profundidad y los 6 metros de ancho, estas dimensiones podrán albergar a un sistema de videoconferencia mediano y una mesa para conferencias para aproximadamente 7 personas.

2.3. CODEC

Las señales de audio y de video que se desean transmitir se encuentran en forma de señales analógicas, por lo que para poder transmitir esta información a través de una red digital, esta debe ser transformada mediante algún método a una señal digital, una vez realizada esta conversión de señales se debe comprimir y multiplexar dichas señales para su transmisión. El dispositivo que se encarga de este trabajo es el CODEC (CODificador/DECodificador) que en el otro extremo de la red realiza el trabajo inverso para poder desplegar y reproducir los datos provenientes desde el punto remoto.

La palabra codec es un acrónimo de Codificador/Decodificador. El CODEC codifica las entradas de audio, video y datos del usuario, y las combina para su transmisión en forma de una cadena digital de datos provenientes del punto remoto, separa o demultiplexa el audio, video y los datos de información del usuario, y decodifica la información de tal manera que pueda ser vista, escuchada o dirigida hacia un dispositivo periférico de salida situado en la sala de conferencia local.

El CODEC comprime el video y actúa como una interfaz entre todo el equipo y la red. La gran mayoría de CODECS pueden seleccionar diversos grados de compresión, dependiendo de la calidad de video requerida, a medida que hay mayor compresión, es mayor la degradación de la calidad del video.

El CODEC está compuesto de tres secciones fundamentales que son:

- a) Codec de la fuente
- b) Codec múltiplex video
- c) Codec de Transmisión

En el codec de la fuente, la señal de video es primeramente digitalizada y filtrada, es decir se ocupa de reducir el ruido.

El codec múltiplex video agrega, a la información de video, señales de sincronización de línea y de trama así como informaciones de dirección y otras.

El codec de transmisión acepta los datos de video, agrega un canal de 64 Kbits/segundo para sonido, un canal a 32 Kbits/segundo para la señalización de codec y canales de datos adicionales.

Los CODECS utilizan una gran cantidad de métodos para comprimir la señal de video en movimiento. En 1990 apareció un estándar internacional llamado H.32P, el cual permite a diferentes proveedores de CODECS, tener una comunicación. El estándar fue desarrollado por lo que entonces era llamado el CCITT, ahora llamado ITU, este estándar también es conocido como PX64 H.261.

La mayoría de los CODECS utilizan este estándar y son generalmente actualizables para incorporar lo último en componentes a medida que el estándar se expande en capacidad.

CAPITULO III
PROCESAMIENTO DIGITAL DE SEÑALES
DE VIDEO

3.1. CONCEPTOS GENERALES DE COMUNICACIÓN DE DATOS

Para el mejor entendimiento de los capítulos subsecuentes, en este punto se darán a conocer algunos términos y conceptos que permitirán al lector de este trabajo, tener una base conceptual de aspectos relacionados con el tema.

3.1.1. Comunicación Analógica y Digital

La diferencia o característica principal de un sistema de comunicación digital comparado con un sistema analógico es que el primero se transmite en forma de onda determinada, de entre un número finito de formas de ondas posibles, mientras que en el caso de la comunicación analógica el número de formas de ondas posibles es en teoría infinito.

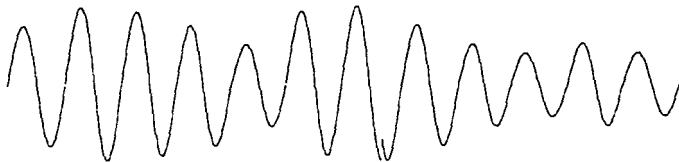


Figura 3.1 Señal Analógica



Figura 3.2 Señal digital

Porqué Modular una Señal?

La modulación es el proceso mediante el cual los símbolos digitales son transformados en formas de onda compatibles con las características del canal de comunicación.

Cuando se habla de señales de banda base, la forma de onda a que se está haciendo referencia son pulsos, mientras que en las señales moduladas o de banda ancha la información a transmitir modula una forma de onda senoidal (llamada señal portadora), la cual es entonces transmitida a través del canal de comunicación.

Una razón importante por la que se utiliza la modulación para la transmisión de señales es la de permitir multiplexar o agrupar varias señales diferentes a través del mismo canal de comunicación

3.1.2. Modulación Digital

El proceso de la digitalización fue desarrollado en los laboratorios Bell en los años sesentas, con la idea de superar algunas limitaciones de la transmisión analógica y de la grabación de imágenes. Las señales analógicas presentan una variación continua de su amplitud con el tiempo.

Las señales analógicas presentan ciertos problemas para su transmisión por el canal.

En primer lugar: la señal va siendo retransmitida por diversos amplificadores y otros transductores. La función de la retransmisión se diseña de modo que sea lo más lineal posible. La linealidad significa que la forma de onda que representa la señal mantiene sus características de un extremo de canal a otro. Una desviación de la linealidad produce distorsiones en forma de onda. Todas las señales analógicas son no lineales en alguna medida. Desgraciadamente, los componentes que intervienen en los sistemas, como los amplificadores, aumentan la no linealidad de las señales.

El segundo problema se relaciona con el ruido presente en el canal, el término ruido se refiere a aquellas señales eléctricas indeseables que se encuentran presentes en cualquier sistema, la presencia de ruido en una transmisión tiene efectos como: degradar la calidad de las señales, limitar la habilidad del receptor para identificar correctamente los símbolos de información. El ruido térmico o también conocido como ruido de Johnson, es ocasionado por el movimiento de electrones en los componentes (alambres, resistencias, etc). En un canal telefónico puede oírse ese ruido que se presenta como un soplo de fondo.

El ruido también está presente en la atmósfera terrestre, y también en la radiación precedente del sol y las estrellas, a este tipo de ruido se le conoce como fuente de ruido natural.

En tercer lugar, si la señal se almacena en un medio de grabación, como un disco o una cinta, el propio medio constituye una fuente de ruido. Por ejemplo las rugosidades de la superficie de un disco o de una cinta y su granulación son también una fuente de ruido.

En cuarto lugar, todas las señales se debilitan (o atenúan) durante su transmisión por el medio. Esta pérdida de señal puede hacer que la transmisión se atenúe tanto que se vuelve ilegible en el receptor. Un cable de alta calidad y de gran diámetro puede reducir en cierta medida la atenuación, pero esta no puede nunca eliminarse del todo.

Los sistemas digitales evitan estos problemas expresando las formas de onda analógicas mediante representaciones digitales binarias (1 y 0). En esencia, la forma de onda analógica se transforma en una serie de números digitales que se transmiten por el canal de comunicación como datos binarios. Los números digitales representan las muestras de la forma de onda, del muestreo hablaremos más adelante.

Las señales digitales sufren de los mismos problemas e imperfecciones que las señales analógicas (atenuación y ruido). Sin embargo, la señal digital es discreta: las muestras binarias de una forma de onda analógica se representan mediante niveles discretos de tensión, a diferencia de las señales analógicas, que toman valores no discretos. A medida que la señal viaja, por el canal, sólo es necesario detectar la presencia o ausencia de un pulso digital binario y no su amplitud como sucede con las señales analógicas. Es mucho más sencillo detectar la presencia o ausencia de un pulso de señal que captar la amplitud de una señal analógica. Por tanto, las señales digitales se pueden restaurar completamente, antes que se deterioren por debajo de un determinado umbral. El ruido y la atenuación se eliminan completamente en la señal reconstruida.

3.1.2.1. Tipos de Modulación Digital

Así como existe una gran variedad de técnicas de modulación para señales analógicas, también la información digital se puede montar sobre una señal portadora de muchas maneras.

Los diferentes tipos de modulación se basan en los tres parámetros básicos de toda onda senoidal:

- 1) Amplitud
- 2) Frecuencia
- 3) Fase

De esta manera, podemos definir también a la modulación, como el proceso mediante el cual la amplitud, frecuencia o fase de una portadora, o alguna combinación de ellas, es variada.

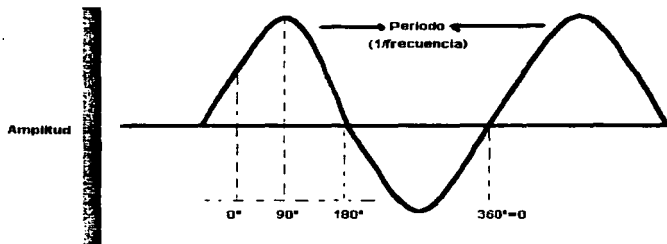


Figura 3.3 Parámetros de una señal

Los tipos básicos de modulación digital son los siguientes:

- 1) Amplitude Shift Keying (ASK) ó Manipulación por cambio de amplitud
- 2) Frequency Shift Keying (FSK) ó Manipulación por cambio de Frecuencia
- 3) Phase Shift Keying (PSK) ó Manipulación por cambio de Fase

3.1.2.1.1. Modulación por Cambio de Amplitud (AM)

En este tipo de modulación el voltaje de pico a pico de la señal portadora varía con la información que quiere transmitir. La modulación de amplitud que se muestra en la siguiente figura, ilustra como los picos de una amplitud representa unos binarios y los picos de la otra amplitud representan ceros binarios. La modulación de amplitud es adecuada para la transmisión de datos y permite usar con eficiencia el ancho de banda (el ancho de banda es la capacidad máxima de transmisión de un enlace) disponible de una línea de grado de voz, sin embargo existe la desventaja de que durante la transmisión es más susceptible al ruido.

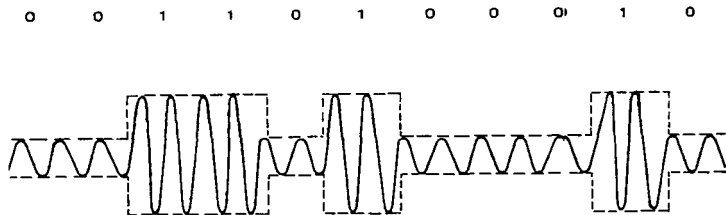


Figure 3.4 Modulación por cambio de amplitud

3.1.2.1.2. Modulación Por Cambio De Frecuencia (FM)

Este tipo de modulación es la más común, aquí la señal portadora se modula a distintas frecuencias. Por ejemplo, la señal puede modularse entre los 1200 y 2200 Hz (sin afectar la amplitud) como respuesta a la señal digital binaria. Las frecuencias específicas utilizadas dependen de equipo receptor y transmisor utilizado. Por ejemplo, una clase de equipo puede representar un 0 con 1200 Hz y 2200 Hz lo puede representar con un 1. La figura 3.5 nos muestra este tipo de modulación.

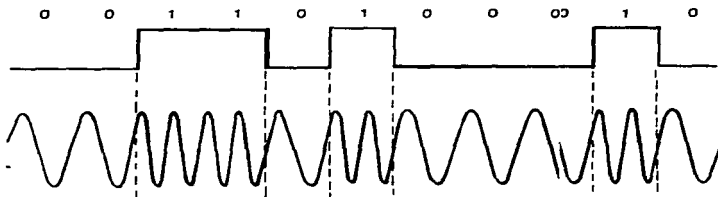


Figura 3.5 Modulación por cambio de frecuencia (aquí los ciclos por segundo varían)

A este tipo de modulación le afecta menos el ruido en las líneas de transmisión que a la amplitud modulada (AM). Por tanto, produce una transmisión de datos menos propensa a errores.

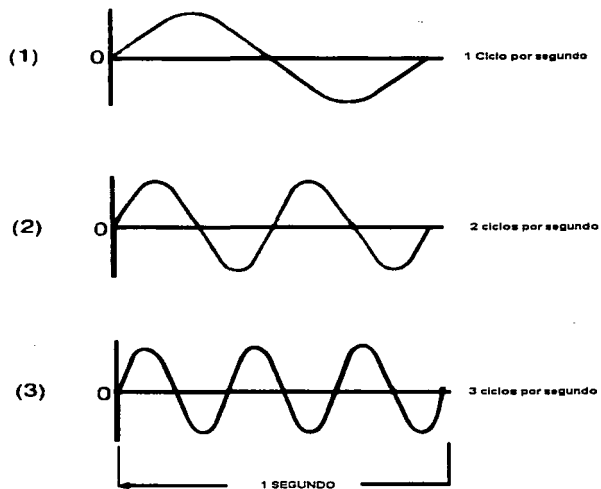


Figura 3.6 Variación de los ciclos de acuerdo al tiempo(segundos)

3.1.2.1.3. Modulación por Cambio de Fase (PM)

En esta modulación la fase de una señal portadora varía de acuerdo con los datos que quieren enviarse. La fase de la señal transmitida se desplaza un cierto número de grados como respuesta al patrón de bits que quieren transmitirse. Aquí la señal se desplaza 180 grados ($360/2$) dependiendo de que se indique un 1 ó 0 binario. Si no hay cambio, la señal se interpreta como un serie de unos o ceros. En esta técnica, siempre que se encuentra un corrimiento o una manipulación de fase de 180 grados (PSK)

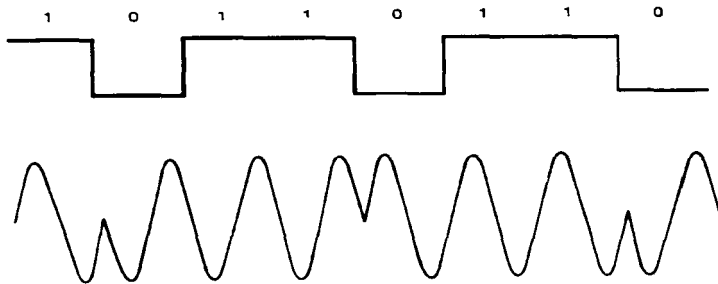


Figura 3.7 Modulación por cambio de fase

Formateado de Información Analógica

Como es de esperarse en el caso de la información analógica, antes de poder ser codificada es necesario su conversión a forma digital; el primer paso para llevar a cabo dicha conversión es el muestreo de la información analógica.

