



Compresión de audio y vídeo en Sistemas Multimedia

❖ *Luis Torres, Asunción Moreno, Enrique Masgrau, Antoni Gasull*

El objetivo de este artículo es proporcionar una visión amplia de los sistemas y standards actuales de almacenamiento de audio e imagen digital, con especial énfasis en los sistemas que tendrán una mayor incidencia en la electrónica de consumo, tales como los sistemas multimedia. Después de una breve introducción, se abordan en el presente artículo, los standards utilizados en compresión de audio e imagen y su aplicación en los diversos sistemas de almacenamiento. A continuación se comentan las diversas tecnologías utilizadas en los sistemas multimedia, donde se describen las características de los principales sistemas y su interacción con los principales standards.

Las tecnologías de almacenamiento y comunicación audio-visual de la información junto con los nuevos sistemas de presentación proporcionados por las pantallas de ordenador, están sufriendo una evolución muy rápida en los últimos años.

En muchos casos el profesional que no trabaja directamente en este campo, desea conocer los últimos avances producidos en dichas tecnologías, por propio interés o por la posibilidad de introducir dichos avances en su propio entorno de trabajo. El problema con que se encuentra es

que la información es tan amplia y tan dispersa que muchas veces es difícil conseguir una visión resumida y a la vez completa de los nuevos desarrollos. Por otra parte, el problema se agrava cuando confluyen en un mismo entorno tecnologías aparentemente tan diferentes como los sistemas analógicos de televisión, los formatos digitales propuestos en televisión de alta definición, los formatos de audio utilizados en sistemas de transmisión digitales, los sistemas multimedia de almacenamiento de audio, imágenes y gráficos, las pantallas de ordenador capaces de visualizar imágenes de alta definición y los nuevos standards de compresión de imagen y video.

Formatos de imagen, vídeo/audio

Las señales de imagen y sonido se adquieren, manipulan, almacenan, transmiten y en su caso, visualizan en una gran variedad de formatos. Ejemplos de dichos formatos en el entorno de la televisión, son las señales analógicas PAL, SECAM y NTSC utilizadas en la transmisión de TV, las señales analógicas Y, R-Y, B-Y empleadas en la industria de la TV, o las R, G, B empleadas en los ordenadores, los nuevos formatos digitales de TV especificados en la recomendación 601 del CCIR, el sistema de televisión de alta

**"Las señales de imagen y sonido
se adquieren, manipulan, almacenan,
transmiten, y visualizan
en gran variedad de formatos"**

definición MUSE propuesto por los japoneses, el sistema HDMAC estudiado por los europeos, o los nuevos sistemas digitales propuestos por los americanos. Por otra parte, hay que tener también en cuenta el standard JPEG (Joint Photographic Experts Group) propuesto para la codificación de imágenes fijas, la recomendación H.261 del CCITT para sistemas de videotelefonía y el standard MPEG (Moving Picture Experts Group) propuesto en aplicaciones de compresión de secuencias de imagen (denominadas vídeo en este artículo) y sonido en el entorno multimedia.

En el campo de audio, la aparición del Compact Disc (CD) a principios de los 80 ha supuesto la validación del audio digital de alta calidad frente a los sistemas analógicos previamente desarrollados. Junto a los sistemas de audio digital ya conocidos CD y DAT (Digital Audio Tape), están surgiendo dos nuevos sistemas con calidades totalmente comparables. El Digital Compact Cassette (DCC) desarrollado por Philips y el Mini Disc desarrollado por Sony capaz de reproducir 74 minutos de sonido digital en un disco de 74 mm (la mitad que el tradicional CD)-

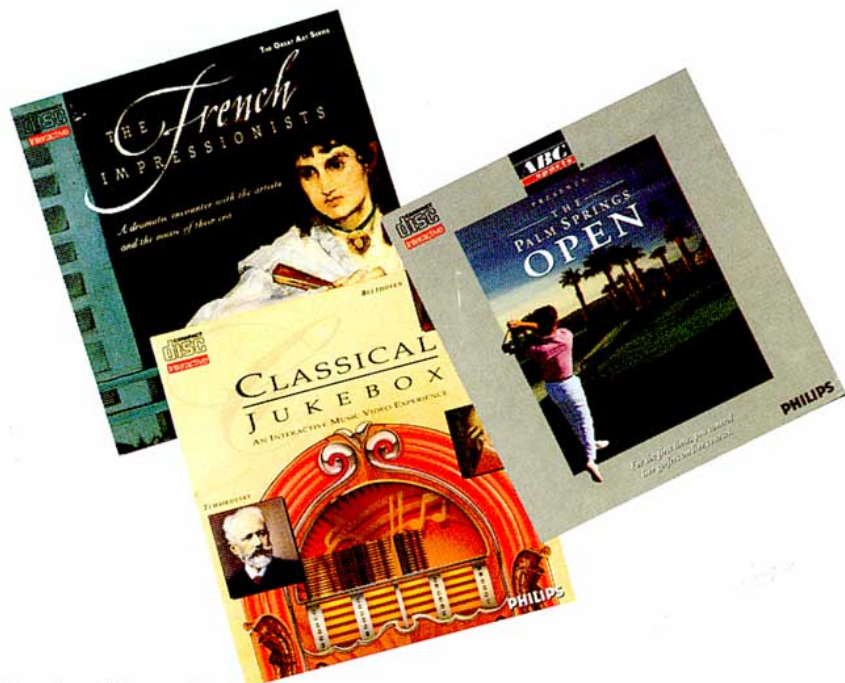
Junto a estos sistemas, han surgido aplicaciones que requieren audio digital de alta calidad. Ejemplos de estas aplicaciones son: telefonía digital, terminales ISDN para teleconferencia, sistemas de almacenamiento de audio, televisión de alta definición HDTV y servicios CD-ROM multimedia.

