

Comunicaciones Eléctricas

Volumen 59

Número 3 · 1985

Comunicaciones Eléctricas

Edición española de ELECTRICAL COMMUNICATION
Revista técnica publicada trimestralmente por ITT Corporation

Comunicaciones Eléctricas presenta las investigaciones, los desarrollos y las realizaciones conseguidas por ITT y sus compañías asociadas.

Publicada desde 1922 en versión inglesa, se edita actualmente en cuatro idiomas y se distribuye en el mundo entero.

Se invita a los ingenieros de ITT a proponer proyectos de artículos, cuyos resúmenes deben enviarse al editor internacional para su consideración.

Director Ejecutivo

Lester A. Gimpelson, Bruselas

Editor, Comunicaciones Eléctricas

Antonio Soto, Madrid

Editor, Electrical Communication

Michael Deason, Harlow

Editor, Elektrisches Nachrichtenwesen

Otto Grewe, Stuttgart

Editor, Revue des Télécommunications

Jean-Pierre Dartois, París

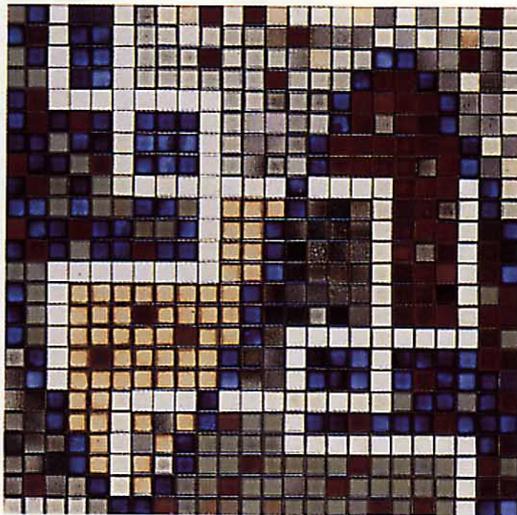
Publicado en 15 de julio de 1985

© ITT Corporation, 1985

Las direcciones de los editores se dan en la página 352

Proceso de señal

- 250 **Presentación**
- 252 **Proceso de señal: tecnología y perspectivas**
J. G. Dunn
- 260 **Codificación del habla a 800 y 400 bit s⁻¹**
J. Carmody y J. Rothweiler
- 266 **Proceso de la voz en centrales telefónicas públicas**
R. Christensen, Ø. Johnsen y B. Patovan
- 273 **Aplicación del reconocimiento y la síntesis de voz a los servicios de las PABX**
H. Mulla y J. F. Vaughan
- 281 **Marcador telefónico activado por la voz**
M. Immendörfer
- 286 **Codificación de señales de vídeo**
G. Bostelmann y P. Pirsch
- 295 **Filtros integrados para señales de vídeo**
A. Ali
- 300 **Nuevo enfoque de la conformación de haz adaptativa**
P. J. Hargrave y C. R. Ward
- 306 **Cálculo de la transformada discreta de Fourier en el procesador celular**
J. M. Cotton y G. E. Masterson
- 312 **Tendencias en procesadores digitales de señales programables de circuito integrado único**
K. N. Kneib
- 320 **Procesador digital de señales para aplicaciones de telecomunicación**
S. Das, S. Jayanthi, F. A. Middleton y S. G. Morton
- 327 **Proceso digital de señales en módems de transmisión de datos**
G. Bång y P. Carlsson
- 333 **Compensador de eco para módems de datos a dos hilos**
T. Alvestad y T. J.-C. Eriksen
- 338 **Compensación de eco adaptativa para transmisión de datos en banda base**
M. Gilsanz, F. Pedrón y J. A. Siles
- 345 **Aplicación del proceso digital de señal a las pruebas e interfaz de enlace del Sistema 12**
M. Fucito y G. Mocerino
- 351 **En este número**
-



Durante largos años el equipo electrónico se ha cimentado sobre una amplia gama de funciones de proceso analógico de señal, contando entre ellas la amplificación, el filtrado y el cálculo. Los dispositivos VLSI digitales pueden ahora realizar muchas de esas funciones con mayor precisión, y además un gran abanico de procesos más complejos, prácticamente inabordable con las técnicas analógicas. El reconocimiento y la síntesis de voz, la reducción de la redundancia de las señales y la conformación de haz adaptativa no son sino algunas de las nuevas aplicaciones del proceso digital de las señales.

Presentación

Entregar señales sin alteraciones apreciables ha sido siempre el fin primario de las comunicaciones, si bien la tecnología analógica no ha hecho fácil su logro. Hoy día, las técnicas digitales ponen a nuestro alcance artificiosas transformaciones de señales complejas, no condicionadas; por otra parte, estas mismas técnicas hacen progresar la función de distribución básica en la telecomunicación, de lo que resulta un provechoso maridaje entre proceso de señal y comunicaciones. El campo de aplicación del proceso de señal es, sin embargo, mucho más amplio en la actualidad, cubriendo el radar, sonar, control de sistemas, televisión digital y reproducción de sonido de alta fidelidad, por citar unos pocos ejemplos.

Tradicionalmente se han utilizado técnicas analógicas para procesar señales también analógicas, sometiéndolas a amplificación, filtrado, atenuación, multiplexación, almacenamiento, transmisión o cálculos aritméticos. En los años 80 se ha experimentado una rápida transición de la tecnología analógica a la digital en casi todos los sectores de la técnica, buscando la prestación de funciones más elaboradas, solicitadas hoy por los usuarios de sistemas, y cuyo logro sería imposible o prohibitivamente costoso por medios analógicos. Aunque la tecnología digital de proceso de señal venga utilizándose muchos años en aplicaciones donde resulta insustituible por su superior comportamiento, sólo recientemente ha resultado económica para uso general merced a los circuitos digitales VLSI. Entre sus ventajas cuenta la mayor exactitud y precisión de cálculo, fabricación de circuitos más perfecta y reducción del tamaño del equipo.

La telecomunicación ejemplifica muy bien la tendencia tecnológica en proceso de señal. La modulación por impulsos codificados (concebida en 1937 por un ingeniero de ITT) lleva muchos años en uso para mejor aprovechar la planta de transmisión, aunque la conmutación continuara siendo analógica. La tecnología de conmutación digital ha suplantado a la analógica en la última generación de centrales, exigiendo con más apremio el proceso digital de las señales y conduciendo a una red de telecomunicación enteramente digital, tanto en conmutación y transmisión como en el tratamiento de las señales.

La central digital Sistema 12 de ITT utiliza profusamente el proceso digital para ofrecer un abanico muy amplio de refinados servicios a los abonados y administraciones explotadoras, en particular cuando las centrales digitales trabajan en un entorno predominantemente analógico, como ahora ocurre en una central interurbana internacional. Los supresores de eco, interfaces de señalización CCITT n° 5, detectores de voz, módulos de conferencia y de prueba de enlaces, todos ellos incorporan el tratamiento digital con evidentes ventajas.

Por el carácter modular y distribuido de las centrales Sistema 12 fue aconsejable utilizar criterios similares para la prueba de líneas y de enlaces. Para lograrlo se ha desarrollado un analizador de señales de prueba, que es un procesador digital de señal con un potente repertorio de instrucciones que proporcionan todas las funciones de generación de tonos y filtrado necesarias para probar líneas de abonado y enlaces, así como para mantener la planta externa desde un centro remoto. La versatilidad de este dispositivo ha determinado su uso en otras aplicaciones, a modo de poderoso instrumento de laboratorio capaz de medir y comprobar conceptos del tratamiento digital de señales en los comienzos del desarrollo de un sistema.

El proceso de la voz es de un creciente interés en la telecomunicación: se incluyen en él la codificación del habla para reducir la velocidad de transmisión o aprovechar al máximo la memoria disponible, así como el reconocimiento y la síntesis de voz. En los últimos años ha quedado patente que no se utilizan habitualmente muchos de los nuevos servicios que las centralitas modernas ofrecen a sus abonados, a causa del gran número de códigos (20 ó más) que pueden necesitarse para acceder a tales servicios. La experiencia demuestra que los usuarios prefieren equipos que se adapten a su modo de trabajo, en lugar de tener ellos que acomodarse a las exigencias del equipo. El reconocimiento y la síntesis de voz jugarán un importante papel en el logro de hacer cómodamente accesibles los servicios de las PABX, así como en el más amplio contexto de la comunicación hombre-máquina natural y eficiente.

Las compañías ITT europeas y norteamericanas investigan las dos principales modalidades de reconocimiento de voz — dependiente y no dependiente del locutor —, así como sus diferentes aplicaciones. El reconocimiento de voz dependiente del locutor requiere que el usuario "enseñe" a la máquina a distinguir su voz, y es muy adecuado para una marcación oral individualizada, sólo pronunciando el nombre de la persona llamada, muy ventajosa en caso de limitaciones físicas. Por el contrario, el reconocimiento no dependiente del locutor es apropiado para las PABX, ofreciendo a todo usuario la facilidad de marcación oral y el acceso a servicios especiales sin tener que volver a adiestrar al sistema para cada nuevo servicio o usuario incorporado. La complejidad mayor de esta solución queda sobradamente compensada cuando el número de usuarios es elevado. Sea cual fuere el método escogido, los usuarios saldrán beneficiados al poder comunicarse con el sistema mediante órdenes verbales y obtener respuestas habladas. No obstante, los factores humanos deberán cuidarse especialmente para que la aceptación de estos sistemas sea máxima.

El tratamiento de imagen es una aplicación relativamente nueva del proceso de señal, cuya importancia crece con rapidez en el marco de los actuales desarrollos de redes digitales de servicios integrados de banda ancha, capaces de transportar y difundir señales de vídeo. La prueba de campo BIGFON de SEL en Alemania ha demostrado la superior calidad y varios de los nuevos servicios que ofrecerán tales redes. Dado que la salida de un sistema de imagen es generalmente analógica, esas señales han de codificarse digitalmente, eliminando toda información redundante para minimizar la velocidad binaria. ITT ha avanzado ya mucho en el desarrollo de técnicas de codificación de fuentes que permitirán transmitir señales de vídeo, de calidad análoga a las de televisión, a 139 Mbit s^{-1} e incluso a 69 Mbit s^{-1} , así como en la obtención de filtros adecuados para las señales de vídeo.

El tratamiento digital de señales requiere un volumen considerable de cálculo con algoritmos complicados, y las aplicaciones futuras exigirán multiplicar por grandes factores las velocidades de cálculo actuales. No parece que las clásicas arquitecturas Von Neumann de proceso en serie puedan rendir los resultados requeridos; para superar este problema dotando a las compañías ITT de la tecnología necesaria, el Advanced Technology Center desarrolla un procesador de estructura celular, en el cual una sola instrucción actúa simultáneamente sobre muchos datos y ofrece así una espectacular mejora en la capacidad de proceso.

A pesar de ser aplicable a todo el campo de la electrónica, no suele concederse a la tecnología de proceso de señal la difusión pública que merece. Esperamos que este número de *Comunicaciones Eléctricas* haga vislumbrar a los lectores los acelerados cambios que están ocurriendo al reemplazar la tecnología digital a la analógica en el proceso digital de las señales, descubriendo un vasto campo de aplicaciones en expansión.

Proceso de señal: tecnología y perspectivas

El proceso de señal es una tarea esencial en muchas áreas tecnológicas, entre las que se cuentan la telecomunicación, visión por ordenador, síntesis y reconocimiento de voz, y radar. En la actualidad las técnicas digitales están sustituyendo a las analógicas, aportando ventajas en prestaciones y permitiendo operaciones hasta ahora imposibles.

J. G. Dunn

ITT Headquarters, Nueva York,
Estados Unidos de América

Introducción

Una señal materializa físicamente la información: no es la propia información, sino sólo su vehículo. Señales pueden ser el movimiento de una mano, una luz que parpadea, los puntos y rayas del código Morse, la secuencia de impulsos del MIC, y muchas otras. Procesar una señal consiste en hacer operaciones con ella para obtener cierto resultado que se desee. Ejemplo es la corrección de degradaciones en la transmisión: la amplificación combate la atenuación, la igualación corrige la distorsión, y la codificación (es decir, expansión del ancho de banda) inmuniza contra el ruido. También se pueden procesar las señales para crear sistemas más sofisticados, como sucede en la modulación y en la multiplexación por división en frecuencia, que ejecuta dos clases de proceso de señal, filtrado y translocación de frecuencia.

Durante muchos años se han utilizado técnicas analógicas para estos tipos de proceso de señal. Hoy día, la tecnología de ordenadores, unida a la tecnología LSI y VLSI, permite importantes avances. Las operaciones antes realizadas por métodos analógicos se pueden efectuar más económicamente usando técnicas digitales. Además, la tecnología de los microprocesadores hace factible la ejecución de operaciones cada vez más complejas.

ITT ha contribuido destacadamente al proceso de señales en las áreas de MIC, invención de un ingeniero de ITT, y de conmutación digital, donde ITT es pionera. Sirva como ejemplo que ITT ha desarrollado nuevas técnicas de proceso digital de señales para mejorar las prestaciones de los circuitos interfaces de línea y de enlace en la central digital Sistema 12.

Proceso digital de señales

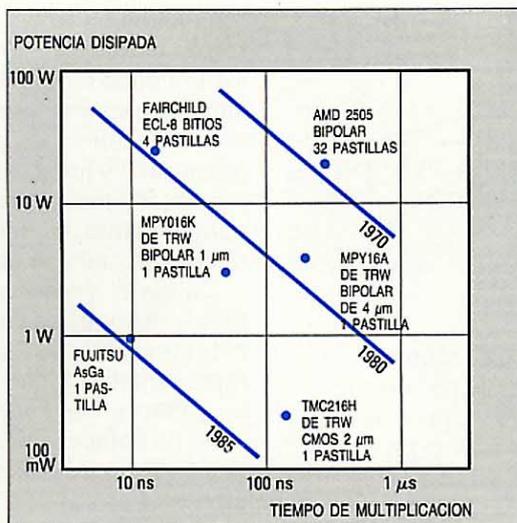
Inicialmente el proceso de señales se hizo con técnicas analógicas, pero la tecnología de ordenadores ha producido una gran mejora en las prestaciones, sobre todo porque se pueden producir dispositivos (p. ej., filtros) con definición muy precisa, en cantidades ilimitadas. Además, las técnicas lógicas y de ordenadores permiten un tratamiento mucho más complejo del que sería posible con los circuitos analógicos.

En este aspecto, el ancho de banda de la señal ha sido el principal factor para determinar la aplicabilidad de las técnicas digitales avanzadas. Tanto es así, que sistemas con ancho de banda vocal, como los de voz y sonar, han utilizado masivamente operaciones de proceso de señal complejas, mientras que los de radar y vídeo, con un ancho de banda mucho mayor, se han limitado a funciones de proceso más elementales. Además, el advenimiento de las tecnologías LSI y VLSI ha enriquecido mucho la capacidad del proceso de señal. Operaciones que fueron impracticables por la gran cantidad de proceso requerido, son ahora factibles; otras que se juzgaban imposibles, están hoy en desarrollo.

El proceso digital de señales se basa en la conversión de las señales de entrada a forma digital. La señal se muestrea en el tiempo, y cada muestra se representa por una "palabra" binaria de ordenador, lo cual da un conjunto o vector de números. La función de proceso de señal consiste en ejecutar secuencias de cálculos numéricos sobre estos conjuntos, y posiblemente transformar el resultado a la salida otra vez a forma analógica.