Muestreo De Una Señal

El proceso de muestreo consiste en tomar y analizar el valor que tiene una señal (muestra) a intervalos de tiempo regulares (velocidad de muestreo).

La señal resultante de este proceso se conoce como señal PAM (Pulse Amplitud Modulation) o modulación por amplitud de pulso ya que como se puede apreciar en la figura 3.8, consiste en una secuencia de pulsos cuya amplitud es aquella de la señal de entrada durante un lapso de muestreo.

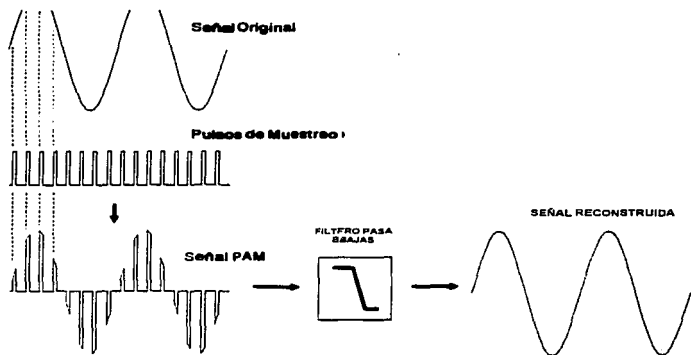


Figura 3.8 Procesamiento de una señal

Una virtud muy importante en este proceso es que la señal puede ser reconstruida a su forma original utilizando un simple filtro.

Representación de una Señal

Para poder procesar una señal analógica en un sistema digital, ésta debe ser representada de ciertas formas. La señal PAM que explicamos anteriormente conserva aún la característica propia de toda señal analógica en el sentido que representa un número infinito de valores posibles, de manera que es necesario someterla a algún proceso que limite el número de posibles valores. Este proceso generalmente se lleva a cabo mediante la siguiente secuencia:

a) MUESTREO Y RETENCIÓN:

Como su nombre lo indica, este proceso involucra la retención del valor muestreado de la señal analógica, el cual es retenido hasta que se efectúa la siguiente muestra.

b) CUANTIZACIÓN DE LOS PULSOS

Consiste en dividir el rango de amplitud de la señal en un número finito de valores discretos, y dependiendo de la amplitud de la señal analógica, asignar el valor discreto más cercano para cada muestra.

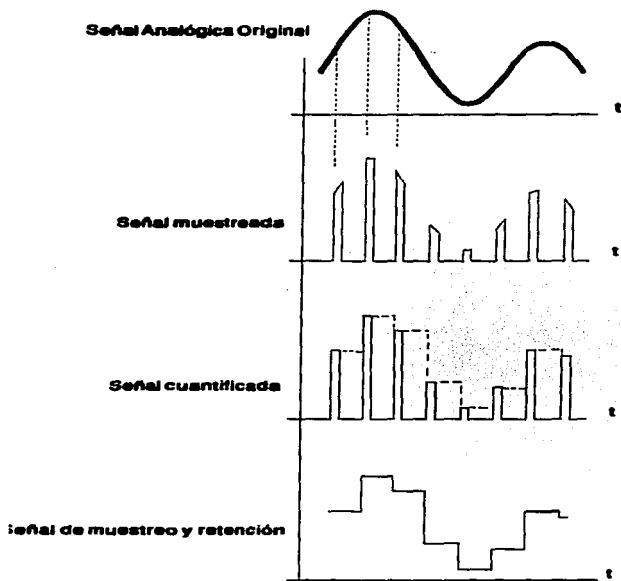


Figura 3.9 Representación de una señal analógica

3.1.2.1.4. Modulación por Codificación de Pulsos (PCM)

Aunque este tipo de modulación consiste en muchos procesos, generalmente éstos se resumen en tres: muestreo, cuantificación y codificación. Los dispositivos que efectúan el proceso de digitalización, denominados bancos de canal o multiplexores, tienen dos funciones básicas:

1. Realizar la conversión de la señal analógica a formato digital.
2. Combinar las señales digitales en una única cadena de datos utilizando la multiplexación por división de tiempo (TDM)⁶.

El PCM se basa en el teorema de muestreo de Nyquist, el que señala que si se muestrea una señal analógica a intervalos regulares con una frecuencia al menos dos veces superior a la frecuencia más alta de dicha señal, las muestras contienen información suficiente para permitir la reconstrucción completa de la señal analógica inicial.

Las muestras se recogen y almacenan en una determinada velocidad y se transforman en una cadena de bits. Cada muestra es un pulso modulado en amplitud (PAM). Una vez efectuado el muestreo, la señal PAM se somete al proceso de cuantificación. El objeto de la cuantificación es asignar un valor a cada pulso modulado en amplitud. Este valor se escoge en el rango de 1 a 128, o de 1 256. Si se escogen 128 valores, se requerirán 7 bits por muestra ($2^7 = 128$). Si se emplean 256 valores, se requieren 8 bits por muestra ($2^8 = 256$).

Una vez que el proceso de cuantificación ha asignado un valor a cada pulso modulado en amplitud, el tercer paso consiste en codificar las muestras formando una cadena de datos binarios.

⁶ Conocido también como multiplexaje por división de tiempos, se lleva a cabo asignando a cada canal un espacio de tiempo específico dentro de una trama de transmisión.

Para la reconstrucción apropiada de la señal, los datos deben pasar por un convertidor digital analógico (D/A) en la misma frecuencia en la que fueron muestreados originalmente.

La transmisión digital no se realiza sin problemas. Las señales digitales están sujetas a distorsiones. Un muestreo inadecuado o deficiente puede acarrear que la señal presente distorsiones, este problema se resuelve mediante un muestreo a mayor velocidad, sin embargo esto requiere componentes más caros y de mayor ancho de banda del canal.

Debido a la naturaleza analógica de las formas de onda, no hay ninguna técnica que permita eliminar completamente las distorsiones que introduce el muestreo. La anomalía fundamental radica en el hecho de aplicar muestras discretas (digitales) a una señal de naturaleza no discreta (analógica).

El segundo problema es debido a los errores de cuantificación. El proceso de cuantificación no representa exactamente las amplitudes de las señales PAM. Como la distorsión que introduce el proceso de cuantificación es proporcional al tamaño del escalón de cuantificación, una posible forma de resolver este problema sería aumentar el número de niveles de cuantificación que se utilizan para representar la señal. Pero esto tiene el inconveniente de un aumento en los costos de los componentes y del número de bits necesarios para representar la señal.

La técnica PCM se puede encuadrar dentro de la familia de análisis o síntesis de forma de onda. Se denomina así porque la forma de onda analógica se analiza y se muestrea para obtener códigos digitales a partir de los cuales podrá reproducirse.

Los sistemas actuales utilizan esquemas más complicados que la técnica PCM. Un sistema muy utilizado es la modulación diferencial por codificación de pulsos

(DPCM). Esta técnica no transmite los valores reales de las muestras de la señal, sino las diferencias entre las mismas. Como las muestras de una forma de onda están muy relacionadas entre sí (la información de una muestra a la siguiente es prácticamente redundante), se requerirán menos bits para representar el rango de las diferencias entre muestras. DPCM utiliza un cuantificado diferencial que almacena el valor previo de las muestras en un circuito de muestreo y retención. Este circuito mide y codifica el cambio entre una muestra y la siguiente.

3.1.2.1.5. Modulación Delta

La modulación delta (DM) es un caso particular de DPCM que utiliza únicamente un bit por muestra. En la modulación delta se mide la polaridad de las diferencias entre muestras sucesivas donde se utiliza un bit 1 para indicar un crecimiento de amplitud, y un bit 0 para indicar una disminución.

La señal se codifica como una escalera de amplitudes ascendentes y descendentes. El código digital se puede utilizar posteriormente para reconstruir la señal analógica (mediante el proceso de conversión digital-analógico) suavizando la escalera para obtener de nuevo la señal original.

La modulación delta es muy simple de realizar. Requiere, sin embargo, velocidades de muestreo mayores de PCM o DPCM, ya que las muestras no llevan mucha información. En la modulación delta se supone que la forma de onda codificada no difiere en más de un escalón de la señal original muestreada, sin embargo, es posible que la señal original varíe más rápidamente que lo que el modulador a escalones pueda reflejar, produciéndose entonces un problema que se denomina sobrecarga de pendiente.

3.2. CONCEPTOS DE VIDEO

En el procesamiento de señales de video han existido tradicionalmente dos partes: El video analógico y el video digital.

El video analógico fue desarrollado en los años 20's haciendo posible la transferencia de imágenes de video en movimiento por medio de señales eléctricas y después usando impulsos electromagnéticos (Radio Frecuencia). Desgraciadamente, las señales de video analógico requieren para su transmisión de un gran ancho de banda debido a la gran cantidad de información que manejan.

El procesamiento de video digital o procesamiento digital de imágenes, consiste en emplear técnicas computacionales para la manipulación y/o modificación de la información de una o varias imágenes.

El desarrollo de la tecnología para el procesamiento de imágenes se ha visto influenciada por el conocimiento en torno al sistema de visión humana.

El sistema de visión humana esta integrado de varios sensores de señales, donde uno de los más importantes es la retina, que contienen dos tipos de foto-receptores llamados bastones y conos. en su funcionamiento ambos foto-receptores proporcionan al nervio óptico una señal representativa de una imagen captada por el ojo. Esta señal es procesada en el cerebro y mediante un sistema de realimentación (compuesto de señales provenientes de distintos órganos) proporciona respuestas de adaptación, del individuo al medio que lo rodea.

Parámetros de una Imagen de Video

La señal eléctrica que representa una imagen puede ser obtenida de dos formas: por rastreo del elemento sensible para cámaras de tubos de rayos catódicos o por la lectura de la salida en el arreglo de sensores para cámaras de estado sólido. En ambos casos el voltaje está en función del tiempo y la señal contiene información sobre el promedio de intensidad de luz en cada punto de la imagen.

Para conseguir una señal de video que contenga todas las variaciones de luz y sombra, se deben explorar todos los detalles de la imagen sucesivamente o sea en un orden secuencial de tiempo. La exploración de la imagen se lleva a cabo de la siguiente manera:

1. Un haz electrónico barre transversalmente una línea horizontal, cubriendo todos los elementos de la imagen de la línea.
2. Al final de cada línea, el haz vuelve muy rápidamente al lado de la izquierda para comenzar la exploración de la línea horizontal siguiente. El tiempo de retorno es lo que se llama retraso o retorno. Durante el retorno no es explorada ninguna información de la imagen ya que durante este periodo el tubo de cámara y el tubo de imagen están inhibidos.
3. Cuando el haz ha retornado al lado de la izquierda, se sitúa en una posición vertical más baja a fin de que explore la línea inmediatamente inferior y no se repita la exploración de la misma línea. Esto se consigue por el movimiento de exploración vertical del haz, el cual esta provisto además de la exploración horizontal.

El número de líneas de exploración de una imagen debe ser grande con la finalidad de que se incluyan el mayor número posible de elementos de imagen y por consecuencia más detalles. El número óptimo de líneas de exploración por

cuadro es de 525 (para un ancho de banda de 6Mhz de los canales de TV) obtener una imagen completa o cuadro completo.

A su vez el haz que mencionamos anteriormente, se mueve hacia abajo al mismo tiempo que efectúa horizontalmente la exploración. Este movimiento vertical en la exploración es necesario para que no sean exploradas las líneas unas sobre otras. La exploración horizontal produce las líneas de izquierda a derecha, mientras que la exploración vertical espacia las líneas a fin de llenar el cuadro entre las parte superior e inferior.