Las aplicaciones multimedia presentan también una variedad de formatos, como son el DVI (Digital vídeo Interactive) de Intel, el CDI (Compact Disc Interactive) de Philips Consumer Electronics, el CDTV (Compact Disc Televisión) de Commodore Electronics y el PhotoCD de Kodak.

La gran abundancia de formatos de imagen, vídeo y audio se unen a la gran cantidad de aplicaciones existentes que utilizan el ordenador como sistema de tratamiento y almacenamiento de información y de visualización de imágenes. Entre ellas destaquemos las aplicaciones médicas, los entornos multimedia para creación de grandes bancos de datos utilizados por muchos usuarios, el diseño de sistemas de ingeniería por ordenador, el tratamiento y producción de gráficos y un largo etc. En muchas aplicaciones, es necesario (o al menos útil) la integración e interactividad de distintos sistemas de imagen. Por ejemplo, son cada vez mas numerosos los sistemas de ordenador que son capaces de tratar

y visualizar secuencias de imágenes de TV convencional, o los receptores de TV convencional que son capaces de visualizar vídeo, gráficos, texto o sistemas de animación, almacenados en discos compactos.

El problema que se presenta es la coexistencia de los numerosos formatos de imagen y vídeo y la diversidad de sistemas de tratamiento, almacenamiento y visualización de imágenes. Se han propuesto diferentes soluciones para la armonización de los diferentes sistemas. Entre ellas, la más extendida es el diseño e implementación de arquitecturas abiertas en sistemas que deberían ser escalables y extensibles. Los sistemas escalables son aquellos que tienen la capacidad de usar diferentes resoluciones, frecuencias temporales, y márgenes dinámicos de intensidad y color con objeto de que, por ejemplo, pantallas diseñadas para satisfacer un conjunto de parámetros, sean capaces de adaptarse a otros tipos diferentes de especificaciones. Los sistemas extensibles son aquellos que tienen la capacidad de ampliación de los diferentes formatos, para tener en cuenta nuevos requerimientos y aplicaciones.



Ya existen CD interactivos

Compresión de datos

Cuando se integran en un solo sistema tecnologías de audio e imagen, es necesario prestar atención también a consideraciones de tiempo de ac-

ceso a la información, velocidad de transferencia de datos y capacidad de almacenamiento. En el caso de sistemas multimedia es particularmente importante la velocidad de transferencia de datos, ya que este parámetro fijara la máxima velocidad de presentación de una secuencia de vídeo. Para sistemas basados en arquitecturas PC, se puede considerar como velocidad media de transferencia de datos la de 1.2 Mbits/seg. Si dicho sistema tiene que proporcionar audio de alta calidad, es preciso conseguir una velocidad de 1.2 Mbits/seg. Con estas cifras, parece que no existe ningún problema en el sistema. El problema surge cuando se desea transferir audio y vídeo al mismo tiempo! Si tenemos en cuenta que la señal de vídeo puede requerir más de 10 Mbits/seg. para una calidad aceptable, está claro que el sistema debería de poseer capacidades de compresión y descompresión de audio y vídeo.

Con respecto a la capacidad de almacenamiento, los discos de ordenador habitualmente utilizados en ordenadores personales pueden almacenar hasta 40 Mbytes de información. Parece claro, que con las necesidades en este sentido de las señales de audio y vídeo se requieren discos de mayor capacidad y/o sistemas de compresión de datos. La tendencia actual en sistemas multimedia es

COMUNICARSE ES COMO RESPIRAR.



Los Disco Laser multiestándar como el modelo V 380 de Philips aceptan directamente todos los tipos y tamaños de discos Láser y LaserVision PAL y NTSC, además de discos de CD Audio. No se necesita ningún tipo de adaptador. El modo de operación correcto se selecciona de forma automática.

a la utilización de discos CD-ROM con capacidad de 600 Mbytes. La ventaja de dichos discos es que son fácilmente extraíbles y se pueden cambiar con lo que se adquiere mayor capacidad de almacenamiento. Se espera que en un futuro cercano,

el precio de los discos de ordenador y el de los CD-ROM sea comparable lo que facilitara su introducción en el mercado multimedia.