Los dos principales métodos para realizar los cálculos numéricos intermedios son la

Figura 1
Efecto de la evolución de los circuitos integrados en multiplicadores de 16×16 bits con un producto de 32 bits.



lógica cableada y los procesadores programables. Ambos se utilizan con profusión, y seguramente se mantendrán en un futuro previsible. La lógica cableada es preferible cuando los requisitos y funciones están bien definidos y no sujetos a cambio, y también parece ser la solución más económica en aplicaciones de gran volumen. Por otro lado, los procesadores programables son más flexibles, y pueden resultar más económicos cuando se disponga de sistemas de proceso que cumplan los requisitos de una aplicación concreta.

Aunque dependa del ancho de banda de la señal original, y de la necesidad de proceso en tiempo real, el proceso de señales suele requerir un gran número de operaciones por unidad de tiempo. A menudo se necesitan procesadores programables especializados para obtener canales de entrada/salida eficientes, rápido almacenamiento de datos provisionales, circuitos multiplicadores, eficaz indexado de conjuntos, etc., y tales procesadores se denominan procesadores de señal; su realización en una sola pastilla equiparable a los microprocesadores se llama circuito procesador digital de señales.

Las posibilidades de proceso digital de señales que encierra la tecnología frecuentemente se miden con referencia a la operación de multiplicar, la cual ha sido históricamente el factor limitador de la velocidad de proceso. La figura 1 muestra la evolución de los circuitos multiplicadores en cuanto a velocidad y consumo. Generalmente, ha habido circuitos de memoria y aritméticos compatibles con los multiplicadores.

Recientemente se ha empezado a utilizar para el proceso digital la tecnología de arseniuro de galio, y cabe esperar que su influjo sea apreciable en un futuro próximo, sobre todo en aplicaciones de banda ancha tales como los sistemas de radar y vídeo.

Dicha tecnología permite combinar circuitos analógicos y digitales en la misma pastilla, pudiendo así integrarse circuitos analógicos de microondas con la lógica digital.

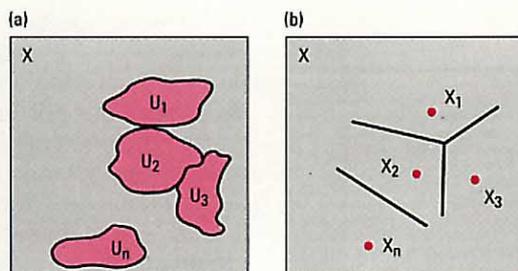
Las funciones de proceso de señal se pueden realizar con una ejecución en serie de los pasos de la función, de acuerdo con la arquitectura en serie de la mayoría de procesadores de señal. Sin embargo, cuando el ancho de banda de la señal es mayor que la capacidad en tiempo real de un solo procesador, se hace imprescindible utilizar procesadores múltiples. Esto obliga a preguntarse cómo dividir la función entre varios procesadores de un modo eficaz, cómo sincronizar procesos concurrentes, y cómo se deben transferir los datos comunes entre procesadores. Para resolver tales problemas se han introducido nuevas técnicas, como las arquitecturas de flujo de datos.

Dado que la capacidad de los circuitos microelectrónicos aumenta, se intenta ahora realizar las funciones de proceso de señal en un circuito único mediante la repetición de elementos de proceso simples, transfiriéndose los datos sólo entre elementos adyacentes; esta técnica minimiza el coste del diseño para un circuito complejo. En las operaciones matriciales puede lograrse tal resultado, ya que intervienen en muchas funciones de proceso intensivo o las estructuran. Se obtiene entonces un conjunto celular de elementos de proceso conectados localmente. Una herramienta útil para analizar la forma de trabajo es un frente de onda simulado en ordenador, donde la acción procesadora se propaga a través del conjunto como una onda luminosa¹.

Reconocimiento de señal

El reconocimiento de señal es la asignación automática de una señal a una categoría predeterminada, con el fin de provocar una acción también predeterminada. Por ejemplo, al reconocer una orden pronunciada, el sistema ejecuta dicha orden. El reconocimiento de señal y la operación inversa, la generación de señal, son las claves del interfaz del hombre con las máquinas. Un ejemplo de generación de señal es la conversión de texto escrito a voz. El texto se almacena en una base de datos de ordenador, como secuencia de caracteres en formato ASCII (código normalizado americano para intercambio de información). Las letras y combinaciones de letras se representan por fonemas; un conjunto de reglas gobierna la asignación a categorías, que también depende del contexto en que aparece un carácter o varios combinados.

Figura 2
Reconocimiento de señal:
 a) partición del espacio de señal para reconocimiento
 b) regiones de decisión por el "vecino más próximo".



Cada fonema se representa por un espectro de sonidos único, y por tanto por un conjunto único de parámetros, asignados conforme a otras reglas, que activan un sintetizador de voz, generando así una señal hablada.

Cada vez adquiere mayor importancia el reconocimiento de señal, al desarrollarse sistemas automáticos para reconocimiento de voz, reconocimiento de imágenes, y clasificación e identificación de señales para radar, sonar, espionaje y guerra electrónica. Puesto que estos sistemas suelen intentar hacer las mismas cosas que el hombre — reconocer el lenguaje hablado, la voz y la imagen —, se pueden considerar como ramas de la inteligencia artificial.

El reconocimiento de señal se puede interpretar como una partición del espacio de la señal de entrada en regiones que no se solapan, correspondientes a cada una de las categorías que han de reconocerse. La figura 2a es una ilustración bidimensional del espacio de entrada, el cual normalmente tendrá un número de dimensiones elevado. En ella, X representa el espacio de todas las señales posibles de entrada al reconocedor, dividido en regiones U_i . Si la señal de entrada cae en la región U_i , la salida del reconocedor declara que la señal pertenece a la categoría i . En los casos prácticos, el número de dimensiones de este espacio de señal es desmesuradamente grande, pero si la señal se representa por parámetros (características), el espacio que les corresponde puede tener bastantes menos dimensiones.

Un método probabilístico es definir cada región basándose en estadísticas conocidas o medidas de las señales y sus categorías. Por ejemplo, con máxima verosimilitud se podría definir la región U_i como el conjunto de señales cuya función de verosimilitud para la categoría i (la probabilidad condicional de la señal observada, supuesto que es de categoría i) es mayor que la función de verosimilitud para cualquier otra categoría. Sin embargo, en la práctica, el número de señales en X es demasiado grande para poder obtener estadísticas suficientes que definan con precisión tales regiones, e incluso para construir el sistema si las regiones se definieran por enumeración directa

(búsqueda en tablas) de cada señal y su región correspondiente. En consecuencia, los métodos probabilísticos tratan de aproximarse a las distribuciones de probabilidad mediante funciones sencillas (p. ej., la gaussiana) y luego estimar los parámetros para esas funciones por estadística de agrupamiento, es decir, por un procedimiento de "adiestramiento".

Quizás el método más sencillo para definir regiones sea por emparejamiento de patrones, en el que se elige una señal representativa x_i para cada categoría, y luego se define la región U_i como el conjunto de señales comprendidas dentro de una determinada distancia a x_i . Alternativamente, U_i puede venir definida por todas las señales que estén más próximas a x_i que a cualquier otra x_j (Fig. 2b). En la práctica, la dificultad se ha transferido a la definición de una medida de distancia que tenga sentido.

Recientemente se han desarrollado programas de ordenador que definen de forma automática regiones en el espacio de características, con base en un conjunto de ejemplos conocidos (es decir, datos de adiestramiento)^{2,3}. Su resultado es un árbol de decisión binaria, cada nodo del cual corresponde a una subdivisión, realizada por un hiperplano, de una región del espacio de características. El programa selecciona sucesivamente tales hiperplanos para dividir mejor los ejemplos de adiestramiento de esa región en las subregiones apropiadas. La selección se puede basar en una medida de la entropía de información, esencialmente tratando de dividir los ejemplos en subconjuntos que separen las categorías tanto como sea posible. Los programas operan bien hasta con algunas docenas de características (parámetros).

Los sistemas de reconocimiento de voz existentes son capaces de reconocer palabras individuales y frases en un vocabulario de varios centenares, así como de ejecutar las acciones apropiadas. Las aplicaciones incluyen la marcación telefónica oral^{4,5}, y también situaciones en que el usuario tiene las manos y la vista ocupadas, como sucede en inventarios, inspecciones y manejo de paquetes. Al ir creciendo el vocabulario, se harán viables aplicaciones más artificiosas, como las máquinas de escribir activadas por voz y los traductores de idiomas.

Técnicas de optimización

Las técnicas de optimización se están ahora introduciendo en diversas aplicaciones de la telecomunicación, como la igualación adaptativa, la compensación de ecos y el rechazo de interferencias, en las que el proceso de una señal se adapta a las carac-

Tabla 1 — Aplicaciones de la aproximación de mínimos cuadrados

Aplicación	Vector de señal a ser aproximado	Vectores de señal base	Ejemplos
Igualación adaptativa	Señal de adiestramiento o señal de decisión realimentada	Versiones retardadas de la señal recibida	Igualación del canal para transmisión de datos ⁶
Compensación de eco	Señal de eco recibida	Versiones retardadas de la señal transmitida	Transmisión de datos en duplex total por un solo canal telefónico a 2 hilos ^{7,8}
Predicción lineal	Señal de entrada	Versiones retardadas de la señal de entrada	Análisis espectral de la voz ^{9,10} , excisión de interferencias ¹¹
Compensación lineal de interferencias	Señal interferente	Versiones retardadas de una señal correlada a la señal interferente	Compensación del ruido acústico con dos micrófonos ¹²
Conformación de haz	Señal procedente de un elemento de referencia del conjunto	Señales procedentes de los restantes elementos del conjunto	Conformación de haz adaptativa para los conjuntos de antena y sensores ¹³

terísticas de la misma (Tabla 1). Así, el sistema determina los parámetros del proceso de tal forma que se mejore la calidad de la señal, para lo cual se compara la señal que se está procesando con una señal "ideal" de referencia almacenada en una base de datos. Por ejemplo, para una aplicación de igualación adaptativa la referencia sería una forma de onda de un impulso ideal; para la compensación de eco la referencia sería el eco que se va a compensar. Las técnicas de optimización utilizadas en

estas aplicaciones comprenden la aproximación de mínimos cuadrados, la estimación estadística de mínimos cuadrados y la estimación de máxima verosimilitud.

Aproximación de mínimos cuadrados

El propósito de esta aproximación (ver el recuadro) es minimizar el promedio del cuadrado del error (diferencia entre la señal procesada y la referencia). Se sigue normalmente este método para procesar señales no aleatorias.

Como ilustración, la figura 3 muestra un igualador lineal en banda base realizado como filtro transversal (respuesta impulsiva finita), al formar una combinación lineal de las sucesivas muestras de una señal en el tiempo. Los vectores base son diferentes versiones retardadas de las señales recibidas que han de igualarse. La combinación lineal de estos vectores es la señal igualada; los coeficientes de ponderación lineal se determinan de modo que la señal igualada se aproxime a una señal de adiestramiento*, en el sentido de los mínimos cuadrados. En aplicaciones como los igualadores de módem y conformadores de haz para antena, cuya señal recibida puede ser una portadora, se emplean igualadores lineales separados para las componentes en fase y cuadratura. La figura 4 muestra el comportamiento de un igualador paso banda.

* La señal de adiestramiento es un duplicado de la referencia, transmitida al comienzo de la llamada con el fin de poder igualar el canal para las señales siguientes.

Teorema de proyección

El teorema de proyección da una base matemática para la aproximación de mínimos cuadrados. Las señales se representan por vectores en un espacio de Hilbert, que es un espacio vectorial lineal en el que se define un producto interno. El producto interno, también llamado producto escalar, entre dos vectores suele definirse como suma (o integral) de los productos de sus componentes. La longitud de un vector se define luego como la raíz cuadrada del producto interno del vector por sí mismo — más comúnmente, la raíz cuadrada de la suma de los cuadrados de sus componentes —.

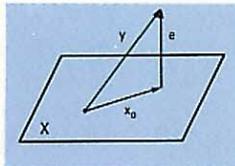
La situación usual es que una señal dada (vector) y se aproxima por una combinación lineal

$$\mathbf{x}_0 = a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_n \mathbf{x}_n$$

de un conjunto dado de vectores $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. Entonces, el vector de aproximación \mathbf{x}_0 debe estar en el subespacio lineal \mathbf{X} que abarcan los vectores en la prueba. El error de aproximación se define como el vector $\mathbf{e} = \mathbf{y} - \mathbf{x}_0$. La aproximación de mínimos cuadrados trata de minimizar la longitud de este vector error.

Teorema¹⁴

Sea \mathbf{H} un espacio de Hilbert y \mathbf{X} un subespacio lineal de \mathbf{H} . A cada vector \mathbf{y} en \mathbf{H} corresponde un vector único \mathbf{x}_0 en \mathbf{X} , tal que haga mínima la longitud de $\mathbf{e} = \mathbf{y} - \mathbf{x}_0$. Además, \mathbf{x}_0 es el único vector de \mathbf{X} que minimiza



\mathbf{e} , siempre en el caso de que \mathbf{e} sea ortogonal a \mathbf{X} (es decir, ortogonal a cualquier vector en el subespacio \mathbf{X}).

El teorema de proyección establece que el vector \mathbf{x}_0 de aproximación de mínimos cuadrados es la proyección ortogonal de \mathbf{y} en el subespacio abarcado por los vectores base $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. Esto se traduce a la condición de que el vector de error \mathbf{e} es ortogonal a todos los vectores base. En notación de producto interno:

$$(\mathbf{e}, \mathbf{x}_i) = 0 \quad \text{para } i = 1, 2, \dots, n.$$

Esto expresa la condición de que el producto interno del vector error por cada uno de los vectores base es cero, lo que equivale a un conjunto de ecuaciones lineales simultáneas con los coeficientes a_1, a_2, \dots, a_n (implícitos en la expresión de \mathbf{e}) y que se denominan ecuaciones normales. Suponiendo que el subespacio tiene un número finito de dimensiones (n finito) y que los vectores \mathbf{x}_i son independientes, estas ecuaciones se pueden resolver directamente hallando los coeficientes de mínimos cuadrados. El cuadrado de la longitud del vector de error mínimo resultante vendrá dado por:

$$(\mathbf{e}, \mathbf{e}) = (\mathbf{y}, \mathbf{y}) - a_1 (\mathbf{x}_1, \mathbf{y}) - a_2 (\mathbf{x}_2, \mathbf{y}) - \dots - a_n (\mathbf{x}_n, \mathbf{y}).$$

El caso de no ser independientes los vectores base, con lo que la matriz de coeficientes de las ecuaciones normales es singular, se examina en el texto.

Como alternativa al cálculo y resolución de las ecuaciones normales, se utilizan varias técnicas iterativas, basadas en que los productos internos $(\mathbf{e}, \mathbf{x}_i)$ son las componentes del vector gradiente del cuadrado de la longitud del vector error, considerado como función escalar de los coeficientes a_i . El vector gradiente indica la dirección en que se deben ajustar los coeficientes para minimizar tal longitud.

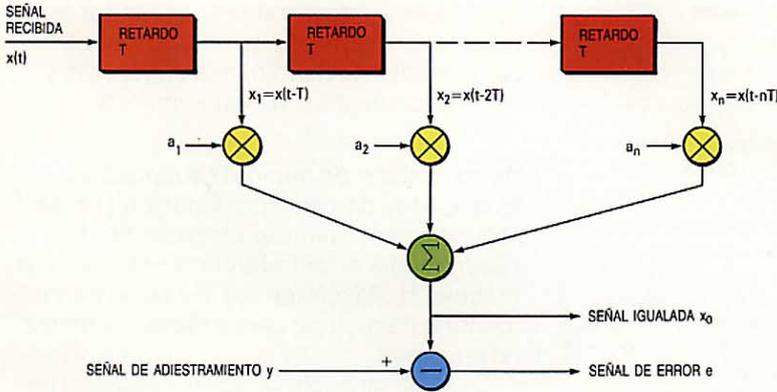


Figura 3 Igualador lineal.