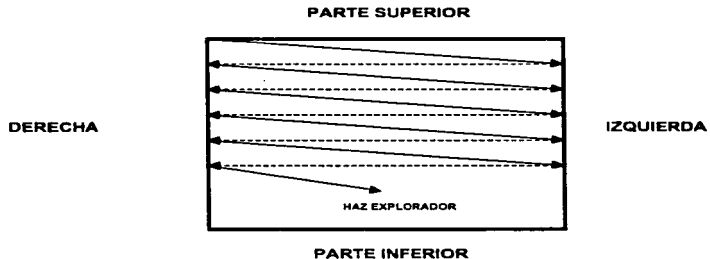


Figura 3.10 Exploración de una pantalla de video

Líneas por cuadro

El número de líneas de exploración de una imagen debe de ser grande con la finalidad de que incluya el mayor número posible de elementos de imagen y por consiguiente más detalles. El número óptimo de líneas de exploración por cuadro para el ancho de banda normal de 6Mhz de los canales de televisión es de 525 líneas.

Cuadros por segundo

El haz se mueve lentamente hacia abajo al mismo tiempo que efectúa horizontalmente la exploración. Este movimiento vertical en la exploración es necesario para que no sean exploradas las líneas unas sobre otras. El tiempo correspondiente a un cuadro completo con 525 líneas es de 1/30 segundos. Entonces la frecuencia de repetición de imagen es igual a 30 cuadros por segundo.

Frecuencias de exploración vertical y horizontal

La velocidad de campo de 60 Hz es la frecuencia de exploración vertical. Este es el ritmo con que el haz electrónico completa su ciclo de movimiento vertical desde la parte superior hasta la parte inferior de la pantalla para volver nuevamente a la parte superior. El tiempo de cada ciclo de exploración vertical de un campo es de 1/60 segundos. El número de líneas de exploración horizontal de un campo es la mitad del total de 525 líneas de un cuadro completo, ya que un campo tiene la mitad de líneas. Esto da por resultado 262.5 líneas para cada campo.

Con el tiempo que corresponde a un campo es 1/60 segundos y cada campo contiene 262.5 líneas, el número de líneas por segundo es de 15,750. Esta

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

Señal de Color

Cuando es explorada la imagen en el tubo de cámara, se producen señales de video separadas para la información de rojo, verde y azul de la imagen. Filtros ópticos de color separan los colores para la cámara. Sin embargo, para el canal estándar de 6Mhz de televisión, las señales de video de rojo, verde y azul son combinadas de modo que se forman dos señales equivalentes, una correspondiente al brillo y la otra para el color.

En esencia las dos señales transmitidas son las siguientes:

1. Señal de Luminancia. Contiene solo las variaciones de brillo de la información de imagen, incluyendo los detalles finos, lo mismo que en una señal monocromática. Esta señal se utiliza para reproducir la imagen en blanco y negro. Esta señal se denomina señal Y.
2. Señal de cromaticidad. Contiene la información de color y es transmitida como modulación en una subportadora de 3.58 Mhz aproximadamente. A esta señal se le denomina señal C o croma.

Estándares de algunos sistemas internacionales para televisión

Los datos manejados en el apartado anterior se aplican al estándar americano, pero existen algunos otros sistemas de televisión usados en otras partes del mundo, en la siguiente tabla se dan algunas características muy generales de ellos como son el Sistema NTSC (ya antes mencionado); PAL (Phase Alternation Line) y SECAM (Estándar de video de color en Francia, Oeste y Medio Oeste de Europa).

CARACTERÍSTICAS

SISTEMA	NTSC Americano	PAL Europeo	SECAM
Líneas por cuadro	525	625	625
Frecuencia de Barrido Vertical	60	50	50
Frecuencia de Barrido Horizontal	15750	15625	15625
No. de Cuadros por Segundo	30	25	25
Ancho de Banda, Video (Mhz)	4.2	5 o 6	6
Sistema de Color	NTSC	PAL	SECAM

Adquisición de la Imagen

Se puede decir el universo de las imágenes está compuesto por imágenes visuales que incluyen fotografías, dibujos y pinturas y las formadas por medios ópticos; las imágenes de representación matemática, que se basan en una función continua o discreta; las imágenes físicas no perceptibles por el ojo humano de manera inmediata como los mapas de densidad, imágenes por rayos X, tomografías computarizadas, imágenes por ultrasonido, por radar, etc.

La adquisición de una imagen es el proceso de transformar la imagen visual de cualquier objeto físico con sus características intrínsecas como su color, su textura, etc. en un conjunto de datos ordenados. En el proceso de adquisición es importante destacar varios aspectos que afectarán la calidad de la información obtenida, estos aspectos son :

A) ILUMINACIÓN:

La iluminación en el video, busca resaltar las características físicas como la textura, el color, los contornos, etc. donde el nivel total de iluminación determina la calidad de los datos obtenidos. Por lo tanto, una iluminación deficiente provoca manifestaciones de ruido en los datos.

B) FORMACIÓN Y ENFOQUE DE LA IMAGEN

Una cámara de video captura la imagen mediante un sensor a partir del cual una imagen visual es convertida a una señal eléctrica. En la formación de la imagen sobre el sensor, se emplea una lente, por medio de la cual la imagen es ampliada o reducida cuidando el enfoque para obtener la mayor nitidez de la misma.

C) DETECCIÓN DE LA IMAGEN

Aquí se convierte la radiación electromagnética proveniente de una imagen en una señal eléctrica. Por ejemplo una cámara de video tiene un elemento sensor localizado en su plano focal, y mediante un barrido de la superficie fotosensible del sensor se produce la señal de voltaje analógica, representativa de la imagen que se forma sobre dicha superficie.

Digitalización de la imagen

La transformación de señal analógica a señal digital necesita pasar por una serie de pasos:

a) El muestreo de la señal de video:

En este proceso se obtienen los valores instantáneos del voltaje analógico representativo de la imagen.

b) Cuantificar el valor de las muestras en niveles apropiados de intensidad de luz:

En este proceso se cuantifican los valores muestreados de la señal de video para establecer el número adecuado de niveles de intensidad de luz. La cuantización es una aproximación de la pérdida de los componentes de compresión del video. La cuantización reduce el número de valores necesitados para representar datos.

c) Digitalizar los valores cuantificados mediante convertidores analógico-digitales:

La etapa de conversión A/D es la que se encarga de obtener la representación digital de la intensidad

3.3 MEDIOS DE TRANSMISIÓN

En este apartado se explicarán algunas de las características de los principales medios físicos de transmisión, ya que para el adecuado diseño de una red de comunicaciones

El éxito o fracaso de un proyecto para la implementación de una red dependerá mucho del conocimiento que se tenga acerca de los medios de transmisión, conocimientos como los costos y tiempos de implementación, así como los beneficios y problemas que traen consigo.

Un medio de transmisión es el enlace o instalación físico que se usa para interconectar equipos y periféricos entre sí que crean una red que transporta información entre los mismos.

Para seleccionar el medio físico más adecuado, se deben tomar en cuenta las siguientes consideraciones:

- ✓ Tipo de ambiente donde se va a instalar
- ✓ Tipo de equipo que se va a utilizar
- ✓ Tipo de aplicación y requerimientos
- ✓ Capacidad económica (es decir la relación costo/beneficio)

Los medios físicos pueden dividirse en terrestres o aéreos

3.3.1. Enlaces Terrestres

- ↺ Par de cables torneados
- ↺ Cable coaxial de banda angosta
- ↺ Cable coaxial de banda ancha

3.3.2. Enlaces Aéreos:

- ↺ Microondas
- ↺ Láser
- ↺ Radio Frecuencia
- ↺ Satélites

A continuación hablaremos de los principales y mencionaremos algunas de sus características:

Par de cables torneados

Estos cables hechos de alambre se tornean o retuercen en pares para minimizar la interferencia electromagnética entre un par y el otro cuando se empaquetan en un cable grande. Es posible agrupar muchos cientos de pares de alambres en un cable grande

Entre sus características principales se encuentran:

- ✓ Un par puede transportar de 12 a 24 canales de grado de voz
- ✓ Son válidos en cualquier topología de red (anillo, estrella, bus, árbol)
- ✓ Pueden transportar tanto señales digitales como analógicas
- ✓ Tiene un alcance de hasta 3 Km, dependiendo del producto

- ✓ Permite trabajar en los métodos de transmisión Half Dúplex o Full Dúplex
- ✓ Es de bajo costo
- ✓ Tiene una alta tasa de error a grandes velocidades
- ✓ Tiene poca inmunidad hacia ruidos e interferencias
- ✓ Requiere protección especial: blindaje, ductos, etc.

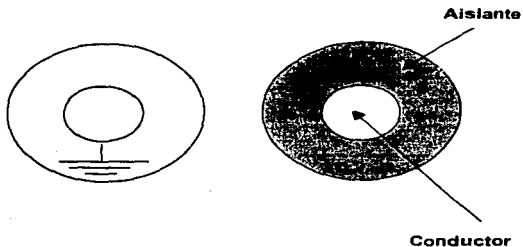


Figura 3.11 Corte de un par de alambres

Cable coaxial

Un cable coaxial puede transmitir a frecuencias mucho más altas que un par de alambres. El cable consiste en un cilindro hueco de cobre u otro conductor cilíndrico, que rodea a un conductor de alambre simple. El espacio entre el cilindro hueco de cobre y el conductor interno. Los cables coaxiales tienen muy poca distorsión, líneas cruzadas o pérdida de señal y constituye un medio de transmisión mejor que el par trenzado

Cable Coaxial de Banda Angosta

Entre sus características tenemos:

- ✓ Existen más de 150 variedades
- ✓ No hay modulación de frecuencia
- ✓ Están diseñados principalmente para comunicaciones de datos
- ✓ Es generalmente utilizado en topologías de red de bus, árbol y raramente configuración en anillo
- ✓ El alcance va desde 1 a 10 Km
- ✓ Su ancho de banda es de 10 Mbps
- ✓ Es de bajo costo y fácil de instalar
- ✓ Tiene poca inmunidad al ruido pero se puede mejorar por medio de filtros

Cable Coaxial de Banda Ancha

Características

- ✓ Es el mismo que se utiliza en redes de televisión
- ✓ Permite combinar voz, datos y video simultáneamente
- ✓ Permite voz y video en tiempo real
- ✓ La señal en el cable es un modo analógico y por lo tanto los datos deben ser modulados antes de su transmisión
- ✓ Las señales son Half Dúplex, pero utilizando dos canales se obtiene una transmisión Full Dúplex
- ✓ Las topologías en que es utilizado son de bus y árbol
- ✓ Su ancho de banda máximo es de 400 Mhz. Puede transportar el 100% de su carga

- ✓ Tiene mejor inmunidad a los ruidos
- ✓ Es un medio resistente que no necesita conducto
- ✓ Su costo es alto

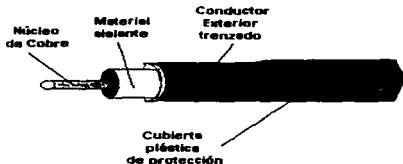


Figura 3.12 Un cable coaxial

Fibras Ópticas y Láseres

Son dos tecnologías nuevas para la transmisión de datos. Un láser genera un haz coherente de luz de frecuencia muy alta que puede transmitir 100 000 veces más información que los enlaces o eslabones actuales de microondas. Las fibras ópticas (filamentos delgados de vidrio o materiales vitroides) constituyen un medio de transmisión experimental que plantea promesas especiales cuando se acoplan con láseres.

Características

- ✓ Consiste en un núcleo central, muy fino, de vidrio o plástico, que tiene un alto índice de refracción
- ✓ El núcleo es rodeado por otro medio que tiene un índice un poco más bajo, que lo aísla del ambiente

- ✓ Cada fibra provee un camino de transmisión único de extremo a extremo, unidireccional
- ✓ La transmisión es generalmente punto a punto sin modulación
- ✓ La fibra óptica no es afectada por interferencia eléctrica, ruidos, problemas energéticos, temperatura, radiación o agentes químicos
- ✓ El ancho de banda es mucho más alto que cualquier otro medio
- ✓ Puede transmitir datos, voz y video
- ✓ Es altamente confiable
- ✓ La fibra es muy liviana, muy fina, durable y por lo tanto instalable en muy poco espacio.
- ✓ Su costo es muy alto
- ✓ Las topologías en las que se utiliza son anillo y estrella
- ✓ Requiere un mantenimiento muy especializado

Microondas

La transmisión de microondas puede portar muchos millares de canales de voz a la vez y no requiere que se tienda un cable. Esta transmisión se logra a través de la atmósfera entre torres de microondas generalmente espaciadas a 40 hasta 48 Km entre sí. Cada torre toma la señal transmitida de la torre anterior, la amplifica y la retransmite a la siguiente torre de microondas. Una antena típica para una torre de microondas tiene unos 3 metros de diámetro, aunque las distancias pueden ser más cortas.

Características

- ✓ La información se transmite en forma digital a través de ondas de radio de muy corta longitud. Pueden direccionarse múltiples canales a múltiples estaciones dentro de un enlace dado o pueden establecerse enlaces punto a punto.

- ✓ Las estaciones consisten en una antena tipo plato y de circuitos que interconectan la antena con la terminal del usuario.
- ✓ La transmisión es en línea recta y, por tanto, se ve afectada por accidentes geográficos, edificios, bosques, mal tiempo, etc.