Las señales de vídeo requieren regímenes binarios de 10 Megabits/seg para vídeo con calidad de

"Para facilitar el intercambio de datos comprimidos entre diferentes usuarios, se han desarrollado diferentes standards"

CADA DIA, SIN DARTE CUENTA, LO HACES MILLONES DE VECES.

CADA DIA, SIN DARTE CUENTA, LAS PALABRAS BROTAN DE TI. MILLONES Y MILLONES. CADA DIA, TAMBIEN, ERICSSON ESTA AHI, A TU LADO, PONIENDO AL SERVICIO DE LA COMUNICACION TODOS LOS MEDIOS TECNICOS Y HUMANOS: DECENAS DE MILES DE PERSONAS



TRABAJANDO, CENTRALES TELEFONICAS, CENTRALITAS, TELEFONOS MOVILES, SOFISTICADOS SISTEMAS Y REDES A LO LARGO Y ANCHO DEL MUNDO ENTERO. AUNQUE, COMO EL AIRE QUE RESPIRAS, CADA VEZ QUE LOS UTILIZAS NO SEAS CONSCIENTE DE ELLO.

ERICSSON 

IDEAS QUE PERMITEN TRANSMITIR LAS IDEAS

distribución y hasta mas de 100 Megabits/seg. para señales de televisión de alta definición. Incluso para la representación de imágenes fijas, se necesita una gran cantidad de información. Por ejemplo, una imagen en color de resolución 1000x1000 pixels y 24 bits/pixel ocupa 3 Megabits de almacenamiento si no se utilizan técnicas de compresión. Este tipo de imagen no cabe en un diskette de alta densidad de 1.2 Megabytes. Un vídeo clip de 10 segundos, digitalizado a partir de una emisión de TV necesita mas de 200 Megabytes de almacenamiento. Para disminuir la gran cantidad de almacenamiento necesitado, se utilizan técnicas de compresión de imagen.

Las técnicas de compresión de imagen se basan en las altas redundancias presentes en las imágenes fijas y en movimiento y en las no linealidades del sistema visual humano. La compresión en el espacio se denomina "compresión intracuadros" y la compresión en el dominio temporal se denomina "compresión intercuadros". Las técnicas de compresión pueden ser con pérdidas o sin pérdidas. Los métodos con pérdidas alcanzan mayor compresión a expensas de no poder recuperar de forma perfecta la imagen original. Los tres standards actualmente propuestos para compresión de imagen se basan en métodos de compresión predictivos y transformados.

Las tasas de compresión de la señal de audio no son tan elevadas como las obtenidas en vídeo debido a que no existe un paralelismo de compresión intracuadros e intercuadros. Fijándonos en la aplicación concreta donde va a utilizarse la compresión de audio, el ancho de banda de la señal a comprimir juega un papel importante. En este sentido, es muy distinto comprimir señal de voz, donde limitando el ancho de banda entre 300 Hz y 3.4 KHz se obtiene una señal perfectamente inteligible que mantiene las características de articulación, que señal de música donde el ancho de banda debe ser mayor.

La compresión se realiza eliminando redundancias en la señal de audio. Dos son los mecanismos para eliminar estas redundancias: la utilización de métodos predictivos como el ADPCM (Adaptive Differential PCM), DM (Delta Modulation), ADM (Adaptive DM) y CELP (Code Excited Linear Prediction), o la utilización de métodos transformados como los bancos de filtros o las transformaciones (análisis frecuencial). Todos los métodos de compresión de audio, se basan en una de estas técnicas o en una combinación de ambas.

Con objeto de acelerar el desarrollo del hardware de compresión de audio e imagen y para facilitar el intercambio de datos comprimidos entre los diferentes usuarios, se han desarrollado diferentes standards. A continuación se describen los principales standards de almacenamiento en audio, imagen y vídeo así como los mas importantes formatos utilizados actualmente en sistemas de almacenamiento, particularmente en el entorno multimedia.

Standards de imagen

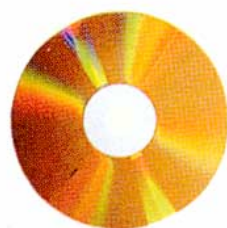
STANDARD JPEG: El standard JPEG es un método de compresión desarrollado para imágenes fijas bajo los auspicios del ISO. Se puede utilizar en aplicaciones tales como: artes gráficas, publicidad, facsímil en color, transmisión de imágenes de prensa, almacenamiento de imágenes médicas, etc. El algoritmo se basa en la división de la imagen en subimágenes de 8x8 y la posterior transformación de las subimágenes mediante la transformada Coseno Discreta. Los coeficientes transformados se cuantifican y se codifican posteriormente mediante códigos de Huffman. El proceso inverso se realiza en el proceso de descompresión de la imagen. El standard se ha desarrollado para dar tasas de compresión desde 0.25 bits/pixel (factor de compresión 16) con una calidad subjetiva aceptable, hasta 1.25 bits/pixel (factor de compresión 6.4) con una calidad visual indistinguible del original.