El caso en que la matriz de coeficientes en la aproximación de mínimos cuadrados es singular (esto es, no posee una inversa), o casi singular, ha despertado últimamente un gran interés^{15, 16, 17}. En este caso, las ecuaciones lineales están mal condicionadas numéricamente, debido a que el número de parámetros es superior al mínimo necesario para aproximarse estrechamente al vector de señal. Esto hace que la solución sea demasiado sensible al ruido y a otros errores. Se han notificado prestaciones muy mejoradas en aplicaciones del tipo de estimación espectral y conformación de haz, utilizando bien un valor singular o bien una descomposición en autovalores. El efecto del ruido se reduce haciendo cero los autovalores pequeños en la descomposición de la matriz¹⁵.

Para la realización de sistemas que utilizan la aproximación de mínimos cuadrados existen dos métodos: el de soluciones directas y el de soluciones iterativas a las ecuaciones normales. Los primeros sistemas estaban basados en la solución iterativa, por ser relativamente sencilla la lógica para la realización física. Este método es satisfactorio cuando la variación de las señales en el tiempo es lenta, como sucede en la igualación adaptativa de canales tele-

fónicos, donde la respuesta del canal suele variar lentamente. Una aplicación en que se necesita la solución directa de las ecuaciones normales es la predicción lineal de la voz, debido a que el tracto vocal que se está modelando cambia muy deprisa. La tecnología de los microprocesadores ha hecho viables estos sistemas, ya que se pueden ejecutar en tiempo real las complejas secuencias de cálculos numéricos.

Se han desarrollado, recientemente, algoritmos para solución directa (exacta) de las ecuaciones normales que resultan de las aplicaciones a series temporales, como la igualación y predicción lineal¹⁸. En este caso, la estructura de la matriz (llamada de covarianza) de las ecuaciones obtenidas de los datos, se utiliza para calcular la solución de n -ésimo orden a partir de la de $(n-1)$ -ésimo orden (recursión en el orden) con un total de operaciones del orden de (n^2) , en vez del orden de (n^3) que necesitan los métodos convencionales. Hay también algoritmos eficientes para calcular la solución a partir de ventanas previas en el tiempo (recursión en el tiempo)¹⁹.

Los ejemplos de la tabla 1 son aplicaciones de la aproximación de mínimos cuadrados *lineal*, es decir, la señal a obtener es una combinación lineal de los componentes de un conjunto dado de señales. La aproximación de mínimos cuadrados es también aplicable a combinaciones no lineales, aunque éstas son de uso más limitado.

Estimación estadística de mínimos cuadrados

Esta técnica de aproximación de mínimos cuadrados se utiliza cuando las señales que se procesan son aleatorias y hay suficiente información en la que basar los estudios estadísticos. La técnica depende del hecho de que las señales implicadas (incluyendo las posibles señales de interferencia, como el ruido) son a menudo probabilísticas, es decir, están caracterizadas por distribuciones de probabilidad. Una función típica de proceso de señal es estimar los parámetros de estas distribuciones, tales como la media o varianza. Un ejemplo del campo de la telecomunicación es la estimación de la media, que es la parte de señal dentro de la señal más ruido recibida. En radar, pueden citarse la estimación de amplitud, frecuencia o retardo de la señal de retorno para determinar el tamaño, velocidad o tipo de un objeto.

En el caso probabilístico de la aproximación de mínimos cuadrados, se supone que la señal y está relacionada con las señales del conjunto x_1, x_2, \dots, x_n . La mejor aproximación general (no lineal), en el sentido mínimo cuadrático de y como función de las x_i , es la media condicional de y dadas las x_i .

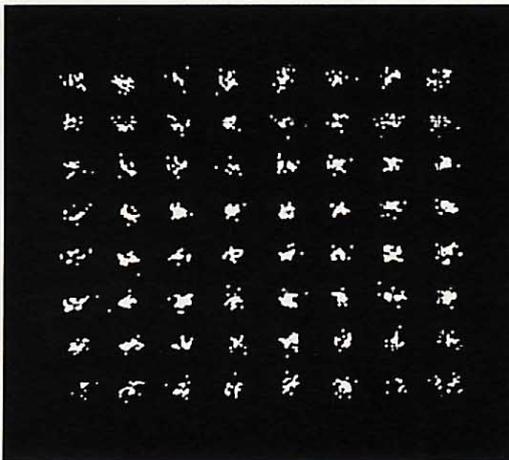


Figura 4 Constelación de 64 puntos modulados en cuadratura de amplitud para un módem de 16 kbit s^{-1} a la salida de un canal telefónico conmutado típico después de una igualación lineal adaptativa con un igualador paso banda de 100 etapas aproximadamente.

Cuando la distribución conjunta de las señales es gaussiana, la media condicional de y (dadas las x_i) se convierte en función lineal de las x_i (hablando estrictamente, la función es lineal sólo si las medias de las señales son cero). Por supuesto, si *a priori* se obliga a que la función de aproximación sea lineal, se aplicará la aproximación de mínimos cuadrados descrita en la sección anterior; esto es, la mejor aproximación es la proyección de la y en el subespacio lineal abarcado por las x_i . Los productos internos para señales probabilísticas se definen como sus covarianzas. La función de aproximación a menudo se obliga a ser lineal, por no conocer la distribución requerida para evaluar la media condicional.

Estimación de máxima verosimilitud

El método de máxima verosimilitud estima un parámetro por el valor que maximiza la función de verosimilitud (es decir, la probabilidad condicional de una señal observada, dado el parámetro). Este, pues, es el valor del parámetro que hace más verosímil la señal observada. Si la probabilidad condicional es gaussiana, la función logarítmica del valor de la función de verosimilitud es (excepto en una constante) el negativo de una función cuadrática. La estimación de máxima verosimilitud, en este caso, corresponde al mínimo de esa forma cuadrática. Esto tiene estrecha relación con la aproximación de mínimos cuadrados, a la cual equivale en algunos casos.

En la transmisión de datos, el receptor debe hallar un compromiso entre igualar la distorsión del canal e incrementar el ruido. El proceso óptimo resultaría de utilizar un filtro adaptado para maximizar la relación señal/ruido y restar (compensar) la interferencia entre símbolos producida por la combinación del canal y el filtro adaptado a partir de la salida de éste. Sin embargo, esto exigirá conocer los símbolos de datos adyacentes al que se está detectando para generar los términos de la interferencia entre símbolos que han de restarse.

Un método basado en la máxima verosimilitud consiste en calcular la función de verosimilitud (salida del filtro adaptado seguida de la sustracción) para cada posible combinación de símbolos de datos, y escoger la más grande. Esto se puede conseguir de modo eficaz obteniendo secuencialmente las decisiones de los símbolos mediante el algoritmo de Viterbi²⁰, lo cual, sin embargo no es práctico cuando la interferencia entre símbolos se extiende a más de unos pocos, ya que el proceso requerido crece exponencialmente en función de este número de símbolos.

Resulta más práctica la igualación asistida por datos, incluyendo la igualación con

decisión realimentada. Este método utiliza estimaciones de los símbolos de los datos (p. ej., por decisiones previas) en vez de los desconocidos datos adyacentes que se necesitan para generar los términos de la interferencia entre símbolos. El resultado es bueno si el número de errores en la estimación de símbolos es pequeños.

Un artículo reciente²¹ ha ayudado a aclarar el fundamento teórico de tal procedimiento, demostrando que el proceso óptimo (en el supuesto de cero errores en la estimación de los símbolos) es un filtro adaptado, seguido de un filtro transversal y de un compensador (Fig. 5). El compensador genera los términos de la interferencia entre símbolos a partir de los símbolos estimados, y los resta antes de proceder a la decisión sobre el símbolo presente. La teoría prevé un conjunto arbitrario de términos del compensador. Si estos términos incluyen sólo los símbolos previos, de ahí resulta un igualador con decisión realimentada convencional. Se demuestra también en el artículo que el filtro transversal tiene términos no nulos sólo donde no tiene términos el compensador. Los coeficientes del filtro transversal se calculan por la aproximación lineal de mínimos cuadrados examinada anteriormente.

Tendencias futuras

Las principales tendencias apreciables son consecuencia de aumentar la complejidad de los interfaces externos y del constante progreso en la tecnología de los circuitos integrados. La mayor funcionalidad aportada por los sistemas requiere interfaces más complejas con las personas que los utilizan o controlan.

En algunos casos, como los robots, el sistema realiza funciones que previamente ejecutaban las personas, como el reconocimiento de imágenes. Los conmutadores y pulsadores son reemplazados por teclados, "ratones" y sistemas de reconocimiento de voz; las luces y tonos, por visualizadores alfanuméricos y gráficos y por síntesis de voz. La necesidad de que el sistema comprenda las órdenes del usuario y de que éste interprete el comportamiento del sistema, se han convertido en los principales condicionantes del diseño.

Por estas razones, en el campo del proceso de señal, el reconocimiento de señales asumirá mayor importancia en el futuro. Aunque a veces se considera el reconocimiento de voz y visual comprendido dentro de la inteligencia artificial, no se han solapado mucho las técnicas y herramientas empleadas en uno y otro campo. Sin embargo, la técnica de los sistemas basados en el

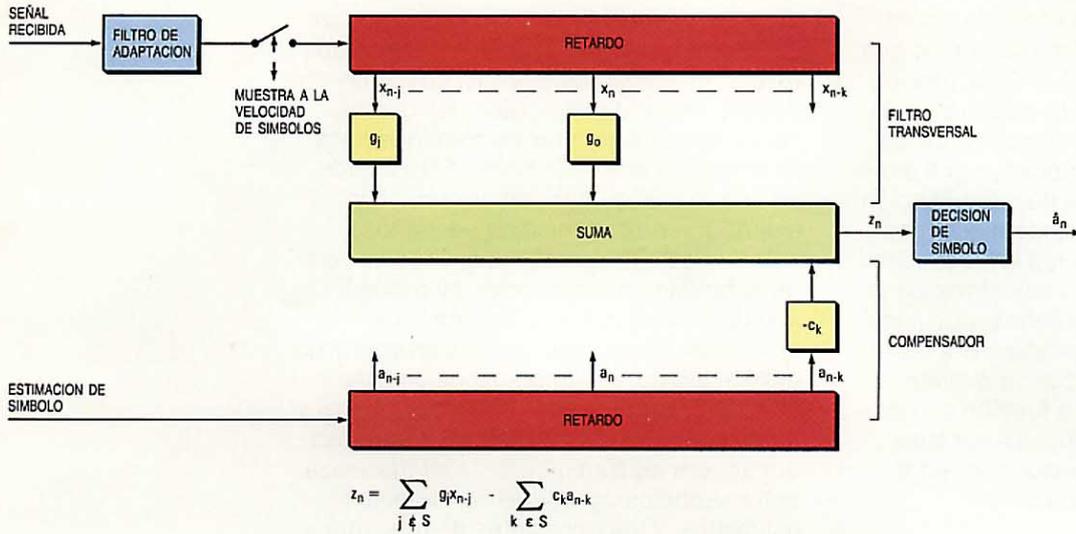


Figura 5 Igualación asistida por datos. S es el conjunto de enteros no nulos correspondiente a las posiciones en que el compensador tiene etapas, y en general puede contener enteros negativos.

conocimiento (sistemas expertos) ha avanzado mucho últimamente en el campo de la inteligencia artificial. En esencia, se trata de una técnica de programación estructurada para ejecutar una inferencia lógica en datos simbólicos, que ha dado lugar a máquinas con capacidad de razonar a un nivel cercano al comportamiento humano, dentro del dominio, claramente limitado, de los datos simbólicos definidos. Hasta ahora, los datos no simbólicos (señales de voz y de imagen, por ejemplo), se han excluido expresamente de los dominios de datos que se consideran apropiados para los sistemas expertos. Sin embargo, ampliar tales dominios para incluir las señales, puede ser un método fructífero para obtener el grado de inteligencia que desde hace mucho tiempo se cree necesario para el reconocimiento automático de señales. Sin duda se seguirá este camino en el futuro.

Conservarán su importancia los dos principales planteamientos del proceso digital de señales: la lógica especializada y los procesadores programables. Los procesadores digitales de señales en un solo circuito integrado, con mayor velocidad y más memoria, serán obviamente aplicaciones clave de la futura tecnología. Sin embargo, tendrán igual importancia los circuitos VLSI diseñados a medida para determinadas funciones de proceso de señal. Estos circuitos especializados serán realizables gracias a las perfeccionadas herramientas de diseño (compiladores de silicio) concebidas para las funciones de proceso de señal, a las técnicas de diseño de conjuntos de funciones de proceso concurrentes, y a la facilidad de acceso a los medios de fabricación (fundiciones de silicio)²².

Se puede esperar que continúe el interés por los conjuntos de circuitos VLSI para realizar operaciones matriciales. El uso de funciones de proceso de señal de compleji-

dad creciente, como las anteriormente descritas, que requieren el análisis de autovalores y autovectores, aumentará la necesidad de una eficaz realización física de estas funciones.

Se acelerará la tendencia de aplicar las refinadas funciones desarrolladas para aplicaciones de banda estrecha a las de banda ancha, a medida que se extienda el uso de la tecnología de arseniuro de galio. Esta tecnología se utilizará para combinar circuitos digitales y de microondas con el fin de realizar muchos tipos de funciones de proceso de señal para receptores y antenas adaptativas.

Referencias

- 1 U. Weiser y A. Davis: A Wavefront Notation Tool for VLSI Array Design: en H. T. Kung, B. Sproull y G. Steele (editores), *VLSI Systems and Computations*, Computer Science Press, 1981, Rockwill, MD.
- 2 L. Breiman, J. H. Friedman, R. A. Olshen y C. J. Stone: *Classification and Regression Trees*, 1984, Wadsworth, Belmont, California.
- 3 J. R. Quinlan: Semi-autonomous Acquisition of Pattern-based Knowledge: *Australian Computer Bulletin*, abril 1980. También en D. Michie (editor) *Introductory Reading in Expert Systems*, 1982, págs. 192-207, Gordon and Breach, Londres.
- 4 M. Immendörfer: Marcador telefónico activado por la voz: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 281-285 (en este número).
- 5 H. Mulla y J. F. Vaughan: Aplicación del reconocimiento y la síntesis de voz a los servicios de las PABX: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 273-280 (en este número).
- 6 G. Bång y P. Carlsson: Proceso digital de señales en módems de transmisión de datos: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 327-332 (en este número).
- 7 T. Alvestad y T. J.-C. Eriksen: Compensador de eco para módems de datos a dos hilos: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 333-337 (en este número).
- 8 M. Gilsanz, F. Pedrón y J. A. Siles: Compensación de eco adaptativa para transmisión de datos en banda base: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 338-344 (en este número).