Satélites

Los satélites de comunicaciones proporcionan una forma especial de transmisión de relevo de microondas. El satélite es una torre de microondas colocada a muchos kilómetros de altitud sobre la superficie de la tierra, generalmente sobre el ecuador. De esta manera puede retransmitir señales a distancias mayores que las posibles sobre la superficie terrestre, ya que la curvatura, montañas y otros obstáculos de la tierra bloquean la transmisión de microondas sobre líneas visuales entre las torres terrestres. Los satélites pueden manejar simultáneamente muchos millares de transmisiones de grado de voz. En la actualidad los satélites comerciales se colocan en órbitas muy altas (35,680 Kilómetros) viajando a la misma velocidad que la rotación de la tierra

3.4. FASES DEL PROCESAMIENTO DE IMÁGENES

Anteriormente los primeros sistemas de videoconferencia empleaban canales analógicos para la transmisión, pero hoy en día el avance de la tecnología en cuanto a la digitalización de las comunicaciones y al mejoramiento y descubrimiento de nuevas técnicas de proceso de señales en cuanto a video se refiere, a traído consigo que la transmisión sea a un menor costo.

Dada la importancia que tiene el procesamiento digital de señales, en un sistema de Videoconferencia a continuación se darán a conocer los principios básicos de ésta técnica, la cual se aplica en el CODEC, que como se ha mencionado anteriormente es el corazón de un Sistema de Videoconferencia.

En el CODEC tiene lugar las siguientes funciones:

- A) **INTERFACES ANALÓGICAS** Que se encargan de realizar funciones de pre y postprocesamiento de la señal analógica, así como la conversión de la señal analógica a digital y viceversa.
- B) **EL PRE Y POSTPROCESAMIENTO DE LA SEÑAL DIGITAL** Estos dos procesos tiene como finalidad realizar funciones de reducción de ruido, prefiltrado digital y demodulación para el preprocesamiento, funciones de interpolación digital, reconstrucción de los componentes de la señal y conversión de estándares para el postprocesamiento
- C) **REDUCIR LA VELOCIDAD DE TRANSMISIÓN EN LA CODIFICACIÓN Y DECODIFICACIÓN.** Aquí se incluyen funciones de aplicación de técnicas de compresión y operaciones inversas para el decodificador

- D) **FUNCIONES PERIFÉRICAS.** Estas realizan tareas múltiples como son la multiplexación y demultiplexación⁷, detección de errores y su corrección e interfaces hacia el canal.

3.4.1. Preprocesamiento

Una vez que hemos descrito las funciones básicas que se realizan en el Codificador-Decodificador (CODEC), a continuación se explicarán más detalladamente las funciones de preprocesamiento y postprocesamiento, que son las que se encargan de filtrar y mejorar la calidad de la imagen de video.

PREPROCESAMIENTO

Reducción del Ruido: Los reductores de ruido han sido empleados en el ambiente, de Videoconferencia ya que sus beneficios no solo se limitan a mejorar la calidad de la imagen, sino que es indispensable para una codificación o compresión eficiente.

Las fuentes de ruido para la Videoconferencia son numerosas, pero las más significativas son las que se generan en la cámara, y esto es por los niveles de luminosidad en la sala y otra es cuando el codec se encuentra físicamente muy lejos de la sala de conferencia y esto significa otra fuente de ruido significativa para la codificación.

⁷ La multiplexación consiste en dividir un canal de transmisión en dos o más canales separados, lo que puede lograrse dividiendo la variación total de frecuencias del canal en varias bandas de frecuencias más angostas, o asignando un canal dado sucesivamente a distintos usuarios en diferentes momentos. La demultiplexación es el proceso inverso.

El error que se genera por fuentes externas puede reducirse con filtros recursivos temporales. Estos sistemas utilizan el concepto " Estimación del movimiento", en donde se filtra la imagen donde no hubo cambios, a este tipo de reductores se les conoce como "reductores de ruido compensado por movimiento"

Prefiltrado Multidimensional: Una de las técnicas empleadas en la reducción de la velocidad de transmisión es el sub-muestreo espacio-temporal. Este proceso involucra la reducción efectiva de las frecuencias tanto espaciales como temporales por medio de filtros pasa bajas.

Demodulación Digital: Las técnicas de filtrado digital también se emplean para separar la señal de color, cuando la señal de entrada es compuesta, generalmente un filtro digital pasabanda es que extrae los componentes de crominancia, posteriormente un proceso de demodulación separa los dos componentes de crominancia.

3.4.2. Postprocesamiento

Las funciones de postprocesamiento son las que generalmente reconstruyen la imagen para ser desplegadas en un monitor, esto ocurre después de haberse convertido la señal digital a analógica. En el postprocesamiento se incluyen funciones de interpolación espacio-temporal, modulación de la señal de video a su formato original.

La interpolación espacio temporal: esta consiste en hacer un submuestreo de la señal, ya que reduce el número de píxeles* a codificar. Dado que una señal tiene tres dimensiones, la elección del patrón del muestreo es sumamente importante para tener el máximo de resolución posible. Por lo tanto el diseño de filtros de interpolación está completamente relacionado con el patrón de muestreo empleada.

Proceso de demodulación: Para los casos en que la señal es compuesta, se requieren procesos de modulación/demodulación y pueden ser aplicables antes o después de la conversión digital a analógica.

CAPITULO IV
MÉTODOS DE COMPRESIÓN DE VIDEO

INTRODUCCIÓN

El video es una de las formas de comunicación más completas que se han tenido. Desafortunadamente este representa también el medio análogo mas difícil de transmitir en una forma digital. El video es por naturaleza un medio analógico, con constantes cambios de señales eléctricas las cuales pueden ser transmitidas o almacenadas en una cinta magnética. Es posible digitalizar el video analógico y almacenarlo como "1" y "0" como lenguaje de computadora. Sin embargo la información de video se lleva una cantidad tremenda de espacio cuando se digitaliza por esta forma.

La videoinformación es difundida a los espectadores en series de imágenes o cuadros, y el efecto del movimiento se lleva a cabo a través de pequeños cambios continuos en los cuadros. Debido a que la velocidad es de 30 cuadros por segundo, los cambios continuos entre cuadro y cuadro son los que dan la sensación al ojo humano de movimiento natural.

Las imágenes de video están compuestas de información en el dominio del espacio y en el dominio del tiempo. La información en el dominio del espacio es provista en cada cuadro, y la información en el dominio del tiempo es provista por imágenes que cambian en el tiempo (por ejemplo las diferencias entre cuadros). Dado que los cambios entre cuadros colindantes son muy pequeños, los objetos aparentan moverse suavemente.

En los sistemas de video digital, cada cuadro es muestreado en unidades de "píxeles"^a o elementos de imagen. El valor de luminancia de cada pixel es cuantificado con ocho bits por pixel para el caso de imágenes blanco y negro.

^a Los píxeles son la unidad mínima por la que está conformada o compuesta una imagen.

En el caso de imágenes de color, cada pixel mantiene la información de color asociada, por primera instancia, en tres piezas de luminancia designados que son: rojo, azul y verde, los cuales son cuantizados también en ocho bits. La información de video que esta estructurada de la forma anterior contiene una gran cantidad de información ; por lo que, para su transmisión o almacenamiento, se requiere de la compresión (o codificación) de la imagen.

Para darnos una idea de que tanto espacio requiere la digitalización del video. Un simple recuadro de video requeriría para su almacenaje casi un megabyte, esto multiplicado por 30 cuadros por segundo, que es lo que se requiere para películas en movimiento, eso hay que multiplicarlo nuevamente por 120 cuadros o más, que es lo que en promedio tiene de duración una película o cinta cinematográfica. Estamos entonces hablando de 216,000 megabytes de espacio, esto refleja que el video necesita ser comprimido.

La técnica de compresión de video consiste en tres pasos básicos, primero el preprocesamiento (mencionado en el capítulo anterior) de las señales provenientes de diversas fuentes como: señales de TV, señales de televisión de alta definición HDTV, señales de videograbadoras VHS, BETA, etc. en este primer paso se realiza el filtrado de la señal de entrada para eliminar componentes inútiles y el ruido que pudiera haber en la señal. El siguiente paso consiste en la conversión de la señal de entrada a un formato común (CIF)⁹. El tercer y último paso es el de compresión, donde las imágenes comprimidas son transmitidas a través de la línea de transmisión digital y llegan al receptor donde son reconvertidas al formato CIF y son desplegadas después de haber pasado por la etapa de post-procesamiento.

Mediante la compresión de la imagen se elimina información redundante, principalmente la información redundante en el dominio de espacio y del tiempo. En general las redundancias en el dominio del espacio son ocasionadas a las diferencias entre pixeles contiguos de un cuadro o frame dado y las redundancias en

⁹ (The Common Interface Format (CIF), establece formatos de pantalla de baja resolución en un ancho de banda bajo y puede ser usado para aplicaciones de videoconferencia).

el dominio del tiempo son debidas a los cambios en cuadros contiguos causados por el movimiento de un objeto.

De aquí se derivan dos técnicas para la compresión de video:

- 1) El método de eliminación de redundancia en el dominio del espacio llamado codificación intracuadros o también llamado intraframe la cual puede ser dividida en Codificación por predicción, codificación de la transformada y codificación de la sub-banda.
- 2) Las redundancias en el dominio del tiempo pueden ser eliminadas por el método de codificación Intercuadros o interframe, que incluye los métodos de compensación/estimación del movimiento, el cual compensa el movimiento a través de la estimación del mismo.

4.1. CODIFICACIÓN INTRACUADROS

La codificación intracuadros utiliza solo la información espacial que existe en cada cuadro de video. Como esta codificación no utiliza ninguna información en el dominio del tiempo, puede ser usada para la codificación de imágenes fijas. La codificación intracuadros de señales de video resulta ser simple y no requiere de memoria que almacene cuadros precedentes o posteriores. Este método se divide en : codificación por predicción, codificación de la transformada, y codificación de la subbanda. En seguida se explicara cada uno de ellos:

4.1.1. Codificación por predicción

La codificación por predicción es uno de las técnicas mas antiguas de compresión de imágenes y esta basada en el hecho de que los errores de predicción son muy pequeños cuando el pixel presente es precedido por los pixeles vecinos. La técnica de codificación por predicción codifica el valor cuantificado de la diferencia entre el valor del pixel presente y el valor predicho (error de predicción). La utilización de un gran número de pixeles contiguos para la predicción puede disminuir el error de predicción y aumentar la efectividad del método. Pero como las ventajas de utilizar un gran número de pixeles vecinos no justifican la complejidad que esto conlleva, el número de pixeles vecinos utilizados para esta técnica no es mayor de cuatro.

La degradación de las imágenes en la codificación por predicción es ocasionada por el ruido granular, declinación de sobre carga y el llamado borde de negocios. El ruido granular se presenta cuando el proceso de cuantificación es muy grande, mientras que la declinación de sobre carga se presenta cuando el proceso de cuantificación es muy pequeño. El ruido granular y la declinación de sobrecarga ocasionan ruido y degradaciones en los límites del restablecimiento de la imagen. El borde de negocios es ocasionado cuando una imagen de la señal es mostrada continuamente en el tiempo. Esta técnica brinda una buena calidad para imágenes fijas y no para imágenes en movimiento.

Para disminuir esta degradación de las imágenes, la cuantificación se puede ajustar para las características visuales de los humanos, un filtro para reducción de ruido puede ser aplicado, y se pueden utilizar diferentes esquemas de codificación y de predicción para las diferentes partes de la imagen. Por ejemplo los límites de los objetos pueden ser tratados de manera diferente que las partes planas.

Esta técnica como mencionamos anteriormente requiere predecir el valor presente del pixel a partir del valor anteriormente codificado. El error resultante de la diferencia del valor real al predicho, es cuantizado dentro de un grupo de niveles de amplitud discretos. Estos niveles son representados como palabras binarias de longitud fija o variable y son enviados al canal.

Un codificador por predicción tiene tres componentes básicos como se muestra en la Figura : A)Predictor , B)Cuantizador y c) Codificador/Decodificador, dependiendo del número de niveles del cuantizador, el codificador es denominado modulador delta (DM) para $n = 2,5$ PCM; (DPCM) Diferencial. En la forma más sencilla, el codificador DPCM emplea el valor codificado del barrido horizontal que corresponde al pixel previo, como base para realizar la predicción. Sin embargo predictores más complejos emplean la línea anterior para realizar la predicción del cuadro.

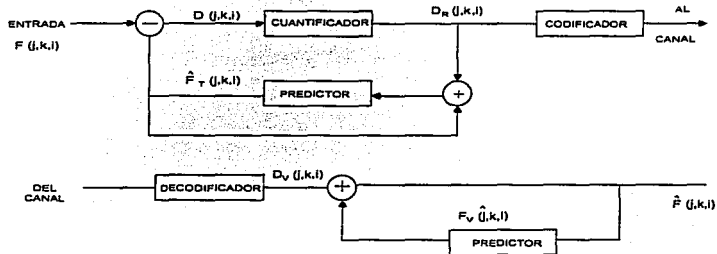


Figura 4.1 Diagrama de bloques de un codificador predictivo

4.1.1.1. Codificación DPCM

El concepto general de la codificación predictiva lineal fue desarrollado a partir de la aplicación de un codificador DPCM (Differential Predictive Coding Modulation) ver figura 4.2 para el procesamiento digital de video.