Varios fabricantes ya han desarrollado circuitos integrados que incorporan el standard JPEG. Entre ellos destaquemos, C-Cube Microsystems Inc. y LSI Logic Corp. Muchos fabricantes de placas incorporan dichos chips en sistemas integrados en PC y estaciones de trabajo.

STANDARD MPEG: El standard MPEG especifica la compresión de señales de vídeo a velocidades de 1.5 Megabits/seg. y de señales de audio a velocidades de 64, 128 y 192 kilobits/seg. en aplicaciones de almacenamiento en forma digital. El standard también proporciona sistemas de sincronización y multiplexado para señales multiplexa-

"Los chips basados en el decodificador estarán antes en el mercado que los basados en el codificador"

Algunos equipos, como el CDV 496 de Philips no hace diferencias, reproduce todos los discos, tanto CD Audio como Laser Disc.



Laser Disc LP
(30 cm/12 pulgadas)
2x60 minutos de vídeo y sonido CD.
Es el disco ideal para la reproducción de conciertos y largometrajes



Laser Disc EP
(20 cm/8 pulgadas)
Grabado en una o ambas caras, este disco contiene hasta 20 minutos de vídeo y sonido CD por cara. Ideal para la presentación de videoclips y dibujos animados



Laser Disc Single
(12 cm/5 pulgadas)
Este single contiene 6 minutos de vídeo y sonido CD además de 20 minutos adicionales de sólo sonido CD. Ideal para la presentación de videoclips y dibujos animados



Compact Disc
(12 cm/5 pulgadas)
Contiene hasta 74 minutos de sonido CD



Compact Disc Single
(8 cm/3 pulgadas)
Ofrece 20 minutos de sonido CD

das comprimidas de audio y vídeo. El MPEG es genérico, es decir, es independiente de la aplicación y se puede aplicar a sistemas de almacenamiento tales como el CD-ROM (Compact disc ROM), discos Winchester, discos ópticos, DAT (Digital Audio Tape) y en sistemas de comunicación como la ISDN (Integrated Services Digital Network) y redes de área local (LAN). Es importante hacer notar que los codificadores y decodificadores no se especifican en el standard. Es decir, un codificador es un codificador MPEG, si puede producir una cadena legal de bits MPEG. Un decodificador es un decodificador MPEG, si puede decodificar de forma satisfactoria una cadena de bits MPEG.

El algoritmo MPEG utiliza métodos predictivos con interpolación y compensación de movimiento para reducir la redundancia temporal y métodos transformados basados en la transformada Coseno Discreta para reducir la redundancia en el dominio espacial. La cuantificación de los coeficientes transformados tiene en cuenta el sistema visual humano, de tal forma que los coeficientes correspondientes a las frecuencias altas se cuantifican de forma mas grosera que los coeficientes correspondientes a las frecuencias bajas, puesto que la percepción subjetiva del error de cuantificación varía con la frecuencia.

El MPEG todavía no se ha convertido en standard y no se espera que se adopte hasta como mínimo dentro de unos dos años. Sin embargo, la parte de vídeo se dejó prácticamente terminada en Septiembre de 1990 y no se esperan grandes modificaciones. A diferencia del IPEG, ningún fabricante de circuitos integrados ha comercializado todavía el MPEG. La razón es clara. La complejidad del desarrollo es tan grande, que ningún fabricante se arriesgara a producir circuitos integrados hasta que el standard este totalmente diseñado. La tendencia es a realizar un chip que sea adaptable a través de una arquitectura microprogramable, que permita introducir las posibles modificaciones que se produzcan en el standard. Ejemplo de esta filosofía es C-Cube Microsystems Inc que ya está trabajando en este tipo de arquitecturas. SGS-Thomson y LSI Logic Corp. han anunciado también chips basados en el MPEG. Por otra parte, debido a la mas alta complejidad del codificador con respecto al decodificador, los chips basados en el decodificador estarán antes en el mercado que los basados en el codificador.

El standard MPEG es extraordinariamente complejo y en este sentido, es de destacar un fenómeno que se está produciendo entre los usuarios de sistemas de compresión de imagen y vídeo. Dado que ya se disponen circuitos integrados y placas de desarrollo basadas en el JPEG, este standard se está utilizando para comprimir se-

¿ QUE ES EL DCC ?



La cassette digital (DCC) es la sucesora natural de la famosa cassette analógica, creada por Philips en 1962, que es uno de los productos que mayor éxito han tenido en la historia de la grabación y reproducción del sonido, hasta el extremo que es el segundo producto, de consumo, de mayor venta en el mundo tras la bombilla.

El DCC, cuya puesta a punto ha requerido varios años de investigación, aporta la calidad del sonido digital al más popular de los soportes: la cassette de audio. Totalmente portátil, puede utilizarse en el hogar, en el automóvil, en la playa o en el tren. Los nuevos DCC serán totalmente compatibles con las "musi-cassettes" analógicas existentes, lo que permite a los usuarios reproducir sus cintas antiguas en los nuevos aparatos, así como realizar grabaciones sobre cintas vírgenes. Esta compatibilidad constituye un factor clave para el éxito del nuevo sistema.