- 9 J. Carmody y J. Rothweiler: Codificación del habla a 800 y 400 bits⁻¹: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 260–265 (en este número).
- 10 R. Christensen, Ø. Johnsen y B. Patovan: Proceso de la voz en centrales telefónicas públicas: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs. 266–272 (en este número).
- 11 N. H. Lu: Performance Characteristics of Adaptive Complex Linear Prediction Filters: *Proceedings of the Institute of Electrical and Electronics Engineers Military Communications Conference*, octubre 1983, págs. 700–707.
- 12 S. F. Boll y D. C. Pulsipher: Suppression of Acoustic Noise in Speech Using Two Microphone Adaptive Noise Cancellation: *Institute of Electrical and Electronics Engineers Transactions*, diciembre 1980, volumen ASSP-28, págs. 752-3.
- 13 P. J. Hargrave y C. R. Ward: Nuevo enfoque de la conformación de haz adaptativa: *Comunicaciones Eléctricas*, 1985, volumen 59, n° 3, págs 300–305 (en este número).
- 14 D. G. Luenberger: *Optimization by Vector Space Methods*, 1969, John Wiley, Nueva York.
- 15 D. W. Tufts y R. Kumaresan: Singular Value Decomposition and Improved Frequency Estimation Using Linear Prediction: *Institute of Electrical and Electronics Engineers Transactions*, agosto 1982, volumen ASSP-30, págs. 671–675.
- 16 J. A. Cadzow: Spectral Estimation: an Overdetermined Rational Model Equation Approach: *Institute of Electrical and Electronics Engineers Proceedings*, septiembre 1982, volumen 70, págs. 907–939.
- 17 D. H. Johnson y S. R. DeGraaf: Improving the Resolution of Bearing in Passive Sonar Arrays by Eigenvalue Analysis: *Institute of Electrical and Electronics Engineers Transactions*: agosto 1982, volumen ASSP-30, págs. 638–647.
- 18 M. Morf, B. Dickinson, T. Kailath y A. Vieira: Efficient Solution of Covariance Equations for Linear Prediction: *Institute of Electrical and Electronics Engineers Transactions*, octubre 1977, volumen ASSP-25, págs. 429–433.
- 19 M. L. Honig: Recursive Fixed-Order Covariance Least-Squares Algorithms: *Bell System Technical Journal*, diciembre 1983, volumen 62, págs. 2961–2992.
- 20 G. D. Forney: Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference: *Institute of Electrical and Electronics Engineers Transactions*, mayo 1972, volumen IT-18, págs. 363–378.
- 21 M. S. Mueller y J. Salz: A Unified Theory of Data-Aided Equalization: *Bell System Technical Journal*, noviembre 1981, volumen 60, págs. 2023–2038.
- 22 P. A. Reatz, S. P. Pope, B. Solberg y R. W. Broderesen: Computer Generation of Digital Filter Banks: *Digest of Technical Papers*, Institute of Electrical and Electronics Engineers International Solid State Circuits Conference 1984, págs. 20–21.

James G. Dunn nació en Filadelfia, Pensilvania, y recibió el grado BSEE en la Universidad de Rutgers y los grados MSEE y ScD en la Universidad de Columbia. Entró en los ITT Federal Laboratories en 1956 y trabajó en transmisión digital y proceso de voz en Nutley, Nueva Jersey, hasta 1977. Se trasladó luego a San Diego, inaugurando el laboratorio de la Costa Oeste de ITT Defense Communications Division, en el que fue director técnico. En 1983 el Dr. Dunn fue nombrado científico ejecutivo de ITT.

Codificación del habla a 800 y 400 bit s⁻¹

Los sistemas de comunicación oral en banda estrecha requieren velocidades binarias menores que las normales de transmisión del habla. Están en avanzado desarrollo métodos más eficaces de compresión de las señales de voz, habiendo conseguido un codificador de voz operacional a 800 bit s⁻¹ y previendo en un futuro próximo una versión a 400 bit s⁻¹.

J. Carmody
J. Rothweiler

ITT Defense Communications Division,
Nutley, Nueva Jersey, Estados Unidos
de América

Introducción

La codificación del habla es importante en muchos sistemas de telecomunicación, como se expone en otros artículos de este número. En particular, cierto número de aplicaciones futuras de los sistemas de comunicación oral en banda estrecha exigirán que las señales de voz se transmitan a menor velocidad binaria para asegurar la integridad de los canales de comunicación frente a un comportamiento dudoso (p. ej., ruido aditivo o interferencia). Para satisfacer esta necesidad, ITT DCD (División de Comunicaciones para la Defensa) está desarrollando métodos más eficientes de compresión de las señales de voz.

La estrategia básica de la División para conseguir velocidades comprendidas en el margen de 800 a 400 bit s⁻¹, se basa en un algoritmo ya existente de modelación del habla para calcular una serie de parámetros, y en realizar después una codificación de fuente adicional para disminuir la velocidad binaria. La primera operación, o cálculo de los parámetros, la realiza un codificador por predicción lineal a 2400 bit s⁻¹.

Revisión de los principios de la codificación LPC

La codificación por predicción lineal (LPC) utiliza un modelo paramétrico del aparato de fonación humano. Este modelo (Fig. 1) considera que la producción del habla está controlada por tres factores: la fuente de excitación, la energía de la señal (o ganancia) y la forma de la cavidad acústica (desde la glotis a los labios). Las señales habladas pueden ser sonoras (como la "a") o sordas (como la "s").

El mecanismo de excitación para sonidos sonoros viene modelado por una serie de impulsos separados por el periodo del tono fundamental, mientras que, para sonidos sordos, su modelo es un generador de ruido. En el codificador de voz LPC, la fuente de excitación está afectada por un indicador que señala si hay que utilizar el modelo sonoro o el sordo. Para muchas aplicaciones, el indicador y el valor del periodo fundamental son suficientes para describir el mecanismo de excitación. La cavidad acústica se modela como un filtro lineal, caracterizado por una serie de coeficientes.

El método LPC reconstruye la voz humana transmitiendo la ganancia, los datos sonoros o del periodo fundamental y los coeficientes del filtro. El modelo de la señal de voz muestreada a la entrada consiste en una señal de voz obtenida en la salida y una señal de error o residual. Si ambas señales, la modelada y la residual, se recogen en el receptor, se consigue una señal de habla con baja distorsión (es decir, de alta fidelidad). El tren de bits LPC utilizado se ajusta al "US Federal Standard 1015". La tabla 1 enumera las características normales para este codificador LPC a 2400 bit s⁻¹.

Figura 1
Modelo paramétrico del habla.

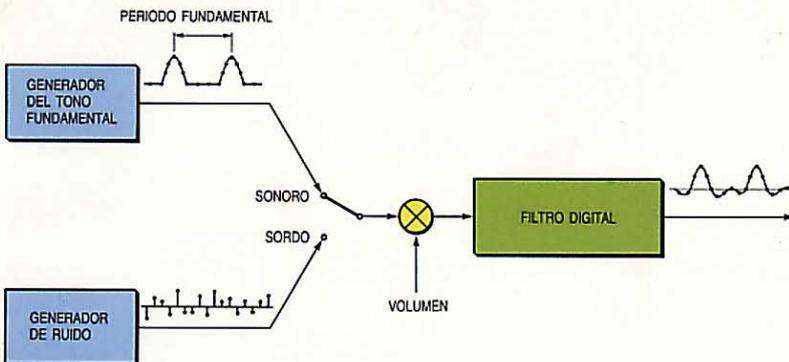


Tabla 1 — Características clásicas LPC

Número de muestras a la entrada:	180
Velocidad de muestreo:	8 kHz
Velocidad de trama:	44,4 tramas/segundo
Número de bits por trama:	54
Longitud de trama:	22,5 ms
Número de coeficientes:	
— (sonoros)	10
— (sordos)	4
Método de cálculo	Covarianza

Tabla 2 — Distribución de los bits para 2400, 800 y 400 bit s⁻¹

	Bits espectrales	Bits del tono	Bits de ganancia
2400	41	7	5
800	10	5	2,6
400	7 (promedio)	1	1

Todos los sistemas LPC actuales en tiempo real utilizan los datos de ganancia, de tono/sonorización y de los coeficientes; típicamente operan a unos 2400 bit s⁻¹. ITT DCD ha construido diversos sistemas LPC en tiempo real. Una de las arquitecturas utiliza procesadores separados para las señales de voz y las que se transmiten por módems. La tabla 2 presenta las principales características del algoritmo de la voz.

La figura 2(a) representa las frecuencias de un fragmento de señal de voz, después del paso por un filtro paso bajo y el muestreo a 8 kHz, sin ningún procesamiento posterior. En cambio, en la figura 2(b) el mismo fragmento ha sido procesado por un LPC de 10 coeficientes (LPC-10). Ambos gráficos se asemejan mucho.

Los datos espectrales LPC se codifican mediante cuantificación escalar, es decir, cada coeficiente se codifica por separado. Los 41 bits que sirven para codificar dichos datos producen una familia de más

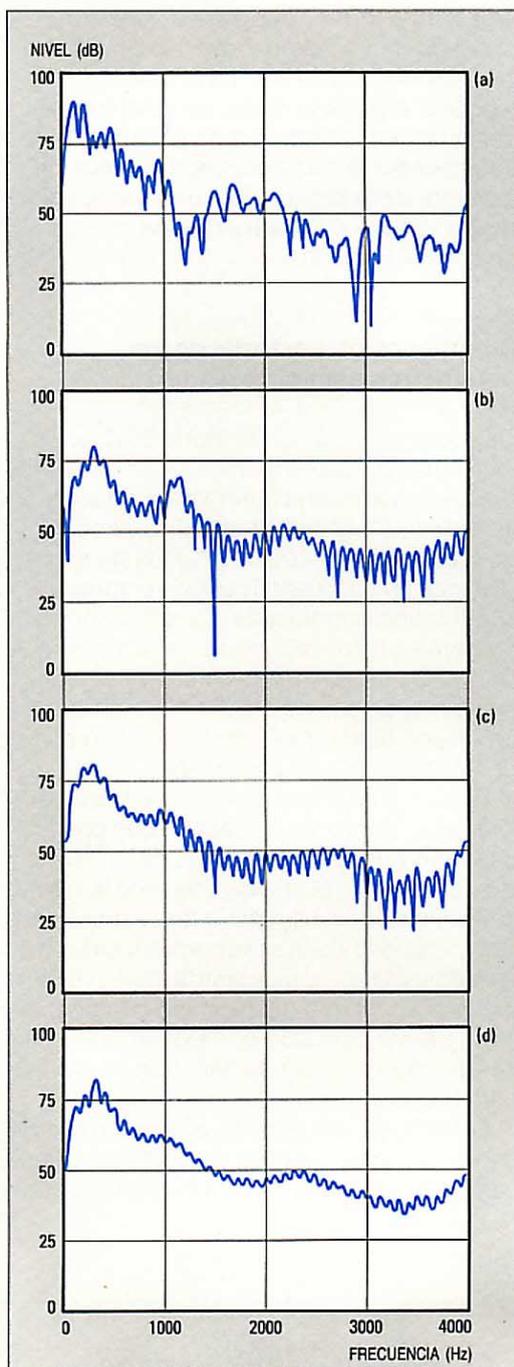


Figura 2
Gráficos de frecuencia de fragmentos de habla, después de someterlos a:
(a) filtrado de paso bajo y muestreo a 8 kHz
(b) procesamiento por filtro LPC de 10 coeficientes
(c) simulación a 800 bit s⁻¹
(d) simulación a 400 bit s⁻¹.

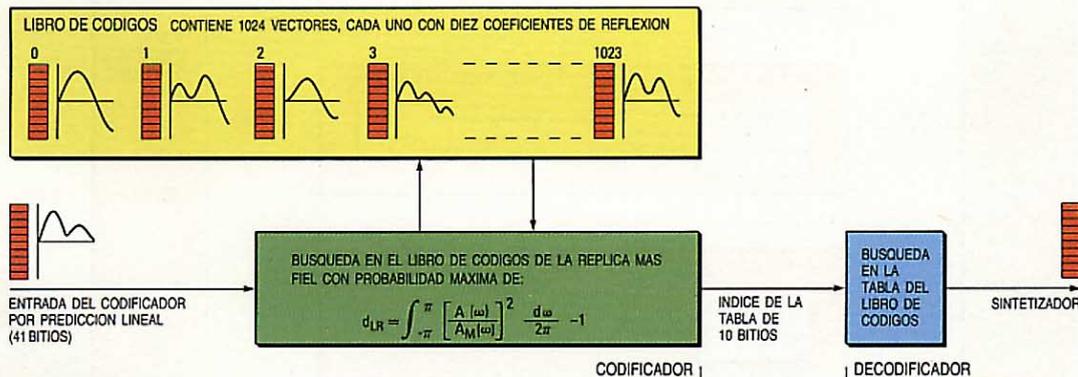


Figura 3
Codificador/decodificador de voz de baja velocidad.

de 2×10^{12} formas espectrales, número muy superior a las encontradas en las conversaciones ordinarias. La codificación en bloques, o cuantificación vectorial, representa al mismo tiempo todo el conjunto de coeficientes. En un cuantificador vectorial, el vector decimensional de los coeficientes se incluye en una palabra de código única (ver figura 3).

Cuantificación vectorial de los parámetros espectrales

Desde la aparición de la teoría LPC, la investigación sobre técnicas más eficientes de codificación de la fuente ha conducido al desarrollo de codificadores de voz en el margen de 800 a 400 bit s⁻¹. Dos de estas técnicas son la cuantificación vectorial y la codificación entrelazada. La reducción del número de bits requeridos para representar un fragmento de información permite mayor capacidad del "canal", ya sea éste un enlace electromagnético, físico, o parte de una memoria de procesador.

La compresión de los datos, o codificación de la fuente, es el proceso que codifica una determinada información de un modo más compacto. Suele decirse que la forma comprimida se ha desprovisto de redundancia. La calidad de la señal reproducida con respecto a la señal de entrada mide la fidelidad o la distorsión del proceso de compresión. Se han aplicado con éxito técnicas de compresión a sistemas para transmisión de datos, voz e imagen.

El habla es, en general, altamente redundante, por lo que ciertas características no cambian radicalmente durante cortos perio-

dos (alrededor de 20 a 25 ms). Esta observación, unida al gran número de posibles espectros LPC, indujo a buscar una codificación más eficaz de tales espectros. En ITT DCD se han investigado tres métodos: una transformada de Karhunen-Loeve, un codificador de voz por formantes y un cuantificador vectorial. La alta eficacia de este último para codificación de la fuente se basa en la teoría correspondiente de Shannon, en virtud de la cual se consigue mayor fidelidad codificando un vector que codificando separadamente los escalares. El cuantificador vectorial es de fácil realización, ya que puede ir unido en tándem a un codificador LPC.

La cuantificación vectorial consiste en buscar el emparejamiento de patrones. Dado un vector de entrada de los datos espectrales, se selecciona la mejor réplica del mismo entre un conjunto finito de vectores almacenados (plantillas). Por dos caminos se mantiene una buena fidelidad: los espectros que representan señales no habladas no se incluyen en la biblioteca, y los grupos de espectros no discernibles al oído se comprimen en un solo vector representativo. Una vez escogido el patrón más fiel, sólo hay que transmitir la dirección, o palabra de código de ese vector.