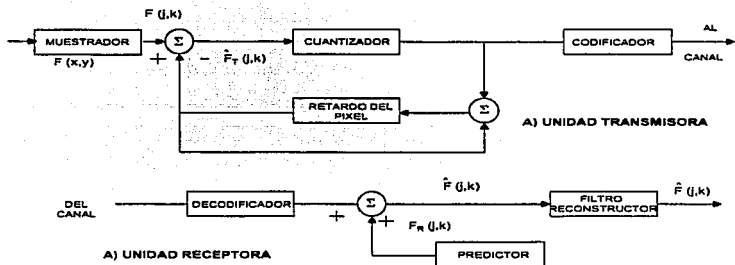


Figura 4.2 Sistema Codificador de Imagen por DPCM

En un principio, se propuso emplear a los integradores como predictores del valor presente del pixel a lo largo de una línea y la diferencia del valor presente y el estimado se cuantiza y se codifica para su transmisión. La diferencia del valor real y el estimado es cuantizada con 8 niveles, con códigos de 3 ó 4 bits, por lo que existe una reducción del ancho de banda de 8 bits por pixel que requiere el PCM a 3 bits por pixel en el DPCM. En el receptor, la diferencia de la señal es decodificada y reconstruida junto con la señal estimada por un predictor idéntico al que tiene el transmisor.

Los cuantizadores que emplean un codificador DPCM son no lineales y de 8 bits, como se muestra en la siguiente figura

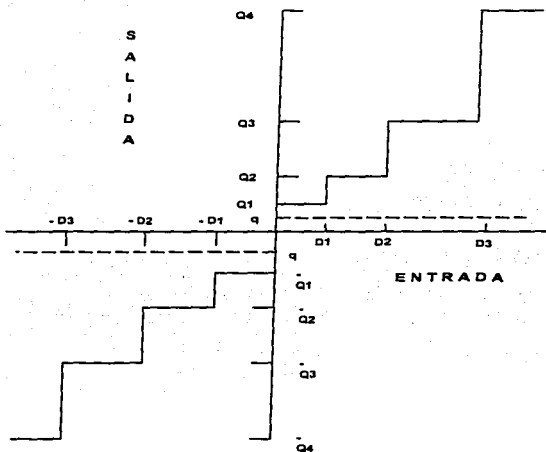


Figura 4.3 Cuantizador DPCM

En un codificador DPCM la probabilidad de ocurrencia de las señales diferenciales es no uniforme; las diferencias pequeñas son más comunes que las diferencias grandes. Por lo tanto, es posible emplear un código estadístico de longitud variable, como el de Huffman (que se explicará en otro capítulo) y de esta forma se obtiene una compresión mayor. Estudios de simulación han mostrado que con un código Huffman se puede llegar a requerir solo de 2.5 bits/píxel.

Los sistemas de videoconferencia actuales que emplean Codecs Predictivos han extendido y adaptado el concepto de modulación DPCM para las imágenes utilizando

más información en la predicción. En la siguiente figura se muestra el diagrama de bloques de esta versión de codificador DPCM, conocido como codificador predictivo espacial, en donde, se numeran todos los píxeles que se emplean en la predicción a conveniencia.

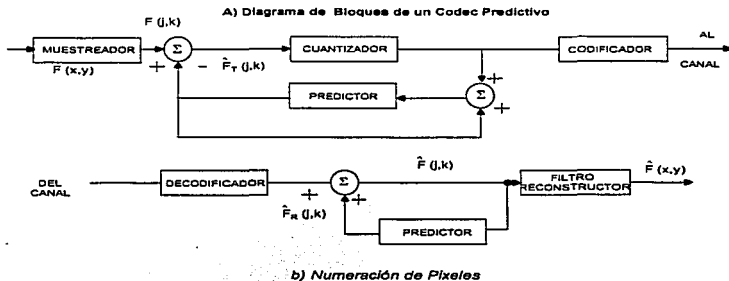


Figura 4.4 Diagrama de Bloques de un codec predictivo

En este sistema un codificador DPCM estándar emplea el píxel anteriormente barrido (S1) a lo largo de una línea de la imagen para la predicción de S0. A este predictor se le llama de primer orden y siguiendo esta nomenclatura, un predictor de segundo orden emplearía dos píxeles barridos previamente a lo largo de la línea (S1 y S2). Un

predictor de tercer orden emplearía a los píxeles S1, S2 y S3 como base para su predicción.

4.1.1.2. Aspectos de Diseño de los Codificadores de Imagen

En el diseño de un codificador de imagen por predicción las variables de diseño que se tienen son: el tipo de predictor a emplear, (número de píxeles usados en la predicción), minimización en la transmisión de errores, el número y distribución de los niveles de cuantización, el tipo de código más conveniente.

- A) Predictores: Los predictores para la codificación por DPCM son clasificados en lineales y no lineales, dependiendo si son función lineal o no de los valores muestreados y transmitidos previamente. Se puede hacer otra subclasificación dependiendo de la localización de los elementos previamente transmitidos: De una dimensión, cuando los elementos previos se encuentran en la misma línea y ; de dos dimensiones cuando emplean información tanto de líneas como de cuadros anteriores. Por otro lado, los predictores pueden ser fijos o adaptivos. Los predictores adaptivos cambian sus características en función de la información, mientras que los fijos mantienen sus mismas características independientemente de la información.

Para los sistemas de videoconferencia se emplean con más frecuencia los predictores intercuadros ya que existe una correlación entre el cuadro presente y el anterior por tratarse de imágenes con poco movimiento.

- B) Efecto de la Transmisión de errores: Uno de los retos que se tienen en el diseño de un predictor es la minimización del error en la imagen reconstruida. Dado que la reedición del próximo píxel se realiza por el codificador DPCM empleando información del valor de los píxeles anteriormente transmitidos, un error en la transmisión afecta no solo al píxel que se está transmitiendo, sino que también a los siguientes. La cantidad de error depende del predictor empleado. Para el caso de los predictores lineales de una dimensión, el error es constante, por lo que es

posible corregirlo. Sin embargo, para los predictores de dos dimensiones, la situación se dificulta. En este caso se requiere emplear necesariamente métodos de detección y corrección de errores. En algunos casos, la línea de barrido con errores es sustituida por la línea anterior o por un promedio de la líneas anteriores.

Para el caso de los predictores adaptivos, la transmisión de errores tiene dos tipos de efectos: uno es la utilización errónea del valor de predicción y la otra en la selección del predictor erróneo. En estos predictores, también se hace necesario emplear códigos detectores y correctores de errores.

- C) **Cuantización:** En un codificador DPCM se tienen 3 importantes degradaciones debidas al mal diseño de los cuantizadores: el ruido granular, la ocupación de borde y la sobrecarga en la pendiente. Los niveles más internos del cuantizador (para pequeñas magnitudes de diferencia), son muy gruesos, por lo que las áreas planas son cuantizadas gruesamente lo que da una apariencia de ruido aleatorio en la imagen. Por otra parte, si el rango dinámico del cuantizador es pequeño y cada borde de contraste alto requiere varios pasos o muestras para alcanzar a la señal de entrada ocasionando una sobrecarga en la pendiente. Los cuantizadores pueden ser diseñados en una base puramente estadística o empleando medidas psicovisuales y pueden ser fijos o adaptativos.
- D) **CODIGOS:** Se mencionó anteriormente que la frecuencia de ocurrencia de los niveles de salida del cuantizador es no uniforme, por lo que resulta conveniente emplear códigos de palabras de longitud variable (Código Huffman). En la siguiente tabla se muestra un código de longitud variable para un codificador DPCM con 16 niveles de Cuantización.

ESTRUCTURA DE LA
CÓDIGO DE LA PALABRA

NIVEL NO.	LONGITUD DE LA PALABRA DEL CÓDIGO	CÓDIGO
1	12	1 0 0 1 0 1 0 1 0 1 0 1
2	10	1 0 0 1 0 1 0 1 0 0
3	8	1 0 0 1 0 1 0 0
4	6	1 0 0 1 0 0
5	4	1 0 0 0
6	4	1 1 1 1
7	3	1 1 0
8	2	0 1
9	2	0 0
10	3	1 0 1
11	4	1 1 1 0
12	5	1 0 0 1 1
13	7	1 0 0 1 0 1 1
14	9	1 0 0 1 0 1 0 1 1
15	11	1 0 0 1 0 1 0 1 0 1 1
16	12	1 0 0 1 0 1 0 1 0 1 0 0

Código de Longitud variable con 16 niveles de cuantización

Como se observa, los niveles más internos ocurren con mayor frecuencia, por lo que se les asigna una longitud de palabra menor que las menos probables.

4.1.2. Codificación por Transformada

La codificación por transformada representa un cambio radical en la codificación de imagen que se realiza con DPCM o con técnicas de interpolación que se tratarán en el punto 4.1.4.. El proceso de codificación por transformada es indirecto, es decir se realiza un tratamiento matemático en la imagen y se obtienen los coeficientes de la transformada, los cuales serán cuantizados y codificados posteriormente.

Como resultado de las investigaciones realizadas, la codificación de la transformada, ha sido elegida como un estándar mundial para compresión de imágenes fijas. El concepto básico de la codificación de la transformada es obtener una relación de

compresión elevada mediante la eliminación de las redundancias a través de las transformadas ortogonales.

Partiendo de la suposición de que las características estadísticas de los datos de la imagen no cambian, la transformada de Karhunen-Loeve (KLT Karhunen-Loeve Transform), ha resultado ser la mejor transformada desde el punto de vista del error cuadrático. Pero debido al hecho de que las funciones fundamentales de la KLT deben ser enviadas al CODEC ya que estas funciones fundamentales son dependientes de los datos, y debido a la dificultad de la computación a gran velocidad que requiere, es impráctico la utilización de la transformada de Karhunen-Loeve (KLT) en las aplicaciones en tiempo real. Una transformada que es muy parecida a la transformada de Karhunen-Loeve es la transformada discreta del coseno, (DCT Discrete Cosin Transform), que se desempeña bien aún cuando no se toman en cuenta las características estadísticas de los datos de la imagen.

La transformada discreta del coseno realiza la transformada utilizando números reales y puede de esta manera emplear algoritmos de computación veloces que ya están implementados. El principio fundamental de esta técnica se explica a continuación:

La imagen de entrada es dividida en bloques de $N \times N$ pixeles, el tamaño del bloque es escogido considerando los requisitos de compresión y la calidad de la imagen. En general, a medida que el tamaño del bloque es mayor la relación de compresión también resulta mayor, esto se debe a que se utilizan más pixeles para eliminar las redundancias. Pero al aumentar demasiado el tamaño del bloque la suposición de que las características de la imagen se conservan constantes no se cumple, y ocurren algunas degradaciones de la imagen, como bordes sin definir en la imagen. Los resultados en la experimentación han demostrado que el tamaño del bloque más conveniente es de 8×8 pixeles. Después de dividir la imagen en bloques, la transformada discreta del coseno se aplica a cada bloque.

La transformada discreta del coseno en bi-dimensional y la transformada inversa se definen de acuerdo a las siguientes ecuaciones:

$$F(u,v) = \frac{1}{4} C(u)C(v) \sum_{i=0}^N \sum_{j=0}^N f(i,j) \cos\left(\frac{\pi u(2j+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right)$$

$$F(i,j) = \frac{1}{4} \sum_{u=0}^N \sum_{v=0}^N C(u)C(v) F(u,v) \cos\left(\frac{\pi u(2i+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right)$$

En las ecuaciones, $f(i,j)$ es el pixel con coordenadas (i,j) de cada bloque, y $F(u,v)$ es el coeficiente de la transformada correspondiente a cada frecuencia. el factor de peso $C(u)$ es $1/2$ cuando $u=0$, y 1 en cualquier otro caso. $F(0,0)$, que es el valor medio de los pixeles de un bloque específico, es a veces llamado el componente de DC o el componente constante.

De esta manera, el pixel $F(i, j)$ primero es transformado en $F(u,v)$ y después comprimido

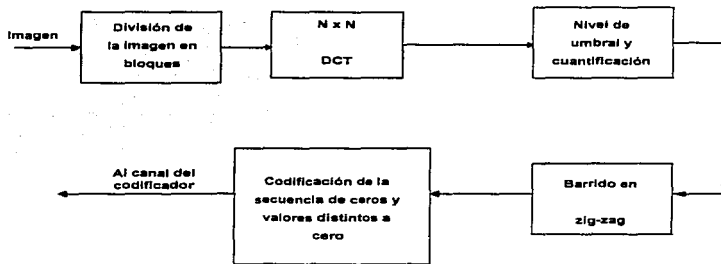


Figura 4.5. Diagrama de bloques de un codificador básico de transformación

Los coeficientes de la transformada $F(u,v)$ son cuantificados en base a un nivel de umbral para crear cuantos ceros sean posibles dentro del rango en el que no ocurran degradaciones en la imagen. Para garantizar continuidad entre los valores medios de los diferentes bloques, los componentes de DC son excluidos de esta cuantificación en base a un nivel de umbral y estos valores son cuantizados utilizando un tamaño de muestra pequeño.