Las nuevas cassettes digitales serán similares a las existentes, si bien incorporan ciertas características distintivas que permiten, por ejemplo, una mejor información gráfica sobre el contenido de la grabación, selección de pista, función auto-reverse, limitaciones para la obtención de copias ilegales de las grabaciones. El tiempo de grabación-reproducción es de 90 minutos.

Entre las características técnicas del sistema destaca el revolucionario método de codificación que se utiliza en el mismo, denominado PASC (Precisión Adaptive Sub-band Coding) que consigue un registro de sonido de hasta 18 bits, tomando como modelo de referencia el comportamiento del oído humano.

Precisamente, el método de codificación de la señal ha sido uno de los elementos clave para conseguir la grabación digital de la cinta, manteniendo una velocidad de desplazamiento igual a la de la cassette analógica.

El lanzamiento del sistema DCC, incluyendo aparatos y cintas —tanto vírgenes como pregrabadas— está siendo apoyado por numerosos grupos industriales en todo el mundo. Polygram, por ejemplo, una de las compañías discográficas más importantes del mundo, editará los primeros 500 títulos para el lanzamiento mundial, que se producirá en el próximo mes de Septiembre.

cuencias de vídeo (recuérdese que el JPEG es un standard de compresión de imágenes fijas). La compresión conseguida con el JPEG cuando se utiliza para comprimir vídeo, es de alrededor de una tercera parte de la proporcionada por el MPEG para la misma calidad de imagen.

Para concluir con la descripción del MPEG, es importante destacar que ya existe una continuación del standard, conocido como MPEG II, donde el objetivo es conseguir un esquema de compresión no para sistemas de almacenamiento, sino para sistemas de transmisión de TV digital en el margen de los 10 Megabits/seg.

Standards de audio

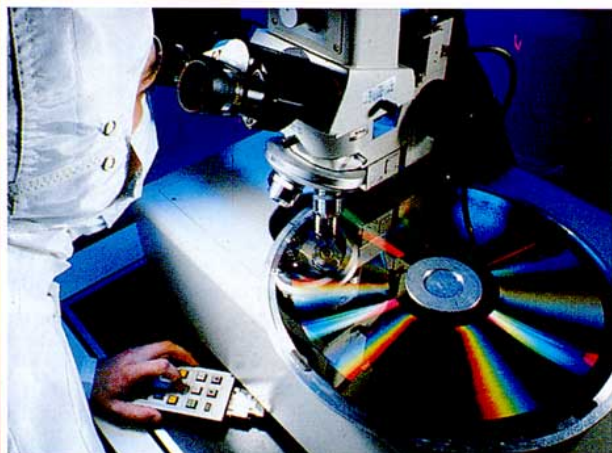
Codificación de voz. Ancho de banda 3.4

KHz PCM: 64 Kb/s (adoptado por el CCITT. Norma G711). La cuantificación es "transparente", es decir no se observa degradación respecto a la señal original.

ADPCM: 32 Kb/s (CCITT G721). Se cuantifica la diferencia entre la señal original y una produc-

"La tendencia es clara hacia sistemas de almacenamiento y tratamiento digital de la información multimedia"

¿ QUE ES EL CDI ?



El Compact Disc Interactivo (CD-I), es el sistema de entretenimiento en el hogar, basado en el Compact Disc, que combina la calidad de sonido del Compact Disc con vídeo, textos, gráficos, animación y capacidades interactivas.

El CD-I permite al usuario interactuar realmente con su televisor, de una forma simple y libre de complicaciones. Usando una unidad por control remoto el usuario puede escoger: repintar un dibujo, visitar un museo, crear una novela o canción, reclamar información sobre un artista o una canción mientras escucha la música a tomar fotografías a su gusto mientras aprende las técnicas de la fotografía.

El CD-I es compatible con las tecnologías existentes, incluyendo CD de sonido y gráficos, así como tecnologías futuras como foto y vídeo CD y televisión de Alta Definición.

Los aparatos CD-I se conectan fácilmente a cualquier sistema de televisión y vídeo. El disco se carga de la misma forma que el disco de CD audio y es controlado por medio de un mecanismo señalizador como un Thumbstick de control remoto, un ratón o incluso de un teclado.

En la pantalla aparecen distintas áreas con palabras y símbolos que indican las posibles órdenes de selección. Situando el cursor en una de estas áreas y presionando el botón de acción permitirá acti-

var la orden escogida.

La mayor parte del software para el CD-I permitirá al usuario interrumpir la secuencia para memorizar una particular elección, regresar al paso anterior, pedir información más detallada o, quizás, solicitar al explicación en otro idioma.

El CD-I utiliza un controlador multimedia (MMC), que contiene los circuitos decodificadores de sonido, imagen y texto y las memorias necesarias para repro-

ducir los discos CD-I. También contiene el poderoso cerebro, microprocesador M 68070, que controla la reproducción y ejecuta el programa interactivo.