Se ha diseñado el cuantificador vectorial como módulo reductor de velocidad capaz de utilizarse en los sistemas LPC existentes de 2400 bit s⁻¹. Se ha decidido también procesar por separado la ganancia y los datos espectrales, con el fin de restringir toda búsqueda de dichos datos a un número relativamente bajo de plantillas. Aunque haya efectos que dependen de la ganancia en diversos sonidos, general-

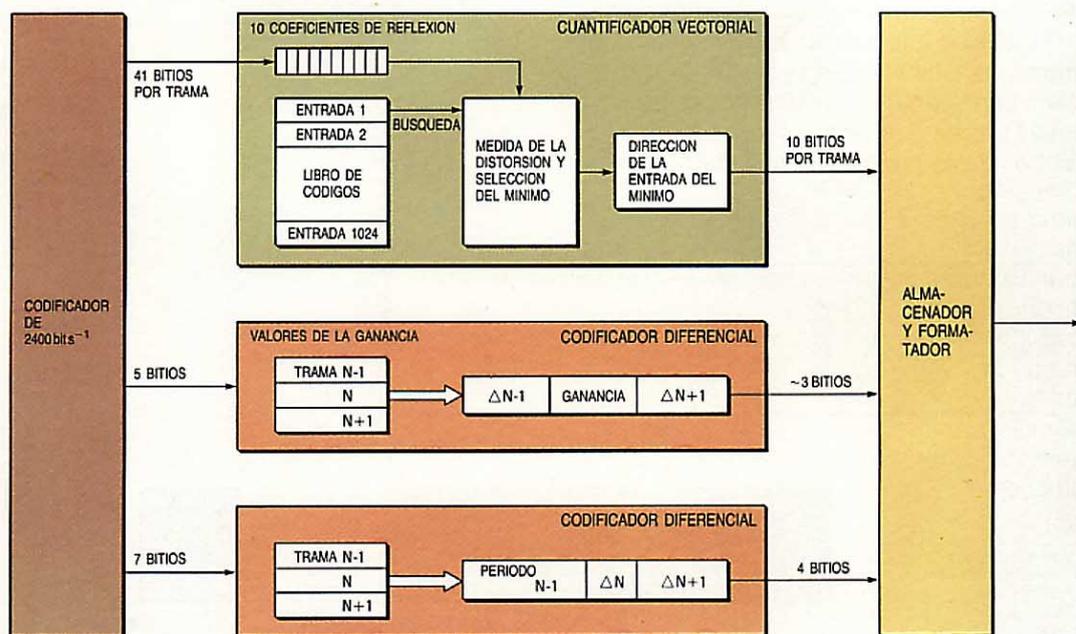


Figura 4
Cuantificador vectorial
a 800 bit s⁻¹.

mente es válida la separación de los datos espectrales y la ganancia. La figura 4 ilustra el método básico seguido: los datos espectrales son la fuente codificada por el cuantificador vectorial, mientras que los del tono y la ganancia se codifican por separado; los datos de sincronización se añaden al tren de bitios de salida.

La cuantificación vectorial reduce el número de bitios a transmitir, ya que muchos posibles vectores de entrada producen el mismo vector de salida. La diferencia entre los vectores de entrada y de salida es un término de error (distorsión) que deteriora la calidad del habla, y que, por tanto, se pretende reducir al mínimo valor posible.

Podían utilizarse distintas medidas de distorsión en el cuantificador vectorial, eligiéndose entre ellas para simulación la LRDM (medida de la distorsión de la relación de probabilidad), la cual mide la similitud en densidad espectral de potencia entre el vector de entrada y el vector patrón. Para la realización en tiempo real, se ha seleccionado la LARDM (medida de la distorsión de la relación logarítmica de las áreas) con preferencia a la LRDM, pues minimiza el cálculo sin afectar apreciablemente a la inteligibilidad.

Consideraciones sobre el diseño del libro de códigos

La calidad de un cuantificador vectorial viene controlada por la calidad del LPC original y el contenido del libro de códigos (los datos de voz utilizados para construirlo). El presente estudio sólo se ocupa del segundo de dichos factores.

Se plantearon algunas preguntas: el tamaño del libro de códigos, las plantillas de voz que debería contener, y su estructuración. Otras cuestiones todavía en estudio son si deberían hacerse los libros de códigos en el entorno en que se van a utilizar y cómo habría que construirlos.

La figura 5 representa el tamaño del libro de código en función de la distorsión. Cuanto menor sea la distorsión entre la señal de voz de entrada y la salida del cuantificador vectorial, mayor será la inteligibilidad*. En dicha figura se aprecia que la distorsión disminuye al ampliar el libro de códigos, lo cual es lógico porque se aumenta el contenido de formas espectrales. No obstante, hay escasa diferencia cualitativa entre un libro con 4096 componentes (palabra de código de 12 bitios) y otro con 1024 (palabra de 10 bitios), y no se

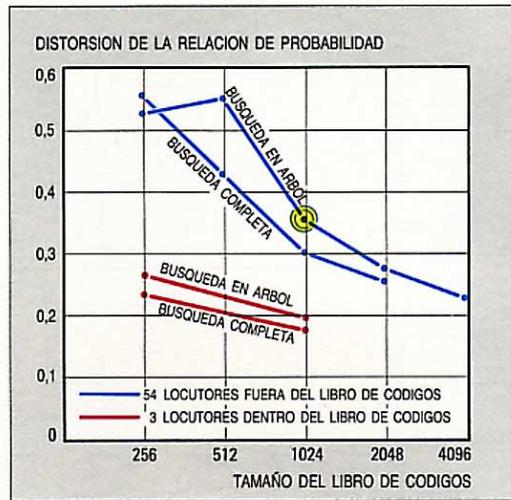
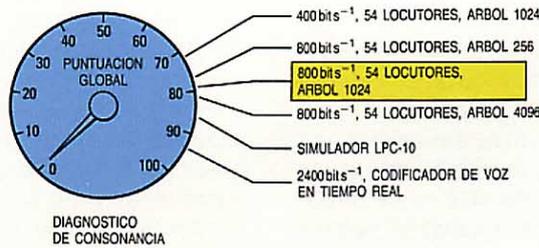


Figura 5
Resumen de los resultados de prueba: (a) pruebas comparativas que relacionan el tamaño del libro de códigos con la distorsión (b) puntuaciones totales en la prueba diagnóstico de consonancia.



puede apreciar ninguna ventaja subjetiva, por lo cual se ha escogido una palabra de código de 10 bitios.

Un proceso de emparejamiento consta de diez MACC (multiplicaciones acumuladas) por cada vector. Así, pues, la exploración completa de un libro de códigos de 1024 componentes requiere 10240 MACC. Una búsqueda en árbol consume menos cálculos; siguiendo este camino se puede realizar una búsqueda tosca sobre cada 32 componentes del libro de códigos completo (Fig. 6), indicando el mejor vector de esta búsqueda el correspondiente cuantificador vectorial de la rama más baja de esos 32 componentes, en la que se hace una segunda búsqueda. La figura 5 indica que la búsqueda en árbol reduce a 640 el número de MACC, a costa de perder muy poca inteligibilidad.

El libro de códigos del cuantificador vectorial codifica la fonación según un plan de

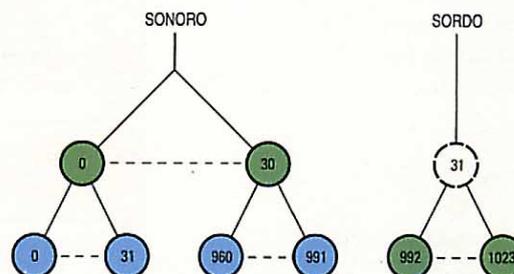


Figura 6
Estructura del libro de códigos para la cuantificación vectorial.

* La inteligibilidad se midió utilizando una prueba típica de elección forzosa entre dos palabras que difieren en un fonema; los resultados fueron registrados por un laboratorio independiente.

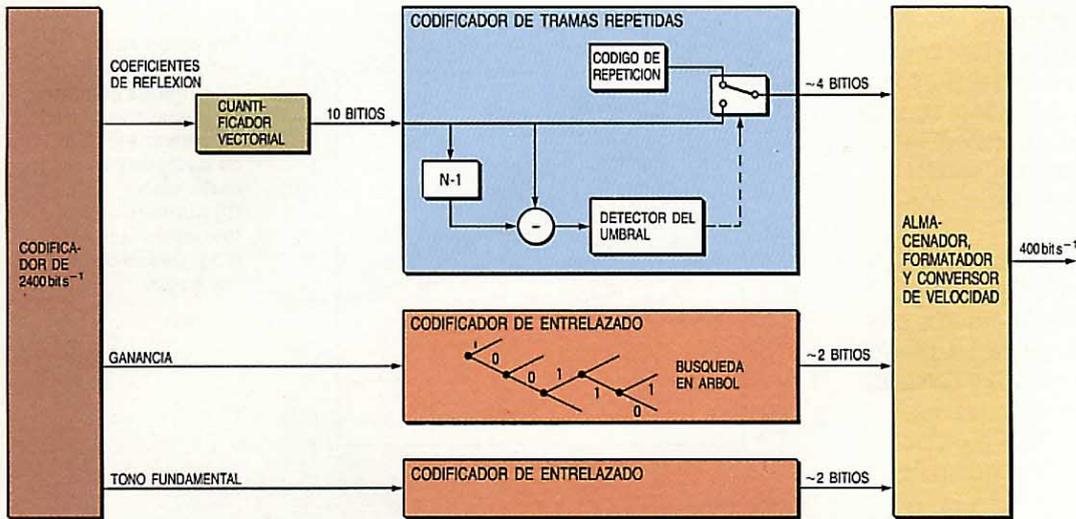


Figura 7
Cuantificador vectorial a 400 bit s⁻¹.

direccionamiento que asigna los vectores 0 a 991 a sonidos sonoros, y los 32 restantes a sonidos sordos. Este esquema permite transmitir los datos de la voz sin utilizar un bitio adicional. La estructura del libro de códigos adoptada para simulación se denominó 31:32/32 (los 31 nodos más altos para sonidos sonoros, con 32 vectores por "rama", más 32 vectores sordos).

La figura 5 muestra también que, cuando el locutor está "en" el libro de códigos (es decir, ha ayudado a producir las plantillas del mismo), la inteligibilidad es mayor que cuando no lo está. Sin embargo, aun en el último caso, el habla es comprensible y hay posibilidad de identificar al locutor.

Decidir quiénes han de estar en el libro de códigos es de gran importancia, pues un sistema independiente del locutor no requiere entrenamiento "in situ" cuando se añaden nuevos usuarios a la red. Si la población de locutores fuera suficientemente amplia en el comienzo, el sistema debería ser independiente del locutor.

Los resultados para los algoritmos de 800 y 400 bit s⁻¹ indican alguna ligera dependencia del locutor para aquéllos que se encuentran fuera del libro de códigos. Sin embargo, para los incluidos en el libro la inteligibilidad es mayor.

La cuantificación vectorial reduce la redundancia espectral dentro de la señal; mediante el libro de códigos 31:32/32 y codificación directa/diferencial se simuló un sistema de 800 bit s⁻¹ con trama fija, que funcionó, en tiempo real, a finales de 1984. La figura 2(c) ilustra una muestra de señal hablada a 800 bit s⁻¹ (simulación).

Cuantificación vectorial a velocidad variable

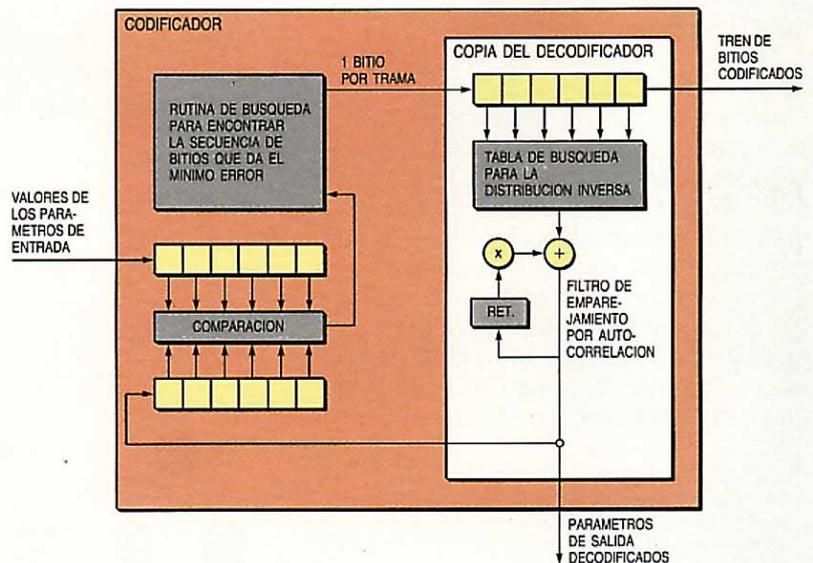
La supresión de la redundancia temporal, unida a la codificación entrelazada del tono

y la ganancia, permitió utilizar velocidades comprendidas entre 400 y 300 bit s⁻¹.

Para eliminar la redundancia temporal se utiliza un bitio de repetición de trama. En efecto, si la forma espectral no cambia apreciablemente de una trama a la siguiente, sólo hay que codificar una pequeña cantidad de información, y en tal caso el bitio de repetición se pone a "1"; en caso contrario, se pone a "0". Para conseguir un sistema de 400 bit s⁻¹, se requiere una repetición del 40%, valor razonable aunque tal sistema ya no sea de trama fija. La figura 2(d) presenta la muestra de voz a 400 bit s⁻¹ (simulación). La figura 7 presenta el sistema de 400 bit s⁻¹ de ITT DCD.

La inteligibilidad de este sistema es aceptable, aunque las puntuaciones obtenidas sean más bajas que las correspondientes al algoritmo de 800 bit s⁻¹. Los efectos dependientes del locutor y del vocabulario pueden hacerse notar a velocidades inferiores a 400 bit s⁻¹.

Figura 8
Codificador de entrelazado para el sistema de 400 bit s⁻¹.



La tabla 2 enumera las características esenciales de los algoritmos de 800 y 400 bit s⁻¹.

Codificación de los parámetros de excitación

La codificación del tono o periodo fundamental y de la ganancia a 800 bit s⁻¹ utiliza una técnica directa/diferencial. Los términos de tono y de ganancia se tratan por separado. Se adquieren primero tres tramas de datos. En el caso del término de ganancia, la correspondiente a la trama central se codifica en cinco bits. Las dos tramas adyacentes se restan luego del valor de la central, y las diferencias resultantes se codifican mediante una tabla de tres bits. El tono se trata de forma similar. La codificación directa/diferencial produce un sistema de tramas fijas.

A 400 bit s⁻¹, se utilizan codificadores de entrelazado para tratar la información del tono y la ganancia a la muy baja velocidad de un bitio por trama de análisis de 22,5 ms (Fig. 8). Estos decodificadores incluyen una tabla que contiene un conjunto fijo de valores de salida. El tren de bitios entrante hace que un registro de desplazamiento seleccione una secuencia de valores de esa tabla que consigan aproximarse a la forma de onda deseada en la salida.

El codificador de entrelazado usa una búsqueda en seis tramas para encontrar la secuencia de bitios que minimice el error medio a la salida del decodificador. Para ello, un algoritmo examina las secuencias posibles y selecciona la que case mejor con la entrada de los datos. La eficacia en la codificación es máxima porque este codificador se adapta tanto a las características espectrales como a la distribución estadística de la secuencia de datos que se está tratando. Además, el procedimiento de búsqueda minimiza el error medio sobre un amplio fragmento de la señal, reduciendo aún más el error de codificación frente a un sencillo cuantificador sin memoria.

Funcionamiento en tiempo real

En 1984, ITT DCD comenzó a utilizar el cuantificador vectorial de 800 bit s⁻¹ en tiempo real. En esta realización se cambia el módulo de reducción de velocidad por un módulo procesador de módem.

El codificador de voz inicial de 800 bit s⁻¹ consta de dos módulos: el procesador LPC-10 y el conversor de velocidad. En su



Figura 9
Arquitectura general para el cuantificador vectorial en tiempo real de ITT DCD.

estado final estará constituido por un solo módulo, capaz de tratar velocidades múltiples (2400 bit s⁻¹ e inferiores).