Por último, los coeficientes que se encontraban formando arreglos de dos dimensiones son reordenados para formar arreglos de una dimensión usando un barrido en zig-zag. Debido a que ocurren largas secuencias de 0's cuando se efectúa el barrido en zig-zag la eficiencia de esta codificación también se incrementa. Los coeficientes diferentes de cero y las secuencias de ceros, se codifican utilizando un libro de código definido en base a los fundamentos de las estadísticas de los datos.

Como se mencionó anteriormente, las degradaciones de las imágenes ocurren cuando el tamaño de muestra de cuantificación es muy grande, por lo que se aplican diferentes tamaños de la muestra para las diferentes partes de la imagen, para los

bordes de los objetos se utiliza un tamaño de muestra pequeño y para las partes planas un tamaño de muestra mayor.

Algunas técnicas de DCT categorizan los diferentes bloques dentro de modelos dependiendo de las características de cada bloque y los manejan de acuerdo a las propiedades de cada modelo.

En seguida se describen brevemente algunos aspectos importantes de este tipo de codificación como son : la Aplicación de la transformada y cuantización.

4.1.2.1. Transformadas

El objetivo principal de obtener la transformada de la imagen es hacer a los elementos de la imagen o píxeles estadísticamente independientes. La mayor parte de las transformadas son lineales y unitarias.

La transformada óptima, es decir, la que provee la imagen con mejor calidad, se obtiene cuando se emplean coeficientes que no están correlacionados. Por ejemplo, si consideramos a los píxeles de una sub-imagen como un vector X , con n componentes, se desearía obtener una transformada "A" en la forma de una matriz de $n \times n$; y el vector de coeficientes sin correlación "Y" se obtendría de la siguiente manera:

$$T = AX$$

La transformada que ha demostrado utilidad es la transformada Hotelling o transformada Karhunen-Loeve (KLT) y se calcula de la siguiente manera:

$$Cx = E \left\{ (x - E(x)) \cdot (x - E(x))^T \right\}$$

Donde Cx es la matriz de covariancia, E es la esperanza, X es el vector de pixeles y T denota la transpuesta. Las filas de la matriz optima son los eigenvectores normalizados de Cx i.e., las soluciones de la siguiente ecuación:

$$Cx = \lambda x$$

Los coeficientes de "Y" obtenidos por esta transformada tienen una matriz de covariancia dada por :

$$C_y = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & \dots\dots\dots 0 \\ 0 & 0 & \lambda_3 & 0 \end{bmatrix}$$

donde $\lambda_1 \dots \lambda_n$ son eigenvalores de los coeficientes de la matriz de covariancia Cx . si ordenamos estos coeficientes conocidos como coeficientes K y si estos se envían al canal se tendría un error de reconstrucción en el receptor cuyo valor cuadrático medio es:

$$\sum_{j=k+1}^n \lambda_j$$

A pesar de que se conoce perfectamente, desde el punto de vista matemático, la transformada óptima, en la práctica se presentan muchos problemas que requieren trabajos de simulación para resolverlos. Por lo anterior, a veces resulta más económico aplicar otras transformadas "sub-óptimas", pero cuya implementación es más fácil, entre estas se encuentran:

A) Transformada de Fourier (FFT)

$$A = \text{Matriz}\{a_{ij}\}$$

donde:

$$a_{ij} = \frac{1}{\sqrt{N}} \exp[-2\pi\sqrt{-1} \cdot (ij)]$$

B) La transformada discreta del coseno (DCT):

$$a_{ij} = \frac{2k(i)}{\sqrt{N}} \cos\left[(2j+1) \frac{i\pi}{2n}\right]$$

Donde:

$$k(i) = \begin{cases} \frac{1}{2} & \text{para } i = 1 \\ 1, & \text{para } i = 2, \dots, N \\ 0 & \text{para otros} \end{cases}$$

C) La transformada de Hadamard:

La transformada simétrica de hadamard de orden $N=2$ esta definida por:

$$a_{ij} = \frac{1}{\sqrt{N}} \cdot (-1)^{b(i,j)}$$

Donde:

$$b(i, j) = \sum_{l=0}^{n-1} i_l j_l$$

los términos i_l y j_l son los estados binarios de i y j , respectivamente.

4.1.2.2. Cuantización

Otro factor importante que determina la calidad de la imagen en un codificador por transformada es la selección del cuantizador

Una vez que se determinan los coeficientes de la transformada, debe diseñarse un cuantizador para cada uno de ellos. Con el fin de minimizar el valor cuadrático medio del numero total de bits, es necesario considerar el mismo promedio de error para cada coeficiente en proporción al logaritmo de sus varianzas. A esta técnica se le conoce como "Submuestreo por zona" o "Filtraje por zona".

Otra alternativa para el diseño de cuantizadores es basarse en estadísticas, realizando histogramas de los coeficientes de la transformada.

4.1.3. CODIFICACIÓN HÍBRIDA (Transformada/Dpcm)

La codificación híbrida surgió de la idea de superar los inconvenientes que presentan los sistemas de transformada por un lado y los DPCM por el otro.

En la codificación híbrida, se consideran pequeños bloques o sub-imagenes; se evalúan los coeficientes y se codifican con DPCM. El siguiente es un diagrama de bloques de un codificador híbrido

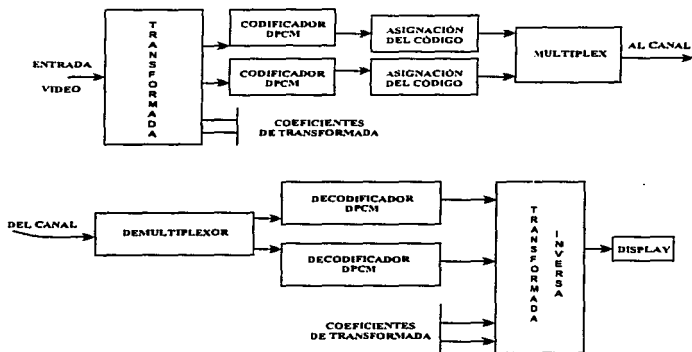


Figura 4.6 Diagrama de Bloques de un codificador-Decodificador Híbrido (Transformada/DPCM)

4.1.4. Codificación por interpolación

En la codificación por interpolación, se transmiten sub-grupos de píxeles y los que restan (que no se transmiten) se interpolan¹⁰, como ejemplo en la figura 4.7 se muestra como se realiza una interpolación con un sub-muestreo horizontal 2:1. Los elementos sub-muestreados son "tambaleados" de una línea a la siguiente y son interpolados por 4 promedios, como muestran las flechas de la figura 4.7 (Por ejemplo, el elemento A es interpolado promediando los valores de B, C, D y E).

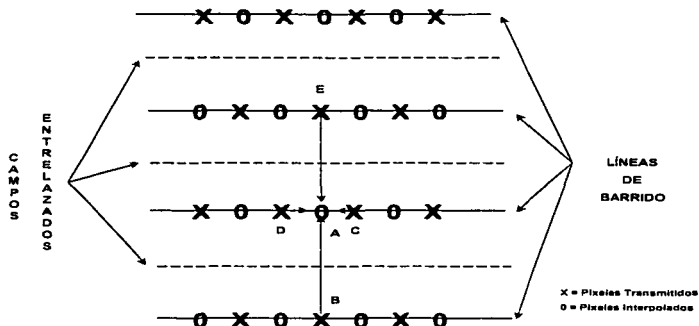


Figura 4.7 Codificación por interpolación

Los codificadores por interpolación se clasifican en dos grupos: fijos y adaptativos. En los codificadores fijos se escogen grupos fijos de píxeles y el resto se interpolan. Por otro lado, en el caso de interpolación adaptativa, se evalúa la interpolación antes

de la transmisión y se escogen puntos para transmitir y se interpolan los que no se transmiten. La evaluación se realiza en base al error. Por ejemplo, en la figura 4.8 se muestra una señal típica de vídeo. Empezando de la muestra "A", se podría interpolar directamente hasta "E" mediante una línea recta sin causar mucho error; pero si queremos interpolar hasta "E" el error resultante sería considerable, por lo anterior se escogen "A" y "E" para transmitir, y "B", "C" y "D" son interpolados.

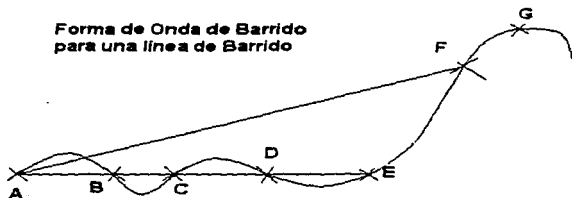


Figura 4.8 Representación de la técnica de Interpolación Adaptiva

¹⁰ Interpolar significa averiguar el valor de una magnitud en un intervalo, cuando se conocen algunos de los valores que toma a uno y otro lado de dicho intervalo.

4.1.5. Codificación de la subbanda

Aún cuando los fundamentos de la codificación de la subbanda son simples, el progreso en esta técnica para compresión de imágenes no se había logrado hasta hace poco. La codificación de la subbanda se compone de dos pasos. El primero de ellos es la filtración de la subbanda, que divide una señal de imagen en sus componentes de frecuencia, y el segundo paso es la codificación, que comprime cada banda de frecuencia de acuerdo a sus características respectivas.

La codificación de la subbanda es acompañada por un filtro de análisis en el codificador y un filtro de síntesis en el decodificador, respectivamente. El filtro de análisis divide la señal de entrada en diferentes bandas de frecuencia utilizando una velocidad de muestreo diferente para cada banda. En contraste el filtro sintetizador combina las diferentes bandas de la señal para sintetizar la señal deseada. La codificación de la subbanda requiere menor tiempo de procesamiento pero utiliza más procesadores, uno para cada banda.

Después de descomponer la señal en bandas de diferente frecuencia usando el filtro de análisis, se aplica un esquema de codificación apropiado para cada banda. Ya que las características de cada banda varían considerablemente y la sensibilidad visual humana también varía de banda a banda, un mejor desempeño se obtiene al tratar a cada una de las bandas de acuerdo a sus características particulares. Uno de los métodos más empleados es una combinación de la codificación intracuadros, de la subbanda y el de la transformada discreta del coseno, que trabaja de la siguiente manera: cada cuadro se descompone en cuatro bandas (LL, LH, HL, HH) aplicando un filtrado y análisis en la dirección horizontal y después en la dirección vertical.

La banda LL incluye la mayoría de los datos importantes excepto las orillas y los límites; por lo tanto, es necesario minimizar las pérdidas asociadas con la codificación de esta banda en particular. Es por esto que la codificación intracuadros es empleada generalmente para la codificación de la banda LL.

Las bandas de las frecuencias altas (LH, HL, HH) contienen la mayoría de la información de los límites de los objetos, los fondos, y las orillas, y los valores de los pixeles son generalmente menores que aquellos de la banda LL, entonces la información total contenida en estas bandas es menor que la contenida en la banda LL. Además de que los ojos humanos no son sensitivos a los cambios pequeños de los pixeles de estas tres bandas. Y partiendo de esto se puede aplicar una cuantificación no uniforme con alguna zona muerta para convertir los pequeños valores de los pixeles a cero sin que se note una degradación perceptible.

4.2. CODIFICACIÓN INTERCUADROS

Como se ha mencionado, existe numerosa información redundante entre cuadros de imágenes continuas, por tanto se puede obtener más información acerca del presente cuadro, que se determina de cuadros previos.

La configuración básica consiste de dos pasos: el primero nos va a dar la estimación y compensación del movimiento, y el segundo nos sirve para la compresión. El movimiento de un objeto es estimado para calcular el desplazamiento relativo entre el cuadro previo y el dato de la imagen correspondiente, generalmente en unidades de bloque. La diferencia entre el dato presente y el movimiento compensado del dato anterior es codificado para ser comprimido. La compresión de imágenes se usa para disminuir tiempos de redundancia.

4.2.1 Estimación del desplazamiento del movimiento

El método de estimación del desplazamiento consiste en un algoritmo recursivo basado en el pel¹¹. Este algoritmo estima el movimiento pel a pel recursivamente. El algoritmo de división de bloque BMA (Block Matching Algorithm) estima el movimiento bloque a bloque de la imagen, y el algoritmo de división recursivo, es una mezcla de los dos algoritmos anteriores. Para la estimación de movimiento generalmente es requerido un largo período de cálculo computacional, así es como el algoritmo de división es ampliamente usado, debido a que sus cálculos son implementables en un tiempo real.