El CD-I procesa la información de un disco óptico de cinco pulgadas y 650 megabytes, que le permite:

- Más de 25.000 hojas de texto mecanografiado.
- Más de 7.000 imágenes con calidad fotográfica.
- 72 minutos de imágenes animadas.
- 19 horas de discurso.
- Capacidad de vídeo para más de 16 millones de variaciones de color.
- 4 planos de efectos visuales.

El CD-I está adaptado para cuatro calidades de sonido:

- CD-DA (Compact Disc sonido digital)
- Hi-Fi (Alta Fidelidad)
- Mid-Fi (Media Fidelidad)
- PalabraEl CD-I también ofrece cuatro modos básicos de codificación para vídeos.
- DYUB
- RGB
- CLUT
- Run Length

El CD-I es capaz de almacenar comentarios en múltiples lenguajes, hasta 16 pistas paralelas con un nivel de sonido de discurso.

El CD-I fue presentado en los EE.UU. y su introducción en el mercado europeo tendrá lugar en el curso de este año.

ción de la misma basada en muestras anteriores. La calidad es análoga a la anterior pero la cuantificación requiere mayor carga computacional.

LD-CELP (Low Delay CELP): 16Kb/s (próximo a ser adoptado por el CCL). Es un sistema más complejo que los anteriores, donde la señal codificada se genera por medio de un filtro variante con el tiempo que se excita por una secuencia código que minimiza una medida de distorsión perceptual. Se obtiene una calidad similar a las anteriores.

CELP: 8Kb/s. Utilizado básicamente en comunicaciones de radio celular, logra una calidad

de "comunicación", es decir, se observan ciertas degradaciones pero la señal es perfectamente útil para comunicaciones.

CELP: 4.8 Kb/s. Utilizado en comunicaciones móviles vía satélite MSAT y en aplicaciones marítimas internacionales vía satélite INMARSAT, obtiene una calidad inferior a los anteriores pero altamente inteligible. Con este método puede almacenarse en un chip de 16 Mb prácticamente una hora de habla.

Codificación de audio de 7 KHz:

El ancho de banda se limita entre 50 Hz y 7 KHz. Su aplicación fundamental está en señal de voz a nivel de conversación. Sobre calidad telefónica aporta mayor naturalidad debido a la inclusión de bajas frecuencias y permite diferenciar sonidos de alta frecuencia tales como s y f. Podemos asimilar la calidad de estos sistemas a radiodifusión AM.

El standard de codificación por el CCITT (G 722) es a 64 Kb/s, admitiendo también 56 Kb/s y 48 Kb/s, y se desarrolló originalmente para teleconferencia por ISDN. El sistema G722 está basado en la codificación de dos subbandas, cada una de ellas por un sistema ADPCM. La banda inferior es ligeramente superior a la de telefonía digital. Se realiza una asignación de bits distinta en la banda superior (6 bits por muestra) que en la inferior (2 bits por muestra). Es robusto frente a errores y su sencillez permite ser implementado en un chip. Las nuevas alternativas consisten en extensiones de los codificadores CELP y LD-CELP a esta aplicación para reducir la velocidad.

Codificación de audio de 15 KHz

El ancho de banda se sitúa entre 20 Hz y 15 KHz. Es útil para música, obteniéndose calidades próximas a lo que conocemos como FM.

Para la codificación se emplea NICAM-PCM (PCM con Nearly Instantaneous Compounding) a 384 Kb/s y ADM, sistema muy similar al ADPCM explicado anteriormente, a 330 Kb/s.

Codificación de audio de 20 KHz

El sistema Compact Disc - Digital Audio (CD-DA) almacena audio a una velocidad de 700 Kb/s por canal, (16 bits PCM a una frecuencia de muestreo de 44.1 KHz). Es el sistema que mejor calidad obtiene ("super HIFI").

En banda ancha de 20 KHz, la compresión de datos se realiza teniendo en cuenta un modelo auditivo que puede resumirse en estos términos: El oído únicamente percibe sonidos por encima de un umbral de audición. Los sonidos con mas

energía enmascaran dinámicamente sonidos más débiles de frecuencias cercanas, cambiando dinámicamente el umbral de audición. Los sonidos por encima del umbral de audición deben ser cuantificados con suficientes bits, siendo innecesaria la cuantificación de los sonidos inaudibles.

Basados en esta filosofía se han desarrollado los sistemas:

PASC (Precisión Adaptive Subband Coding) para aplicación en DCC (Digital Compact Cassette) donde la asignación de bits se realiza dividiendo la señal en 32 subbandas y alcanzando una compresión de 4:1 respecto a PCM

ATRAC (Adaptive Transform Acoustic Coding) con aplicación en Mini Disc, donde la asignación de bits se realiza en el dominio transformado (análisis frecuencial) logrando una compresión de 5:1.

MUSICAM: Basado en un análisis por subbandas, trabaja con velocidades de muestreo de 48 KHz, 44.1 KHz y 32 KHz pudiendo trabajar en un margen de 64 a 192 Kb/s.