En la figura 9 se muestra la operación del codificador de voz referido. El tren de bitios LPC tradicional de 2400 bit s⁻¹ se decodifica para obtener los datos del tono fundamental, la ganancia y la relación logarítmica de las áreas. Estos últimos se leen directamente a partir de una tabla de conversión que toma los coeficientes del filtro como índices. La cuantificación de voz en estas relaciones de las áreas se efectúa mediante la LARDM. En el codificador de voz se utiliza la misma estructura del libro de códigos que en la simulación (31:32/32). El conversor de velocidad añade una trama de retardo a la salida, tanto en el transmisor como en el receptor; este retardo proviene de la decodificación de la trama LPC por el conversor de velocidad para la cuantificación vectorial, y de la función inversa en el lado receptor. No se requerirán tramas adicionales en el sistema de la nueva generación. El tono fundamental y la ganancia se codifican mediante el método directo/diferencial.

El codificador de 800 bit s⁻¹ estuvo disponible para pruebas de campo en noviembre de 1984; para 1985 está programada la operación en tiempo real de un codificador de voz de 400 bit s⁻¹.

Jack Carmody pertenece con categoría superior al gabinete técnico de ITT DCD, en Nutley, Nueva Jersey. Ha trabajado para ITT durante más de seis años en el grupo de ingeniería de sistemas para el proceso de voz, donde es un destacado investigador en la tecnología de codificación del habla a bajas velocidades binarias. Recientemente el Sr. Carmody ha recibido un Premio de la Asociación para el Desarrollo de ITT; durante 1985 en la Universidad de Stanford se dedicará a desarrollar un nuevo algoritmo para la codificación del habla a muy baja velocidad.

Joseph Rothweiler se graduó BS en 1974 y MEng en 1977, ambos por la Universidad de Louisville. Desde 1975 hasta 1982 estuvo asociado al Laboratorio de Comunicación Avanzada de RCA, donde fue responsable del desarrollo de sistemas de proceso digital de señales y de reconocimiento de patrones, así como de la programación para aplicaciones de comunicaciones y vigilancia. En 1983, entró en el departamento de ingeniería de sistemas de proceso de voz en ITT DCD, donde ha participado en el desarrollo de técnicas para la codificación del habla a bajas y medias velocidades. El Sr. Rothweiler es miembro del IEEE.

Proceso de la voz en centrales telefónicas públicas

El proceso de la voz perfeccionará el interfaz humano entre el abonado y la central, facilitando el que los abonados disfruten de las ventajas de los nuevos servicios, que se introducirán en los años inmediatos. Las técnicas de codificación a bajas velocidades binarias están demostrando ya su buen rendimiento económico en aplicaciones tales como máquinas parlantes, reemplazando a los dispositivos electromecánicos existentes.

R. Christensen
Ø. Johnsen

Centro de Investigación de Standard Telefon og Kabelfabrik A/S, Oslo, Noruega

B. Patovan

Bell Telephone Manufacturing Company, Amberes, Bélgica

Introducción

Cuando se inventó el conmutador telefónico, su función primaria era sencilla: que pudieran comunicarse dos personas situadas en lugares diferentes. Al correr los años, tecnologías nuevas han suscitado el desarrollo de una serie de servicios concebidos para hacer más eficaz la comunicación. De ahí que el acceso a los diversos servicios telefónicos se haya complicado cada vez más, resultando que rara vez se hace uso de muchos servicios útiles, como el de desviación de llamadas. La lección es clara: pese al esfuerzo aplicado a crear nuevos servicios, sólo los utilizarán unos pocos abonados, a menos que se mejore el interfaz con el usuario.

Gracias a las ventajas de la tecnología de proceso de la voz se puede ahora hacer mucho más agradable al usuario dicho interfaz, fomentando con ello el uso de esos nuevos servicios. Por otra parte, esta tecnología va a dar origen a otros servicios de

comunicación, tales como mensajes hablados para ausentes y el correo de voz.

Sistemas de proceso de la voz

El proceso de la voz en la conmutación para telecomunicación se puede dividir en dos categorías: codificación y reconocimiento del habla.

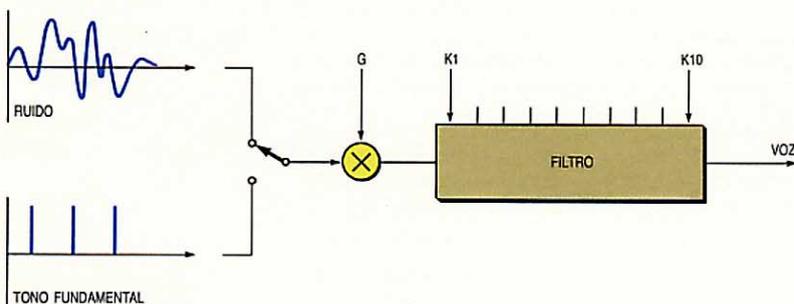
Codificación entre 1 y 4 kbit s⁻¹

Cuando se codifica a 1 ó 2 kbit s⁻¹, es posible almacenar las señales de voz económicamente en una memoria digital, facilitando la producción de voz de modo totalmente electrónico. La calidad del habla es bastante menor que cuando se codifica en MIC a 64 kbit s⁻¹, pero es suficiente para aplicaciones tales como las máquinas parlantes.

Una de las técnicas de codificación más atractivas es la de LPC (codificación por predicción lineal), basada en la forma de producirse el habla. Los sonidos se pueden dividir en dos grandes clases: sonoros y sordos. Se produce un sonido sonoro cuando una señal periódica procedente de la glotis se conforma a lo largo del tracto vocal. En cambio, el sonido será sordo cuando el tracto vocal conforma el ruido creado por turbulencia en la boca.

En la figura 1 se observa un modelo eléctrico de este mecanismo. Se hace pasar una señal procedente de una de las dos fuentes a través de un filtro, descrito por diez coeficientes ($k_1 \dots k_{10}$), que representa el tracto vocal. Una de las fuentes

Figura 1
Principios de síntesis de voz utilizando la codificación por predicción lineal.



produce ruido blanco, mientras que la otra produce una señal periódica caracterizada por dos parámetros. Las fuentes y el filtro se pueden describir mediante 12 parámetros (aproximadamente, 50 bits); si se cambian éstos cada 25 ms puede obtenerse una señal de voz bastante aceptable a unos 2 kbit s^{-1} , cuya calidad real dependerá del acierto en el cálculo de esos 12 parámetros.

Con el fin de lograr la mejor calidad de voz posible, el Centro de Investigación de STK ha instalado el sistema representado en la figura 2. El primer procesador calcula los 12 parámetros que se cargan al mini-ordenador para su corrección manual. La calidad del habla resultante es lo bastante buena para el uso en máquinas parlantes destinadas a sistemas de telecomunicación.

Codificación entre 8 y 64 kbit s^{-1}

Las nuevas técnicas de codificación a 64 kbit s^{-1} son capaces de producir un habla de alta calidad, necesaria en aplicaciones como las teleconferencias, generando el espectro completo de frecuencias vocales desde 20 Hz a 20 kHz.

Las técnicas que operan entre 8 y 32 kbit s^{-1} , darán en el futuro una calidad del habla semejante a la de la MIC a 64 kbit s^{-1} . Igual que los métodos de codificación de voz de alta calidad, estas técnicas de menor velocidad se basan en estadísticas de la señal de voz. La principal aplicación de los codificadores de 8 kbit s^{-1} será a sistemas de radio móvil digitales y para almacenamiento digital del habla en sistemas de correo de voz. Las nuevas reducciones de coste previstas en las memorias permitirán utilizar estos métodos de codificación en máquinas parlantes dentro de los próximos cinco años.

Reconocimiento del habla

Las máquinas generadoras de voz que utilicen métodos de codificación a velocidades comprendidas entre 1 y 10 kbit s^{-1} perfeccionarán el interfaz entre los abonados y el sistema de conmutación telefónico. El sistema podrá "hablar" con el abonado y pedirle que seleccione el servicio requerido, utilizando el abonado los 12 botones del teclado telefónico para responder a las preguntas hechas por el sistema. Sin embargo, cuando los sistemas de reconocimiento de voz resulten económicos, el abonado podrá contestar con órdenes verbales, y la gama de respuestas posibles será más amplia.

Las técnicas de reconocimiento de voz se basan esencialmente en variaciones temporales de la distribución de frecuencias de los sonidos en palabras y frases, o en variaciones de los coeficientes (k_1 a k_{10} en la figura 1). Podrá reconocerse una palabra

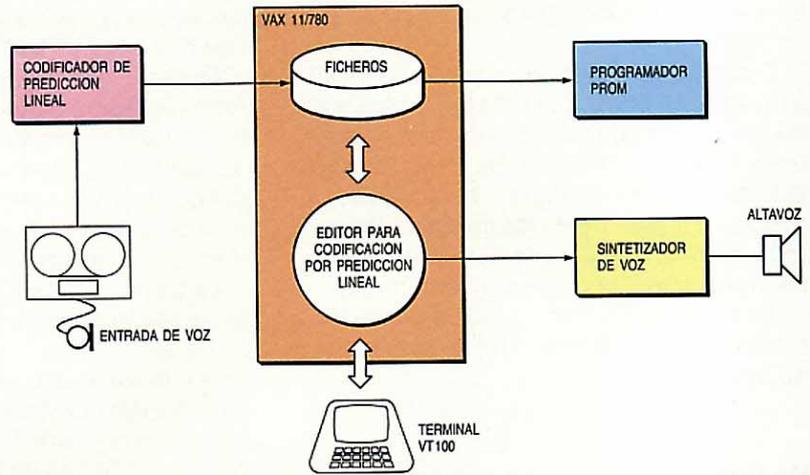


Figura 2
Laboratorio de voz
en STK.

pronunciada comparando sus parámetros con los de palabras incluidas en un vocabulario. Si el sistema no depende del locutor, este vocabulario constará de una gran cantidad de datos, y la comparación exigirá una considerable potencia de proceso. En consecuencia, tales dispositivos de reconocimiento habrían de instalarse en la central, si bien pueden equiparse en los teléfonos de abonados dispositivos más pequeños, dependientes del locutor¹.

El uso de sistemas de reconocimiento y generación de voz en las centrales facilitará mucho al abonado la elección entre diferentes servicios telefónicos, lo cual es importante dada la gran variedad de servicios de comunicación que están apareciendo como resultado de la nueva tecnología.

Realización tecnológica

Hoy día se realiza codificación de la voz a velocidades de 16, 32 y 64 kbit s^{-1} , aunque la calidad de voz dada por la modulación delta a 16 kbit s^{-1} es bastante baja si se compara con las técnicas de codificación actuales. La normalización de las nuevas técnicas en la banda de los 10 kbit s^{-1} no se completará hasta finales de los 80, cuando se hayan introducido los sistemas digitales de telefonía móvil. A corto plazo, se utilizarán técnicas no normalizadas en los sistemas de correo de voz para reducir el tamaño de memoria requerido y mejorar la calidad del habla.

La codificación a velocidades de 1 a 4 kbit s^{-1} se utilizará en máquinas parlantes que requieran vocabularios de más de 30 segundos de duración (unas 30 palabras), a fin de minimizar los costes de memoria. Las que tengan vocabularios menores pueden resultar económicas aun con técnicas de codificación a 64 kbit s^{-1} . En Noruega, las primeras máquinas parlantes a

baja velocidad de codificación se instalarán en 1985.

El desarrollo de las técnicas de codificación de voz viene impulsado por la limitación de las radiofrecuencias disponibles y el tamaño de memoria requerido en los sistemas de almacenamiento del habla. Análogamente, lo que fomenta las aplicaciones de reconocimiento de la voz es la necesidad de una entrada de datos conveniente. El coste y el tamaño físico de los equipos de reconocimiento determinará la extensión de su uso.

Máquinas parlantes con voz codificada por LPC

Hasta la fecha, las administraciones han utilizado magnetófonos y otros equipos electromecánicos para informar a los abonados sobre cambios en la guía, mediante mensajes como los siguientes:

“El número de este abonado ha cambiado. Por favor, marque el 003, servicio de información, o consulte la nueva edición de la guía telefónica”.

“El número llamado está fuera de servicio”.

Al utilizar métodos convencionales de grabación, la calidad de voz se degrada con el tiempo. Además se necesita un mantenimiento frecuente y una renovación regular de las cintas que conduce a altos costes del servicio.

El problema dominante en hacer enteramente electrónicos estos sistemas ha sido el coste de la gran cantidad de memoria que se requiere para los mensajes hablados. La técnica de LPC reduce la memoria necesaria en un factor de 30, posibilitando el desarrollar una máquina parlante económica. En STK (Standard Telefon og Kabelfabrik) se ha desarrollado un aparato basado en esta técnica, cuyo primer sistema de prueba se instaló en la red telefónica pública de Noruega en 1984.

Desarrollo del nuevo producto

Hace alrededor de dos años se inició un programa de desarrollo en el Centro de Investigación de STK para diseñar un aparato parlante experimental basado en la técnica LPC, con un vocabulario compuesto de palabras y frases relacionadas con centralitas y centrales telefónicas públicas. Concatenando relativamente pocas palabras y frases se puede generar un gran número de mensajes. Por ejemplo, “El abonado tiene un nuevo número de teléfono; el nuevo número es siete, nueve,

cinco, seis, cuatro, dos” se ha construido con dos frases y seis palabras.

Durante el programa de desarrollo, la Administración de teléfonos noruega y STK trataron del suministro del equipo de desvío de llamadas y servicio hablado. Se ofreció el equipo NST (nuevos servicios telefónicos) desarrollado por Standard Telefon und Radio, asociada suiza a ITT.

La tecnología LPC utilizada en el dispositivo experimental exige una velocidad binaria de sólo 2 kbit s⁻¹, lo cual parecía muy atractivo. Se probó el modelo experimental ante la Administración noruega, demostrando que la calidad de voz era satisfactoria para aplicaciones de telecomunicación y, como resultado, se decidió desarrollar un nuevo módulo de voz para el NST basado en el método LPC. En la figura 3 vemos cómo se ha montado este dispositivo de síntesis de voz en el equipo NST.

Aplicaciones posteriores

La nueva máquina parlante es el primer producto de telecomunicación en Noruega que utiliza técnicas de codificación a baja velocidad binaria. En un futuro próximo se estudiarán otras aplicaciones, en las que el habla podría sustituir o complementar a los tonos en sistemas de conmutación.

El modelo experimental se ha demostrado también a usuarios de centralita, interesados por algunos de los servicios hablados que puede ofrecer una PABX. Hasta ahora pocos usuarios potenciales han comprendido que las instrucciones orales dadas por la PABX producirán una mayor utilización de los nuevos servicios. Sin embargo, es probable que pronto se den cuenta de que sólo un pequeño grupo de abonados usan los nuevos servicios de

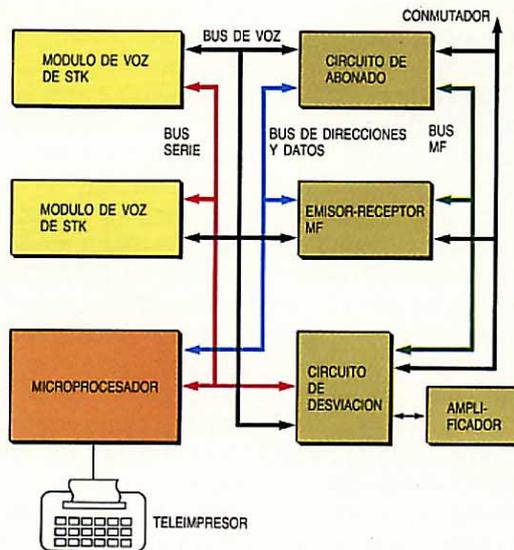


Figura 3 Señal de voz sintetizada incorporada al equipo NST. MF - multifrecuencia.