¹¹ Acrónimo dado para nombrar el elemento de imagen

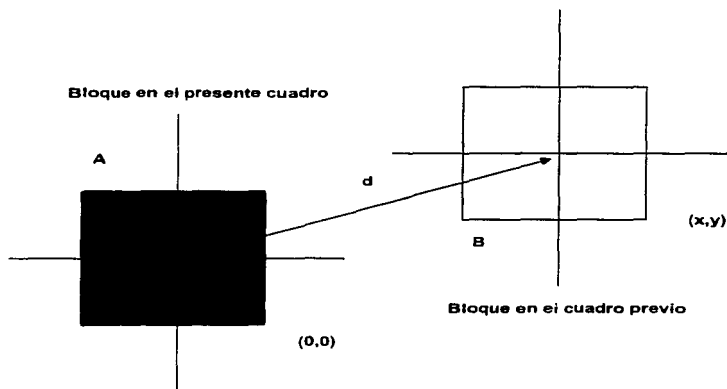


Figura 4.9 Estimación del vector de movimiento

EL BMA estima el movimiento de un bloque principal denominado bloque base. En este algoritmo todos los píxeles en el bloque simulan el movimiento en una dirección, la asociación de hardware y software es sencilla. La operación BMA consiste en dividir cada cuadro dentro de bloques tamaño $N \times N$, y el desplazamiento se realiza entre el cuadro presente y el cuadro anterior. La referencia para la estimación puede ser normal significando un error de cuadratura o error absoluto de diferencia, y el bloque con el mínimo de error es escogido como el vector de movimiento. El área examinada del cuadro previo es preespecificado, así la estimación de movimiento se da en todos los bloques dentro de esta área examinada.

El propósito de la estimación de movimiento es el de estimar los datos (bloque) de la imagen, presente en la pantalla de los cuadros previos o vecinos en orden para reducir el tiempo usado en las redundancias.

4.3. COMPRESIÓN CON PÉRDIDAS

Para ciertos tipos de imágenes y de aplicaciones puede permitirse cierta degradación de la información. Esto es el resultado de la compresión, la imagen no tendrá la misma calidad que el archivo original, pero la diferencia va a carecer de importancia. En esos casos se pueden alcanzar relaciones de compresión de 100:1 y más. Los algoritmos que proveen esta compresión se denominan con pérdida por que toleran la pérdida de alguna parte de la información para alcanzar compresiones elevadas.

Los programas que utilizan algoritmos con pérdidas a menudo le permitirán decidir la cantidad de pérdida que se puede tolerar, lo que se denomina factor Q. Existen numerosos esquemas para conseguir la compresión, eliminando redundancias en el texto o en las imágenes y reemplazándolas con codificaciones más cortas. Las pérdidas que se pueden aceptar, en una rutina de compresión de archivos de imágenes, se refiere al número de colores soportados o la resolución general aparente de la imagen, los límites de la imagen se vuelven menos definidos y se pierde detalle. La pérdida es aceptable, dependiendo de la aplicación que se requiera utilizar. Las transparencias de 35 mm de pequeño formato pueden tolerar factores altos de Q (factor de pérdida). Por otra parte, trabajos de formato grande, como películas de cine, no pueden aceptar pérdidas de ninguna especie.

4.4. COMPRESIÓN SIN PÉRDIDAS

La compresión sin pérdidas de información es lenta e ineficiente. Se puede conseguir compresión mucho más rápida si se utilizan procesadores dedicados para tal fin, particularmente si han sido diseñados para la compresión. Sin embargo, la compresión sin pérdidas es realmente un lujo innecesario en el caso de la mayoría de los archivos de imágenes y muchas aplicaciones pueden tolerar cierta cantidad de pérdida de información.

CAPITULO V
MODELO DE COMPRESIÓN H.261

INTRODUCCIÓN

El modelo del sistema PX64, el cual consiste básicamente de cinco etapas:

Una etapa de compensación del movimiento, una etapa de transformación, una etapa de cuantificación "lossy", ¹²(con pérdidas), y dos etapas de codificación del tipo "lossless", ¹³(sin pérdidas). La etapa de compensación del movimiento sustrae la imagen corriente de la vista cambiada de la imagen previa si ambas se asemejan. La etapa de transformación concentra la energía de la información en algunos de los primeros coeficientes de la transformada. Un cuantificador origina una pérdida controlada de información y las dos etapas de codificación proveen de compresión adicional de los datos.

El modelo PX64 es considerado un compresor del tipo lossy, debido a que la imagen reconstruida no es idéntica a la original. Los codificadores lossless, los cuales crean imágenes idénticas a la original alcanzan muy poca compresión debido a que los bits menos significativos de cada componente de color llegan a ser progresivamente más fortuitos, lo que los hace más difíciles de codificar.

¹² Lossy Compression: Es un tipo de compresión, donde los datos de la imagen original se conservan.

¹³ Lossless Compression: Es un tipo de compresión, donde los datos de la imagen no se conservan, y la calidad de la imagen se degrada.

5.1. ESTIMACIÓN Y COMPENSACIÓN DEL MOVIMIENTO

Debido a que la mayoría de los cuadros en una secuencia de imágenes se observan muy similares exceptuando los cambios debidos al movimiento, como los son un "paseo" o movimiento de la cámara a través de la escena, podemos evitar el codificar el mismo bloque dos veces enviando la codificación de este a través de vector de desplazamiento de la imagen previa causado por el movimiento de "paseo".

La estimación del movimiento compara a 16 por 16 macrobloques en la luminancia en un área pequeña de búsqueda de la imagen previamente transmitida. El rango para esa comparación esta entre ± 15 pixeles basados en el componente de luminancia de la imagen. El desplazamiento con la diferencia mas pequeña del macrobloque, determinada por la suma de los valores absolutos de la diferencia pixel a pixel a través del bloque, es considerada el vector de compensación de movimiento para ese macrobloque en particular. (El vector de movimiento de la crominancia es el vector de movimiento para la luminancia dividido a la mitad).

El bloque compensado en movimiento es la diferencia o error entre el bloque de mejor acoplamiento y el bloque actual a ser codificado.

La operación de la compensación del movimiento se muestra en la figura 5.1. El bloque "A" es un bloque en la imagen actual que será codificado. El bloque "B" es el bloque en la misma posición de "A" pero de la imagen que fue previamente almacenado tanto en el codificador como en el decodificador. Debido al movimiento de la imagen, el bloque "A" se asemeja más a los datos de los pixeles del bloque "C" que aquellos del bloque "B". El desplazamiento del bloque "C" desde el bloque "B" medido en pixeles y en direcciones X y Y , es el vector de movimiento. Las diferencias pixel por pixel entre los bloques "A" y "C" es transformada y codificada. El vector de movimiento y los datos codificados son transmitidos al decodificador, donde los datos del bloque inversamente transformados son agregados a los datos

en el bloque "C", apuntados por el vector de movimiento y situados en la posición del bloque "A".

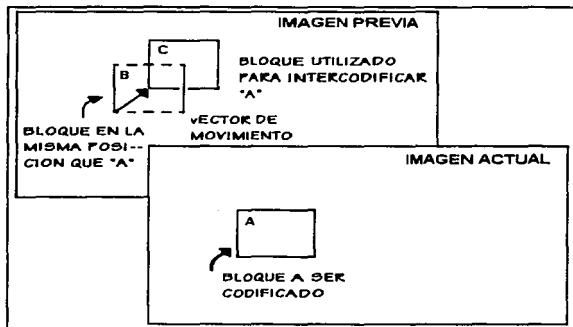


Figura 5.1. Compensación del movimiento.

El uso de vectores de movimiento es opcional en el codificador, donde el cálculo de los vectores de movimiento óptimos es complejo, pero si se requiere en el decodificador, en el cual, la reconstrucción del movimiento es relativamente simple. El estándar H.261 no define todos los aspectos de la codificación y decodificación de la imagen. Si lo anterior es más bien una especificación de interoperabilidad, que garantiza que cualquier codec manufacturado de acuerdo a el estándar sea capaz de comunicarse con otro de diferente marca y fabricante, esto todavía permite una cierta libertad para los fabricantes de ofrecer una mejor calidad, y nuevos desarrollos pueden ser incorporados, (esto es un contraste del estándar de audio G.722, donde el algoritmo de codificación está ya plenamente definido). La estrategia del codificador no está definida. Cuáles bloques serán codificados, con qué tipo de

codificación y con que exactitud dependen exclusivamente del diseñador, así como también dependen de esta la forma en que se realizará el filtrado o interpolación de las imágenes. Aún así, la H.261 permite a dos codecs el comunicarse uno con otro

5.2. ETAPA DE TRANSFORMACIÓN

Si la energía de la imagen de video es de baja frecuencia espacial, es decir con variaciones lentas, entonces una transformada puede ser utilizada para concentrar la energía dentro de algunos coeficientes. El método de transformada elegido por la CCITT es el de la transformada discreta del coseno (DCT) bidimensional de 8×8 , una transformada extensivamente estudiada para compresión de imágenes.

Una fórmula explícita para la DCT bidimensional 8 por 8 puede ser escrita en términos de valores de píxeles, $f(i,j)$, y los coeficientes de la transformada en el dominio de la frecuencia, $F(u,v)$.

$$F(u,v) = \left(\frac{1}{4}\right)C(u)C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i,j) \cos((2i+1)u\pi/16) \cos((2j+1)v\pi/16)$$

Donde,

$$C(x) = \begin{cases} 1/\sqrt{2} & x = 0 \\ 1 & x \neq 0 \end{cases}$$

La salida transformada de la 2D DCT será ordenada de tal manera que el mínimo valor, el coeficiente DC, se sitúe en la esquina superior izquierda y los coeficientes de mas alta frecuencia se sitúen progresivamente a partir del coeficiente DC. Las mayores frecuencias verticales representadas por los número mayores de las filas y las mayores frecuencias horizontales representadas por los números mayores de las columnas.

La transformada bidimensional inversa es escrita de la siguiente manera:

$$f(i, j) = \left(\frac{1}{4}\right) \sum_{u=0}^{i-1} \sum_{v=0}^{j-1} C(u)C(v)F(u, v) \cos((2i+1)u\pi/16) \cos((2j+1)v\pi/16)$$

5.3. CUANTIFICACIÓN

Los coeficientes de la transformada Inversa son cuantificados para reducir su magnitud y para incrementar el número de coeficientes con valor cero. La ecuación para el cuantificador puede ser escrita en términos de el factor cuantificador de el macrobloque, Q algunos veces llamado MQUNT:

Si Q es par, Si Q es impar,

Donde C(u,v) es el coeficiente cuantificado, F(u,v) es el coeficiente de frecuencia DCT y donde ± es positivo para F(u,v)>0 y negativo para F(u,v) El cuantificador inverso está también en términos de par/impar Q. Si Q es par, entonces

Y si Q es impar, entonces

El signo ± indica un mas para un coeficiente positivo, C, y un menos para un coeficiente negativo. El representa la situación contraria.

La cuantificación es la etapa lossy en el esquema de codificación PX64. Si se cuantifica muy "ásperamente", se obtendrán imágenes demasiado "pixeladas", en cambio si se cuantifica muy "finamente", se pueden agregar bits menos utilizados como los de ruido por ejemplo.

5.4. CODIFICACIÓN DE LOS ELEMENTOS CUANTIFICADOS

La codificación del modelo reordena los coeficientes DCT cuantificados dentro de un patrón de zig-zag, con las frecuencias menores primero y al último las frecuencias mayores. El patrón de zig-zag es utilizado para reordenar a los coeficientes de tal manera que se agrupen consecutivamente la mayor cantidad de coeficientes cuyo valor sea cero. Se asume que las frecuencias mas bajas tienden a tener coeficientes mayores y las frecuencias mayores son, por la naturaleza de la mayoría de las imágenes, predominantemente cero.

La figura siguiente muestra el patrón de reordenamiento de los coeficientes DCT cuantificados, donde el primer coeficiente (0,0) es llamado el coeficiente DC y el resto de los coeficientes son llamados AC. Los coeficientes AC son recorridos por el patrón de zig-zag desde la localidad (0,1) a la localidad (7,7).

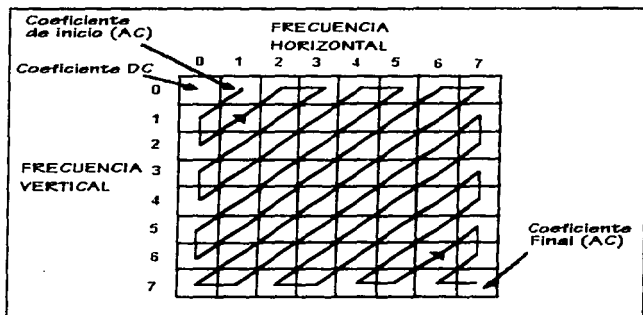


Figura 5.2. El patrón de reordenamiento de los coeficientes DCT cuantificados.