MPEG: Basado también en análisis por subbandas, tiene la opción adicional de un análisis por transformada. Trabaja con velocidades de muestreo de 48 KHz, 44.1 KHz y 32 KHz pudiendo trabajar en un margen de 32 a 192 Kbits con distintas calidades.

Sistemas y tecnologías Multimedia

Los sistemas multimedia representan aquellos sistemas que integran en una sola plataforma las imágenes, el audio, los textos, los gráficos y combinaciones entre todos ellos bajo el común denominador de la tecnología digital. Aunque si bien es verdad que existen sistemas de almacenamiento analógico como los videodiscos que pueden actuar como soporte de sistemas multimedia, la tendencia es clara hacia sistemas de almacenamiento y tratamiento digital de la información multimedia. Bajo esta perspectiva, solo se tratarán los sistemas con soporte digital. De entre los más utilizados, se han elegido el DVI y el CDI.

Digital vídeo Interactive - DVI

La tecnología DVI, desarrollada por Intel e IBM, integra imágenes, audio, textos y gráficos en sistemas basados en ordenadores. La tecnología DVI no es un sistema, es una tecnología lo que significa que se pueden construir diferentes sistemas utilizando dicha tecnología.

La tecnología DVI consiste en cuatro elementos fundamentales: 1) Un circuito integrado VLSI dedicado que es el corazón del sistema vídeo, 2) Una especificación para un interface software,

3) Formatos de ficheros de audio y vídeo y 4) Algoritmos de compresión y descompresión. Utilizando estos cuatro elementos, la tecnología DVI se puede implementar en cualquier ordenador que tenga suficiente potencia de cálculo y capacidad de almacenamiento. Hasta el momento, el DVI se ha implementado en sistemas IBM PS/2 y ordenadores PC AT compatibles. Intel, está trabajando en la implementación del DVI en otras plataformas hardware.

El coprocesador de vídeo es el denominado i750TM. Este procesador consta de dos circuitos integrados, el 82750PA denominado procesador de pixels y el 82750DA denominado procesador de visualización. El procesador de pixels ejecuta 12.5 millones de instrucciones por segundo y puede descomprimir en tiempo real imágenes vídeo. Dicho procesador, puede tratar 72 minutos de vídeo almacenados en CD-ROM utilizando toda la pantalla de visualización. El tiempo de visualización se puede incrementar si se reduce la definición de la pantalla o la definición temporal de las secuencias de vídeo. El procesador de visualización tiene una definición de 256x240 pixels para imágenes de vídeo y una definición máxima de 1024x512 para imágenes fijas.

Un sistema de desarrollo típico DVI consiste de un sistema PS/2 o PC AT compatible, placas de digitalización de vídeo, discos de almacenamiento CD-ROM, sistemas de almacenamiento de gran capacidad (actualmente se utilizan unidades de cinta) y un monitor analógico RGB dotado de amplificador de audio y altavoces.

Existen dos sistemas de compresión de señales de vídeo. El primer sistema, denominado RTV (Real Time vídeo) se puede utilizar por el usuario en la plataforma DVI. El segundo sistema, denominado PLV, puede ser utilizado por el usuario, únicamente como un sistema decodificador en tiempo real. EL RTV comprime y almacena en tiempo real, secuencias de vídeo adquiridas en el propio sistema de desarrollo. El algoritmo de compresión trabaja a una cadencia de 10 imágenes por segundo (se comprime una imagen de cada tres). Una imagen RTV es absoluta, es decir, no es relativa con respecto a la imagen anterior. Esta propiedad permite una gran facilidad de implementación de operaciones de edición de vídeo. Cuando la aplicación que se está implementando en el DVI llega a la etapa de diseño final, en la que se requiere vídeo de alta calidad, el usuario envía las imágenes desarrolladas en el RTV, a un servicio de compresión DVI para su posterior compresión con el algoritmo PLV. Una vez comprimidas, el usuario recibe las imágenes digitalizadas y substituye sus imágenes RTV por las PLV en su sistema de desarrollo DVI. A partir de este momento, el usuario tiene capacidad con su propio

"El futuro no distinguirá entre un ordenador que incorpore un sintonizador de televisión y un televisor que incorpore CPU muy sofisticadas"



Luis Torres

Ingeniero de Telecomunicación por la Escuela Técnica Superior de Ingenieros de Telecomunicación de Barcelona en 1977. Se doctoró en 1986, en la Universidad de Wyoming, Estados Unidos. Desde 1988 es Profesor Titular del Departamento de Teoría de la Señal y Comunicaciones de la Universidad Politécnica de Cataluña donde imparte la asignatura de Televisión. Sus áreas de interés son el Procesado y la Codificación de Imagen.



Asunción Moreno

Dr. Ingeniero de Telecomunicación por la UPC desde 1987.