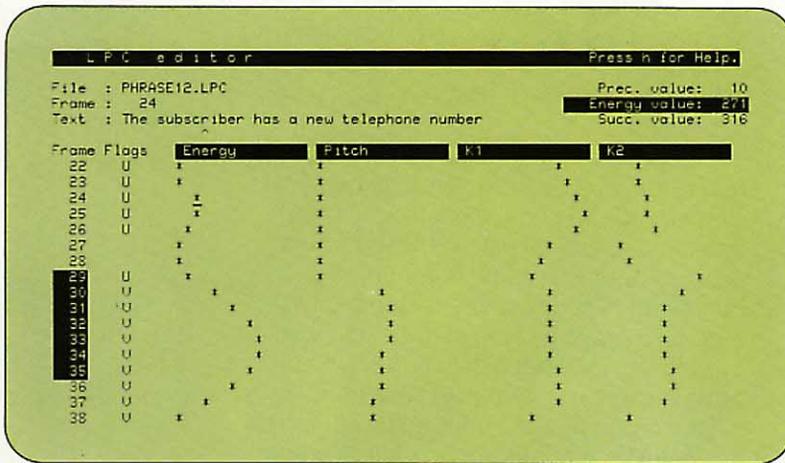


Figura 4
Visualizador en pantalla ilustrando cómo se utiliza el editor LPC.

conmutación (p. ej., repetición automática de llamada cuando un abonado llama a un número ocupado), y soliciten un mejor interfaz de usuario, basado en síntesis y reconocimiento de la voz.

Editor de la codificación por predicción lineal

La técnica LPC permite alcanzar una velocidad media de unos 1650 bit s^{-1} para voz codificada. Sin embargo, para asegurar que el habla sea inteligible a tan baja velocidad, la codificación debe ser correcta.

Dicha codificación se realiza en un primer procesador, mediante un algoritmo iterativo que no siempre converge al valor correcto en el tiempo de proceso concedido a cada muestra. Ello causa errores y degrada la calidad del habla, la cual puede no obstante mejorarse editando la señal de voz codificada. El editor LPC es una herramienta semi-gráfica e interactiva que STK ha desarrollado para corregir, modificar y analizar el habla codificada en LPC, mejorando así su calidad.

El editor está escrito en Pascal. La naturaleza discreta de los códigos LPC hace que sea suficiente la resolución de un terminal convencional de textos; en la figura 4 aparece una presentación típica en pantalla. El editor consta de dos partes fundamentales, la presentación y los modificadores.

Parte de presentación

Esta parte ofrece al usuario una presentación semi-gráfica de los parámetros numéricos LPC para energía, tono fundamental y los 10 coeficientes k . Estos coeficientes son argumentos de entrada a un filtro digital en celosía que simula el tracto vocal humano, y sus valores se pueden considerar como las áreas de las diferentes secciones transversales del tracto vocal, comenzando con k_1 en los labios. Los valores de los

parámetros se representan en función del tiempo, y existe un conjunto de parámetros, denominado trama, por línea, correspondiente a 25 ms de señal de voz.

Se ha procurado por todos los medios simplificar el uso del editor. La disposición en la pantalla es consistente, figurando todas las entradas de usuario en la línea superior de la misma. La zona de estado da una información relevante sobre el valor actual, el número de trama, el parámetro y el fichero. El texto correspondiente a la voz codificada se visualiza con un marcador que señala la posición relativa de la trama en curso en el texto. Como el editor es interactivo, muestra inmediatamente la nueva posición del valor modificado.

La columna del indicador señala si la trama es sonora, sorda, de silencio, o si se ha insertado una trama "mágica". Las tramas mágicas se utilizan para limpiar el filtro en celosía en el sintetizador de voz.

El editor puede trabajar en uno de los tres modos que definen el alcance de las órdenes:

- Modo fichero: todas las órdenes afectan a todas las tramas.
- Modo bloque: un bloque es un subconjunto del fichero definido por el usuario. Los números de tramas de un bloque se presentan en vídeo inverso. Todas las órdenes afectan a todas las tramas comprendidas en el bloque.
- Modo trama: las órdenes sólo afectan a la trama en curso. En el modo de trama única, el cursor permanece en la representación gráfica del valor actual.

El usuario es libre en un momento dado de maniobrar sobre una trama del fichero, sobre un carácter del texto, o de acceder directamente a cualquier número de trama. La edición se ejecuta sobre un parámetro cada vez; el usuario puede elegir qué parámetro se va a editar y cuántos se deberían visualizar. La escala a que se representan estos parámetros depende del número que se haya escogido, de su margen de variación y de la resolución del terminal. Si se visualizan los 12 parámetros a la vez, se aprecia bien la situación pero la resolución es escasa; en cambio, al visualizar sólo unos pocos parámetros, éstos se distinguen mejor pero no se tiene visión de conjunto.

Una de las propiedades más eficaces del editor es su capacidad para manejar bloques (es decir, un cierto número de tramas contiguas definido por el usuario). Como ya se ha mencionado, un bloque restringe el campo de acción de cada orden a las tramas incluídas en él, pero los bloques se pueden también desplazar dentro del fichero, escri-

bir en un nuevo fichero, leer o borrar. Esta capacidad de manejar bloques es especialmente útil cuando se juntan palabras de diferentes frases para construir otras nuevas. Para aumentar la calidad de voz, se puede cambiar una palabra de una frase por una versión mejor de la misma palabra procedente de otra frase.

Usando el sintetizador de voz de STK, el usuario puede escuchar, en cualquier momento, la versión actual del fichero. Es posible ralentizar el habla sintetizada para examinar con mayor detalle cada parte de la frase. Pueden sintetizarse también tramas aisladas para que el usuario escuche los sonidos que representan. La misma función sirve para preparar un fichero con el formato correcto para el programador de la memoria programable de sólo lectura (PROM). La parte de presentación facilita también ciertas estadísticas, como el valor promedio y la desviación típica del parámetro actual. En el caso de que el usuario encuentre dificultades en el uso del editor, existe una función de ayuda que señalará todas las órdenes permisibles en tal situación.

Modificadores

La parte de presentación ofrece un entorno uniforme para los modificadores, varios de los cuales se han realizado ya. La modularidad del soporte lógico simplifica la adición de nuevos modificadores. Se diferencian éstos de las funciones de la parte de presentación en que cambian los valores de los parámetros para una o más tramas. Es posible borrar o insertar tramas individuales; la trama insertada puede ser de tres

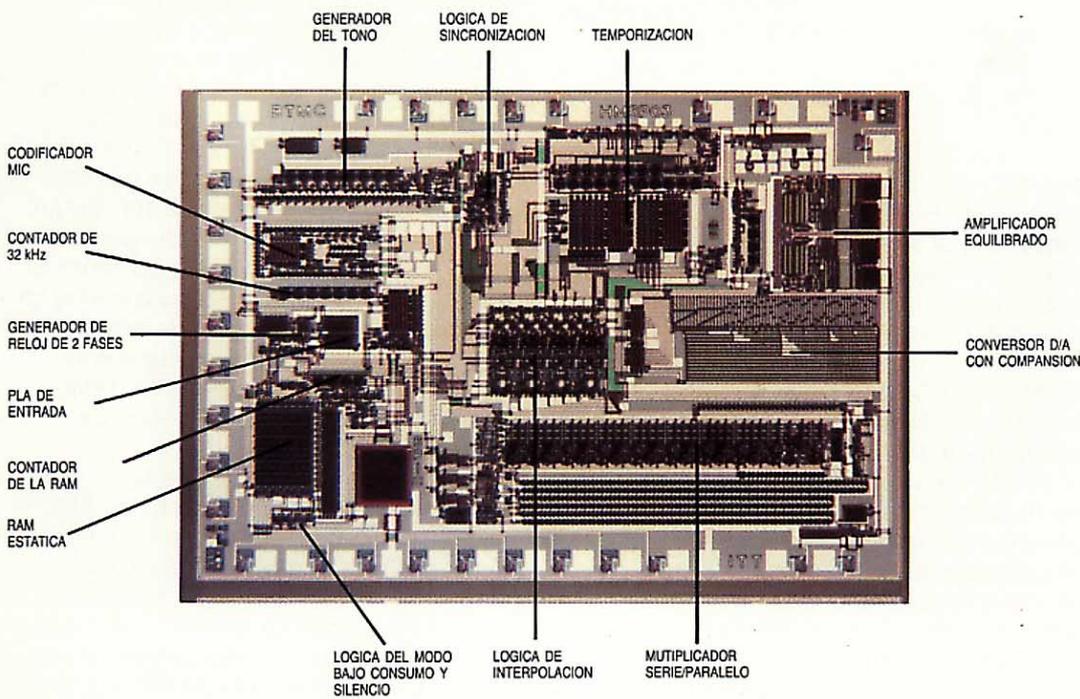
tipos: una trama de silencio que produzca una pausa, una copia de la trama precedente para prolongar un sonido, o una trama mágica para limpiar el filtro en celosía del sintetizador.

Al representarse gráficamente los valores, la mayoría de los errores de codificación se revelan como elevaciones o descensos bruscos de los valores de tramas simples. Esto se produce por un error común en la codificación del tono, cuyo efecto es grave, si bien se detecta y corrige fácilmente. La detección y corrección de otros errores requiere más experiencia.

Una orden puede cubrir en su campo de acción el fichero entero, un bloque o una trama. El efecto de cada modificador es más fácilmente observable cuando se trabaja en modo de trama simple. Puede aumentarse o disminuirse en uno el valor del parámetro actual, supuesto un valor especificado por el usuario, o bien calcularlo como la media entre los valores precedente y siguiente. Estos sencillos modificadores permiten corregir la mayoría de los errores en las tramas. Hay otros modificadores sólo utilizables en los modos de fichero y de bloque.

Los valores de los parámetros se pueden alterar en un factor especificado por el usuario, multiplicando por él todos los valores de modo que cambie el nivel global. Esto sirve principalmente para amplificar la energía, modificando no sólo el valor medio sino también la desviación típica de los valores de parámetros.

Hay, sin embargo, un modificador capaz de alterar el nivel de un modo más controlado, informando al usuario de los valores



Pastilla VLSI, realizada en tecnología NMOS de 3 μ m, para el proceso de voz de elevadas prestaciones. PLA - ordenación lógica programable.

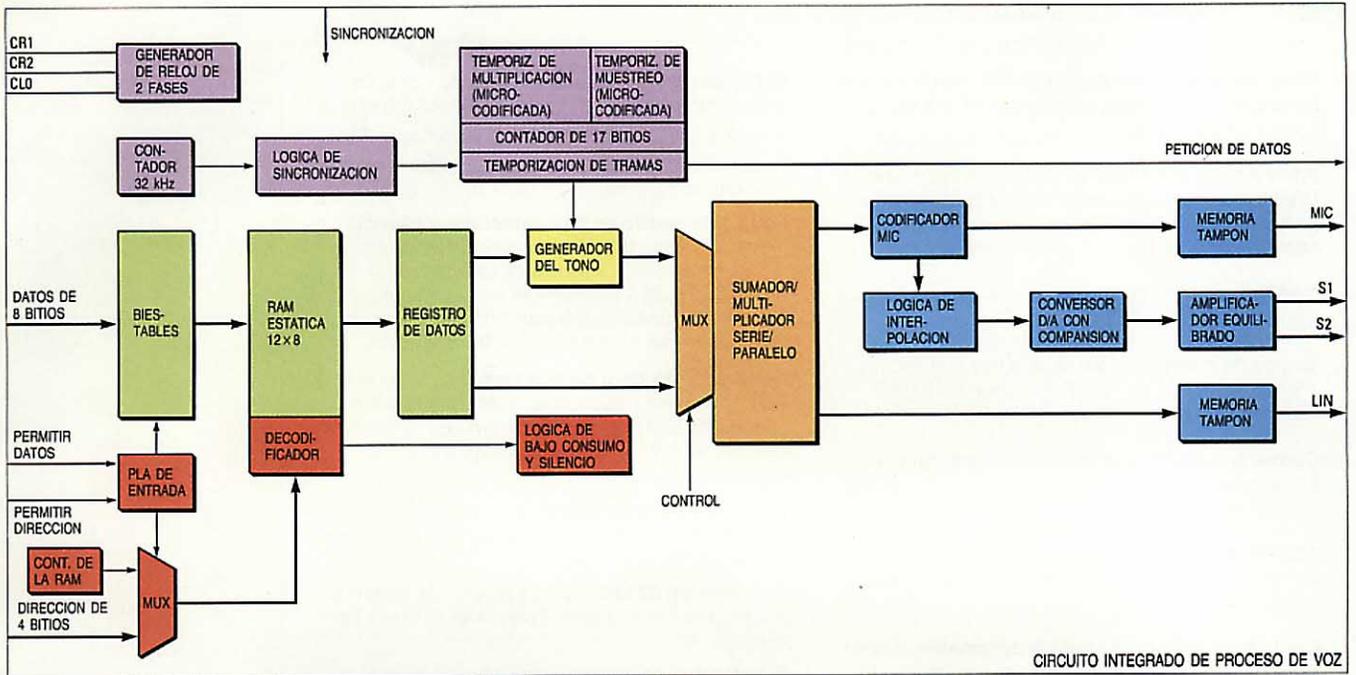


Figura 5
Diagrama de bloques del circuito integrado para proceso de voz.

estadísticos actuales y solicitándole que introduzca valor medio y desviación típica nuevos, tras de lo cual se cambian todos los valores de trama para adaptarse a los nuevos valores introducidos.

El efecto de estos cambios queda muy bien ilustrado por el parámetro del tono fundamental; al cambiar su valor medio, la voz se hace más grave o más aguda, mientras que cuando aumenta su desviación típica se amplía la gama tonal comprendida entre el tono más alto y el más bajo.

Estos modificadores son importantes, tanto para hacer más inteligible el habla como para equilibrar los niveles de los parámetros en frases diferentes.

Desarrollo en curso

Se está actualmente potenciando la presentación mediante una función que recoge y visualiza los fonemas que describen el sonido de la trama correspondiente.

Circuito integrado sintetizador LPC

Se ha desarrollado un circuito integrado para proceso (síntesis) de voz de elevadas prestaciones, basado en la técnica LPC, el cual utiliza tecnología NMOS de 3 µm. Esta pastilla puede utilizarse en una configuración de tres circuitos que, además del sintetizador, comprende un microprocesador y una memoria ROM externa de vocabulario. La calidad del habla se puede adaptar a los requisitos del usuario, variando la velocidad binaria entre la ROM del vocabulario y el microprocesador, desde 1,1 a

8,5 kbit s⁻¹. Como características específicas, están la síntesis sincrónica del tono fundamental, la posibilidad de interpolar parámetros de la voz y el modo de bajo consumo y silencio. En la figura 5 se aprecia el diagrama de bloques del circuito sintetizador, y en la tabla 1 se describen las funciones de los distintos bloques.

Para eliminar la necesidad de filtrado fuera del circuito se interpola la salida del filtro digital a una elevada frecuencia de muestreo (32 kHz). Se consiguen una salida MIC de 8 bits (ley A) y una salida de codificación lineal de 16 bits. El circuito sintetizador se puede entregar en dos configuraciones físicas distintas para usarlo, bien en pequeños sistemas (de tres pastillas), o bien en sistemas mayores.

Este nuevo sintetizador de voz de altas prestaciones incorpora características especiales, como la capacidad de interfaz con una gran variedad de microprocesadores y de ajustar la calidad de voz a los requisitos del usuario, variando la velocidad binaria dentro de un amplio margen. La interpolación de coeficientes se lleva a cabo a la mitad de una trama (11,25 ms). El área de la pastilla es sólo de 15 mm² y el filtrado de aplanamiento dentro del propio circuito integrado, previo a la conversión de digital a analógico, minimiza el número total de componentes externos.