Los coeficientes DC son codificados directamente como un número de 8 bits para los bloques intra-codificados; para la predicción de movimiento de bloques, el coeficiente DC es tratado al igual que los coeficientes AC . El barrido en zig-zag pudiera arrojar códigos con coeficientes AC diferentes de cero conteniendo entre ellos ceros consecutivos. Por lo tanto, se puede codificar los códigos resultantes del barrido en zig-zag mediante una técnica que permita reemplazar los dígitos idénticos consecutivos con el número y tipo, (por ejemplo, 555555599999999 por 7589), esta técnica es conocida como run length. Así, los coeficientes AC son codificados basados en el número de ceros contenidos entre dos coeficientes diferentes de cero. Para combinaciones de ocurrencia frecuente, un código de longitud variable único es utilizado.

La codificación run length inversa traduce la cadena codificada de entrada en un arreglo de salida de coeficientes AC basado en la codificación run length, toma la posición actual en el arreglo de salida y define un número de ceros seguido por el próximo coeficiente diferente de cero.

5.5. CODIFICACIÓN HUFFMAN

En este formato, el archivo para comprimir es examinado y se crea una tabla de referencia con las frecuencias en las que aparecen los caracteres. Caracteres comunes como las letras a, o, e, etc, estos caracteres reciben el byte más corto en ASCII, reemplazando a una asignación de ocho bits por un código de tres bits. Así es como se consigue una compresión de 2:1 o 3:1. Rutinas Huffman aún más rápidas utilizan una tabla preasignada de referencia basada en un idioma especial, pero a cambio de la rapidez se pierde algo de eficiencia en la compresión. Las fuentes Post-Script están codificada con el algoritmo Huffman. Este algoritmo es sólido y sin pérdida de información, pero es lento.

Para el método de PX64 el codificador de Huffman es utilizado para efectuar esta codificación. Una razón para utilizar el codificador de Huffman es que es muy fácil su implementación en hardware. Para comprimir los símbolos de los datos, la codificación de Huffman crea códigos mas cortos para símbolos que se repiten frecuentemente y códigos mas largos para símbolos que ocurren ocasionalmente. Por ejemplo, vamos a codificar un fragmento de la canción que interpreta Michael Jackson: Bad

```

Because I'm bad, I'm bad -- come on
Bad, bad-- really, really bad
You know I'm bad, I'm bad--
you know it
Bad, bad-- really, really bad
You know I'm bad, I'm bad-- come on, you know
Bad, bad-- really, really bad

```

El primer paso para crear el código de Huffman es la creación de una tabla que asigne un valor de frecuencia a cada frase. En la letra de la canción, se ignoran las mayúsculas. Se muestra lo antes mencionado en la tabla siguiente:

Frase	Símbolo	Frecuencia
Because	B	1
I'm	L	6
Bad	B	15
Come on	C	2
It	I	1
Really	R	6
You Know	Y	4

Tabla de Frecuencia de repetición de las palabras en la canción

Inicialmente se designan los símbolos, como los nodos formados por la unión de una hoja y la rama de un árbol. Ahora empezando por los dos nodos de menor peso, agregue el par de menor valor en un nuevo nodo. Por ejemplo, en la carta de frecuencia de arriba, Because e it serán agregados primero. Repita este proceso para una nueva serie hasta que la serie de símbolos este representado por un sólo nodo . El resultado se muestra en la figura siguiente:

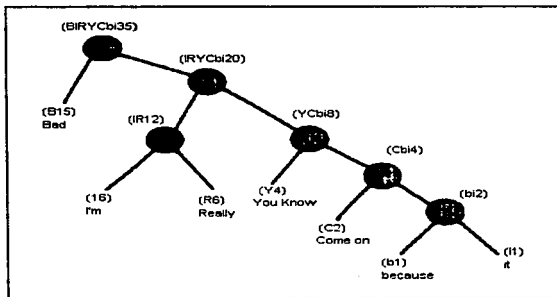


Figura 5.3. Diagrama de árbol para la canción

Un código de Huffman puede ser generado para cada símbolo mediante la asignación de un dígito binario para cada rama. vamos a asignar el dígito binario 1 para cada rama del lado izquierdo y el dígito binario cero para cada rama del lado derecho. El código de símbolos es generado siguiendo el patrón de ramificaciones desde el nodo superior a el símbolo del nodo de la hoja. Para el caso de la palabra bad, el código de Huffman es 1, y para la palabra I'm el código de Huffman es 011.

Una tabla con los valores correspondientes a cada palabra se muestra a continuación.

Frase	Símbolo	Frecuencia	Longitud del Código	Código
Because	B	1	5	0 0 0 0 1
I'm	I	6	3	0 1 1
Bad	B	15	1	1
Come on	C	2	4	0 0 0 1
It	I	1	5	0 0 0 0 0
Really	R	6	3	0 1 0
You Know	Y	4	3	0 0 1

Tabla Código de Huffman para las palabras en la canción

Para decodificar una cadena de bits, inicie desde el nodo superior, y sígala hacia la ramificación izquierda o derecha dependiendo del valor tomado por la cadena de bits; continúe hasta que un nodo de hoja es alcanzado. El símbolo decodificado es el símbolo asociado con esa hoja.

En el ejemplo de arriba los primeros bits en la cadena de bits de salida son:

00001-011-1-011-1-...

La eficiencia de nuestro código puede ser calculado comparando el número de bits requeridos para realizar la letra de la canción. Para el código de Huffman arriba, la longitud de $15(1) + 16(3) + 2(4) + 2(5) = 81$ bits. En comparación, para un código de 3 bits, la longitud es de $35(3) = 105$ bits; y para un código ideal de 7 símbolos la longitud es de $35(\log_2 7) = 98.3$ bits. El código de Huffman comprime la letra en cerca

de 20 por ciento, pero esta figura no incluye el costo de transmitir la tabla inicial del código de Huffman a el decodificador.

5.6. EJEMPLO DE CODIFICACIÓN DE BLOQUES

La figura 5.4 nos da un ejemplo de cómo cada bloque de 8×8 es codificado: En este caso, la codificación intracuadros es utilizada, pero el principio es el mismo para la codificación intercuadros. La figura 5.4a muestra el bloque original a ser codificado. Sin compresión, este tomará 8 bits para codificar cada uno de los 64 píxeles, o un total de 512 bits. primero el bloque es transformado utilizando la transformada discreta del coseno (DCT) bi-dimensional, obteniéndose los coeficientes de la figura 5.4b.

Nótese que la mayoría de la energía está concentrada dentro de los coeficientes situados en la esquina superior izquierda de la matriz de coeficientes. Enseguida, los coeficientes de la figura se cuantifican con un valor $Q = 3$. (El primer término (DC) utiliza un valor $Q = 4$). Esto nos da los valores de la figura 5.4c, los cuales son más pequeños en magnitud que los coeficientes originales y la mayoría de los coeficientes llegan a cero. Entre mayor sea el valor de Q , más pequeños son los valores producidos resultando en una mayor compresión.

Los coeficientes son entonces reordenados, usando el orden del barrido en zig-zag. Todos los coeficientes cero son reemplazados con la cantidad del número de ceros anteriores a cada coeficiente diferente de cero, (a esta cantidad se le denomina RUN). Cada combinación de RUN y valores diferentes de cero, produce un código de longitud variable (VLC) que es enviado al decodificador. El último valor diferente de cero es seguido por un código de fin de bloque (EOB). El número total de bits utilizado para describir el bloque en nuestro ejemplo es 25, una compresión de 20:1.

En el decodificador, (y en el codificador para producir la predicción de la imagen), el valor de Q y los valores diferentes de cero son utilizados para reconstruir los coeficientes cuantificados inversos, los cuales, como se muestra en la figura 5.4e son similares, pero no exactamente iguales, a los coeficientes originales. Cuando estos

coeficientes son transformados (DCT) el resultado se obtiene (figura 5.4f). Nótese que las diferencias entre este bloque y el original son pequeñas.

75	76	77	78	79	80	81	82
77	78	79	80	81	82	83	84
79	80	81	82	83	84	85	86
81	82	83	84	85	86	87	88
83	84	85	86	87	88	89	90
85	86	87	88	89	90	91	92
87	88	89	90	91	92	93	94
89	90	91	92	93	94	95	96

76	76	77	79	80	81	82	83
77	77	78	80	81	82	83	84
79	79	80	81	83	84	85	86
81	82	83	84	85	87	88	88
84	84	85	87	88	89	90	91
86	87	88	89	91	92	93	93
88	89	90	91	92	94	95	95
89	90	91	92	93	95	96	96

Figura 5.4a Bloque original (8X8X8)= 512.

Figura 5.4f Bloque reconstituido.

684	-19	-1	-2	0	-1	0	-1
-37	0	-1	0	0	0	0	1
0	0	0	0	0	0	0	0
-4	-1	-1	-1	-1	0	-1	-1
0	0	0	0	0	0	0	0
-2	0	0	0	-1	0	-1	-1
0	0	0	0	-1	-1	-1	-1
-1	-1	-1	0	-1	0	-1	0

Figura 5.4b Bloque de coeficientes cuantificados.

688	-21	0	0	0	0	0	0
-39	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 5.4e Coeficientes inversamente transformados.

83	-3	0	0	0	0	0	0
-6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

5.4c Niveles de los coeficientes

Nivel	Código
86	01010110
-3	001011
-3	001000011
EOB	10

Longitud total del código= 25

5.4d Coeficientes en orden "zig-zag" y cuantificados y longitud de la variable codificada

CONCLUSIONES

CONCLUSIONES

Por mucho tiempo medios como el teléfono, el fax o un módem han sido utilizados para satisfacer necesidades de comunicación individual o por instituciones y empresas, hoy en día, la Videoconferencia ofrece una solución a esa necesidad de comunicación ya que permite transmitir información visual y sonora entre zonas distantes evitando de esta forma, los gastos y pérdida de tiempo que implican el traslado físico de la persona o personas haciendo posible que los gastos se reduzcan y la calidad de la información transmitida (señales visuales y audiovisuales) sea de mejor calidad. Es por eso que la Videoconferencia es el segmento en el área de las telecomunicaciones que observa un mayor crecimiento.

Con la introducción de una confiable y efectiva videocompresión, el video puede ahora almacenarse en archivos mucho más pequeños y con muy poca degradación en la imagen, además el tiempo y la velocidad requeridos para transferir información de un equipo a otro tiende a disminuir.

El éxito en la implementación y estandarización de las técnicas de compresión en un equipo de Videoconferencia va a depender de las necesidades y aplicaciones que requiera el usuario, entre más bajo el rango de información a transmitir, será menor la calidad y más bajo el costo de transmisión. Se debe tener en cuenta, que casi la mayoría de la información visual puede aceptar cierta pérdida en la calidad de la imagen, ya que ésta no tiene una importancia muy significativa y trae consigo mayor rapidez en su transmisión y un menor costo. El efectuar una compresión sin pérdida de calidad significa un lujo que muy pocos se pueden dar.

Falta todavía mucho por hacer en cuanto a los Sistemas de Videoconferencia se refiere, pero el primer paso ya está dado, el siguiente es seguir investigando y aplicando nuevas tecnologías y mejores esquemas de compresión, para el mejoramiento de este servicio, tanto en la calidad como en el costo.

BIBLIOGRAFÍA

BIBLIOGRAFÍA

TELEDATA, Fundamentos de Comunicaciones Digitales, México D.F., 1993

A. BAXES Gregory, Digital Image Processing. Principles and Applications, EUA, Ed. John Wiley & Sons, Inc., 1994, 452 p.

ANDREW S. Tanenbaum, Redes de Ordenadores De. Prentice-Hall Hispanoamericana S.A., 759 p. (2ª edición)

D. WATKINS, Christopher, SADUN Alberto y MARENKA Stephen, Modern Image Processing: Warping, Morphing and Classical Techniques, De. Academic Press Professional, EUA, 1993, 234 p.

FITZGERALD Jerry y TOM S. Eason, Fundamentos de Comunicaciones de Datos, Ed. Limusa, México D.F., 1989, 275 p. (4ª edición).

J. BROWN Telecomunicaciones, De. Esasa, 435 p.

JÄHNE Bernd, Digital Image Processing, Concepts, Algorithms and Scientific Applications, De. Springer-Verlag, Germany, 1991, 383 p.

SCHAPHORST Richard, Videoconferencing and Videotelephony, Technology and Standards, Ed. Artech House, 1996, 199 p.

DIRECCIONES CONSULTADAS EN INTERNET:

<http://video.comserv.ipn.mx>

<http://lisa.ee.nd.edu/rls/old/research/compression.html>

<http://www.columbia.edu/cu/moment/110195/video.html>

<http://www.worldguide.com/Tech/videocapture.html>

http://www.dse.doc.ic.ac.uk/nd/surprise_96/journal/vol11/srd2/article1.html

<http://www.e-town.com/news/articles/MPEGRC.html>