Entre 1980 y 1984 trabajó en la fábrica de televisión de Philips. En la actualidad es profesor Titular en la E.T.S.I. Telecomunicación de Barcelona. Sus áreas de interés son Reconocimiento de voz y Codificación.

sistema de descomprimir en tiempo real las imágenes codificadas en el sistema PLV. El algoritmo PLV es original de los propietarios del DVI y no se adapta, por ahora, al standard MPEG. Por otra parte, el sistema es capaz de descomprimir imágenes fijas comprimidas con el standard JPEG. Intel ha anunciado la presentación de un circuito integrado con el standard MPEG para finales del año 1992. Por otra parte, el objetivo final de Intel en lo que se refiere a compresión de imagen, es que el esquema RTV alcance calidad cercana a la de la televisión y que el esquema PLV alcance una calidad comparable a la televisión de alta definición.

El DVI utiliza un sistema de compresión de audio, que consta de cuatro niveles distintos de calidad: Modo CD- Digital Audio: Ofrece el sistema CD-Audio Format (44.1 KHz, 16 bits estéreo) sin compresión. Modo HI-FI: Frecuencia de muestreo: 37.8 KHz, 8 bits por muestra. en cuatro canales mono o dos estéreo. Modo Mid-HIFI: Frecuencia de muestreo: 37.8 Hz, 4 bits por muestra. en ocho canales mono o cuatro estéreo. Modo Voz: Frecuencia de muestreo: 18.9 KHz, 4 bits por muestra. en dieciséis canales mono u ocho estéreo. Los tres últimos modos utilizan compresión por ADPCM.

Compact Disc Interactive - CDI

El CDI es una especificación que proporciona las características que debe cumplir un sistema basado en ordenador, con posibilidades de audio y vídeo interactivo y que utiliza el CD como medio de almacenamiento. El CDI no está diseñado para su uso como ordenador personal. Se puede contemplar como una especificación de un nuevo y diferente ordenador personal. El sistema ha sido desarrollado por Philips y Sony. Los principales elementos de un sistema CDI son 1) Una unidad de CD, 2) Una placa que contiene una CPU de la familia 68000 con memoria RAM y ROM 3) Circuitos integrados para aplicaciones de audio y vídeo y 4) Un interface para un monitor RGB o un televisor. El CDI se debería ver como el equivalente

de un VCR (reproductor de vídeo convencional) que se conecta al aparato de televisión del usuario.

El hardware asociado a la parte de vídeo, proporciona diversos modos de funcionamiento en secuencias de imagen y en gráficos. El CDI tiene la capacidad para visualizar vídeo en parte de la pantalla con baja resolución temporal. El sistema está limitado por la máxima velocidad de transferencia que puede proporcionar el CD y por la ausencia de algoritmos sofisticados de descompresión de imagen. El algoritmo de compresión de imágenes que se utiliza está basado en técnicas de run-length, que proporciona imágenes muy comprimidas con la apariencia de dibujos animados. Esta es una de las razones por la que se dice que el CDI está especialmente diseñado para su uso por niños. Para compensar las posibles desventajas con respecto a la parte de vídeo, Philips ha anunciado un circuito integrado basado en el MPEG para finales del año 1992 que se incorporará en el sistema CDI.

El sistema de almacenamiento de audio es el mismo que el utilizado en el DVI.

El presente artículo ha intentado dar una visión global de los actuales standards utilizados en compresión de audio y vídeo en aplicaciones multimedia con especial énfasis en los sistemas DVI y CDI. Con respecto al futuro, es de esperar que los avances en ordenadores y el desarrollo de nuevos circuitos integrados continúen desarrollando nuevas aplicaciones y mejorando las presentes en el entorno de compresión de audio y vídeo. La aceptación de los standards propuestos en las dos áreas contribuirán a acelerar la introducción en el mercado de productos más baratos y asequibles a un mayor número de aplicaciones. La perspectiva a muy corto plazo es la intersección de tecnologías, hasta hace poco muy disjuntas e independientes, como los ordenadores, la televisión y el audio de gran calidad. El futuro no distinguirá entre un ordenador que incorpore un sintonizador de televisión y un televisor que incorpore CPU muy sofisticadas.

Antoni Gasull

Ingeniero de Telecomunicación (1978) y Doctor Ingeniero de Telecomunicación (1983) por la Universidad Politécnica de Cataluña (UPC). En la actualidad es Profesor Titular de Universidad del Departamento de Teoría de Señal y Comunicaciones y profesor de la Escuela de Ingeniería de Telecomunicación de Barcelona de la UPC. Sus áreas de interés son el Procesado y Análisis de Imagen.

Enrique Masgrau

Ingeniero de Telecomunicación (1978) y Doctor Ingeniero de Telecomunicación (1983) por la Universidad Politécnica de Cataluña (UPC). Es Profesor Titular de Universidad del Departamento de Teoría de Señal y Comunicaciones y profesor de la Escuela de Ingeniería de Telecomunicación de Barcelona de la UPC. Sus áreas de interés son el Filtrado Adaptativo de Señal en Comunicaciones, Codificación de Voz y Audio y Procesado Robusto de la Voz.



Sus áreas de interés son el Filtrado Adaptativo de Señal en Comunicaciones, Codificación de Voz y Audio y Procesado Robusto de la Voz.