Conclusiones

La tecnología de proceso de la señal de voz está permitiendo a los abonados acceder a

Tabla 1 — Funciones de los bloques del circuito sintetizador LPC

<p>Generación del tono: genera un impulso para tramas sonoras y ruido para las sordas. Las salidas se envían al filtro digital.</p> <p>Filtro digital: proporciona muestras de voz de 8 kHz. Comprende un multiplicador-acumulador en serie-paralelo y registros de desplazamiento para los coeficientes y los resultados intermedios.</p> <p>Codificador MIC: codifica muestras obtenidas del filtro digital.</p> <p>Lógica de interpolación: calcula tres resultados intermedios entre las muestras x y $x + 1$ a 32 kHz (acepta datos codificados).</p> <p>Conversor digital-analógico con compensación: convierte los datos interpolados codificados en MIC a forma analógica.</p> <p>Amplificador equilibrado: este amplificador con alimentación única (+5 V) genera la salida analógica del sintetizador.</p> <p>Conjunto lógico de entrada programable: proporciona un interfaz asíncrono entre el circuito y diferentes tipos de microprocesadores.</p> <p>RAM estática: recibe 12 multibitios por trama de voz, que corresponden a los 10 coeficientes del filtro LPC, el tono y la ganancia.</p>	<p>Multiplexado de dirección: conjunto lógico de entrada programable, que selecciona lectura/escritura – lectura: contador de dirección interna de la RAM – escritura: dirección externa enviada por el micro-procesador.</p> <p>Lógica del modo de bajo consumo y silencio: reconoce cuatro direcciones especiales. La de <i>prueba</i> todavía no se usa; la de <i>bajo consumo</i> obliga al circuito a permanecer en espera; el <i>modo silencio</i> produce una limpieza dinámica del circuito. El <i>modo voz</i> es el que se utiliza normalmente.</p> <p>Generador de reloj de dos fases: en la configuración de sistema pequeño, éste acepta un cristal para generación de reloj interna del circuito; en sistemas grandes, acepta un reloj mono-fase generado externamente.</p> <p>Lógica de sincronización: permite usar diferentes frecuencias de reloj sin cambiar el tono de la voz de salida.</p> <p>Contador de 32 kHz: genera la señal de 32 kHz en la configuración de sistema pequeño, utilizada para interpolación.</p> <p>Temporización general: consiste en un contador de 17 bitios, una parte de alta velocidad (18 ciclos de reloj por multiplicación), una de velocidad media (20 multiplicaciones por muestra), y una de velocidad baja (90 muestras por trama, pero en sincronismo con el tono fundamental).</p>
---	--

los servicios telefónicos mediante entrada oral y recibir una respuesta hablada. El resultado será que el usuario tenga un interfaz mucho más agradable con los refinados servicios de telecomunicación que ahora están apareciendo, garantizando así una generalizada aceptación por parte de los abonados.

STK ha desarrollado una máquina parlante que hace uso de la codificación por predicción lineal de la voz para minimizar la memoria requerida, asegurando al mismo tiempo una calidad de voz adecuada para esta aplicación. Dicha calidad se optimiza por medio de un nuevo editor LPC.

Referencia

1 M. Immendörfer: Marcador telefónico activado por la voz: *Comunicaciones Eléctricas*, 1985, volumen 59, nº 3, págs. 281 – 285 (en este número).

Rune Christensen nació en Sandefjord, Noruega, en 1954. Asistió al Colegio Técnico de Sarpsborg en 1975 y a la Universidad Técnica de Trondheim en 1977. En 1981 obtuvo el grado MS y entonces ingresó en el centro de Investigación de STK, donde inicialmente trabajó como ayudante investigador en proceso de voz. El Sr. Christensen es ahora director de proyectos del equipo de redes digitales y proceso de voz.

Øyvind Johnsen nació en Holmestrand, Noruega, en 1958. En 1982 obtuvo el grado BSc en informática por la Universidad de Heriot-Watt, en Escocia. Después de su graduación, ingresó en el Centro de Investigación de STK en Oslo, donde trabaja como ingeniero en el grupo de redes digitales y proceso de voz. El Sr. Johnsen es miembro de la British Computer Society y de la Association for Computing Machinery.

Bob Patovan nació en Borgerhout (Bélgica) en 1944. Recibió el grado de ingeniero químico en 1968; dos años después, ingresó en el departamento de formación de BTM como iniciador e instructor de un curso sobre soporte físico de ordenadores. En particular, ha organizado un curso de un año en Bandung, Indonesia. El Sr. Patovan pasó luego dos años adquiriendo experiencia sobre proceso en el laboratorio de investigación de semiconductores de BTM, antes de entrar en el grupo de diseño de VLSI.

Aplicación del reconocimiento y la síntesis de voz a los servicios de las PABX

La madurez de las tecnologías de reconocimiento y síntesis de voz hace que los diseñadores de sistemas de comunicaciones para oficinas las tengan muy en cuenta en el desarrollo de nuevas opciones para el interfaz hombre-máquina. Sin embargo, se debe prestar una atención esmerada a los factores humanos implicados para que estas nuevas tecnologías ofrezcan al usuario un interfaz agradable.

H. Mulla

ITT Advanced Technology Center, Shelton, Connecticut, Estados Unidos de América

J. F. Vaughan

ITT Business and Consumer Communications Division, Raleigh, Carolina del Norte, Estados Unidos de América

Introducción

La PABX (centralita automática privada) ha evolucionado, desde su humilde comienzo como cuadro conmutador, hasta convertirse en un sofisticado instrumento de comunicación que actuará como núcleo de los futuros sistemas telefónicos; el desarrollo de las PABX debe resolver el dilema de cómo conseguir facilidades avanzadas y mantener, a la vez, la simplicidad de la operación.

El problema básico es que, aun siendo ahora la PABX un potente procesador programable, se mantenga inalterado el enlace de comunicación entre ella y el usuario. Este ha de enviar órdenes mediante el gancho conmutador y secuencias de dígitos preestablecidos, y recibir señales de información relativa al estado por medio de tonos y cadencias de tono (tono de marcar, tono de ocupado, etc.). Una PABX moderna típica ofrece facilidades tales como reenvío de llamadas, retención por ocupación, conferencia y transferencia de llamadas. Aun así, la mayoría de los usuarios no utiliza estos servicios porque las combinaciones de teclas pulsadas requeridas para activar estas funciones son difíciles de memorizar. En muchos casos, los abonados ni siquiera conocen todas las facilidades que proporciona la PABX.

Los diseñadores del equipo telefónico han tratado de simplificar la situación introduciendo aparatos de abonado con teclado multifunción.

Siendo el habla el medio natural y más desarrollado de comunicación humana, el proceso de voz puede desempeñar un impor-

tante papel en las PABX y en los sistemas de comunicación para oficinas, realizando el interfaz hombre-máquina y consiguiendo sistemas de uso más agradable.

Algunas preguntas básicas

La utilización del proceso de la voz en los sistemas telefónicos plantea ciertas cuestiones básicas:

- ¿Debe incorporarse la capacidad de procesar la voz en el teléfono de abonado o de modo centralizado en la PABX?
- ¿Se necesita que el reconocimiento de voz sea independiente del locutor, o es razonable pedir a cada usuario que "adiestrate" al sistema a reconocer su voz?
- ¿Hace falta reconocer palabras enlazadas o basta con identificar palabras aisladas?

Las respuestas a estas preguntas tienen un gran impacto en el diseño del sistema.

Dos conceptos alternativos

Es fundamental decidir dónde localizar la capacidad de reconocimiento de voz: cada respuesta conduce a un concepto totalmente diferente. Un teléfono de abonado con reconocimiento de voz incorporado podría ser útil para ciertas aplicaciones, particularmente si el usuario estuviera físicamente incapacitado o tuviera inconveniente en utilizar las manos para manejar el teléfono. Tal instrumento no

tendría un interés universal y, para mantener razonable el coste, la complejidad y por tanto las prestaciones, tendrían que ser modestas. A menos que el consumo energético de los circuitos adicionales fuera mínimo y que incorporara algún tipo de batería, no sería posible utilizar la modalidad oral durante un fallo de alimentación.

Incluir la capacidad de procesar la voz en la PABX tiene varias ventajas. Como recurso compartido, un solo procesador de señales de voz puede servir para un gran número de usuarios y, por tanto, puede justificar un coste y grado de complejidad elevados, permitiendo mejores prestaciones. Por ser parte de la PABX, la unidad de voz es capaz de ofrecer un amplio conjunto de facilidades refinadas.

Esta última opción es la elegida en el Advanced Technology Center y en la Business and Communications Division, dado su extenso campo de aplicaciones y su gran alcance para los futuros sistemas de telecomunicación.

Independencia del locutor

Consideramos necesario que el reconocimiento de voz en el caso de las PABX sea independiente del locutor. Cuando se utiliza reconocimiento dependiente del locutor, las plantillas acústicas de cada usuario individual han de almacenarse en una base de datos central y ser reclamadas por el sistema al identificarse el usuario marcando un código especial o pronunciando su nombre. Para sistemas de tamaño medio o grande, sin embargo, estas soluciones están sujetas a los problemas de robustez y de adiestramiento, además de exigir una base de datos de tamaño prohibitivo.

Robustez

Las unidades para reconocimiento de voz que se basan en el adiestramiento por parte de los usuarios individuales suelen ser más sensibles a los cambios en las características de las señales orales recibidas. Tales cambios pueden obedecer a causas tales como tensión, fatiga o resfriado del usuario, o bien a diferencias en cuanto a propiedades electroacústicas entre el teléfono utilizado para el adiestramiento y el que se está usando para acceder al sistema. Las unidades independientes del locutor tienen una "robustez" intrínseca que, generalmente, les permite soportar tales variaciones, de lo que resulta un mejor comportamiento global.

Adiestramiento

El proceso de adiestrar un sistema dependiente del locutor en un vocabulario de

tamaño razonable, de unas 75 palabras, es una tarea monótona que requiere unos veinte minutos de constante concentración. La experiencia demuestra que, durante una sesión tan prolongada, una persona comete varios errores típicos de pronunciación que quedarían sin corregir. Además, a no ser que se usen técnicas especiales para obtener respuestas naturales del individuo, la forma de hablar en el adiestramiento es sustancialmente diferente de la utilizada normalmente, a causa de la ansiedad o del aburrimiento. De ahí que la calidad de las plantillas sea inferior a la que se obtendría en condiciones mejor controladas, como sucede en la toma de datos de voz para el reconocimiento independiente del locutor. En este último caso, el tratamiento estadístico utilizado en la generación de plantillas contribuye también a minimizar los efectos de los errores.

Si bien es cierto que en condiciones controladas la precisión intrínseca de las unidades de reconocimiento dependientes del locutor sobrepasa generalmente la que puede lograrse en aquéllas que no dependen del locutor, los dos problemas antes señalados harían que, en el entorno relativamente descontrolado de una aplicación práctica de PABX, los sistemas dependientes dieran resultados menos satisfactorios.

Uno de los mayores logros de los estudios sobre sistemas de ITT ha sido conseguir que el reconocimiento de voz no dependa del locutor.

Antes de dejar el tema de la independencia del locutor, es preciso aclarar el significado del término. En opinión de los autores, es inconcebible que cualquier sistema de reconocimiento de voz trabaje igual de bien para todos los locutores, sean cualesquiera las características de su tracto vocal, su acento o su forma de hablar. Quizás un término más apropiado sería el de "no sensible al locutor". En la práctica, las características exigidas son:

- No se requiere un adiestramiento explícito por cada individuo que desee utilizar el sistema. Este, sin embargo, puede adaptarse a la voz del usuario durante el uso normal.
- El buen resultado del reconocimiento superará un nivel mínimo aceptable para un alto porcentaje de la población de usuarios posibles. El resultado se medirá por el número de operaciones realizadas con éxito antes que por el número de palabras reconocidas correctamente, puesto que un sistema bien diseñado puede utilizar la sintaxis para corregir errores de reconocimiento. Es admisible, e incluso conveniente, adaptar la base de datos de las plantillas a una población de

usuarios específicos, en sistemas prácticos.

Pronunciación de palabras aisladas

Sin menoscabo de otros factores, sería deseable el reconocimiento de palabras enlazadas. Sin embargo, hasta ahora no es factible conseguirlo de un modo independiente del locutor con resultados aceptables y una razonable realización del equipo físico. Puestos a elegir entre una alta precisión en el reconocimiento y la capacidad de enlazar palabras, hay que escoger lo primero. En la aplicación actual, se puede admitir la exigencia de palabras aisladas, no desprovista de ventajas. Cuando el usuario sabe que debe pronunciar separada y distintamente cada palabra, suele articular más claramente que cuando no se impone requisito alguno. El uso de realimentación, el almacenamiento intermedio y las frases cortas para órdenes hacen bastante aceptable esta restricción.

La aplicación final a un producto combinará probablemente el reconocimiento "no sensible al locutor" con plantillas específicas dependientes del mismo. Estas se crearían "in situ" para formar una pequeña lista de números de marcación abreviada, específicos de cada usuario del sistema. La rapidez de la llamada aumentará la productividad de los usuarios del sistema hasta un nivel que justificará, casi por completo, el pequeño costo que supondrá añadir la posibilidad de marcación oral a la PABX.

Los ingenieros de ITT han desarrollado un sistema experimental de marcación oral con el que se pretende hallar respuestas a varios de los interrogantes más difíciles, capaz de evolucionar al compás de la tecnología y de la mejor comprensión de los aspectos humanos implicados en el reconocimiento de voz. La segunda generación del sistema se exhibió por primera vez en Telecom 83 (Ginebra), y después en Comunicasia 84. Una tercera versión, que se mantiene semejante a las anteriores, se ha integrado en una PABX Sistema 3100, en la primera de dos pruebas de campo proyectadas con usuarios reales.

Sistema de ATC

Instalación demostrativa

La instalación presentada en Telecom 83 y Comunicasia 84 consistió en una centralita con 10 teléfonos para una supuesta oficina comercial. La PABX se conectó a una central digital Sistema 12 y a la red pública telefónica suiza en Ginebra. Ocho

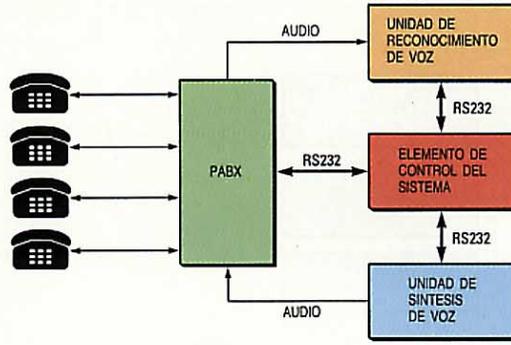


Figura 1
Principales elementos funcionales del sistema de ATC.

de los teléfonos tenían rótulos con nombre de persona, nombre de departamento y un número de cuatro cifras. Por ejemplo:

Mr. Jackson
Dept. de Personal
4165

Los dos teléfonos restantes sirvieron para demostrar las facilidades gobernadas por la voz, tales como:

- efectuar llamadas a personas, departamentos, compañías, líneas aéreas, servicios especiales, etc.
- retención por línea ocupada
- transferencia de llamadas de una extensión a otra
- petición de ayuda a la agenda
- petición de instrucciones para usar determinados servicios del sistema.

Aunque las anteriores posibilidades no sean todas las que ofrece una PABX, ni sean las más idóneas para aplicación a un producto, ilustran, sin embargo, la naturaleza del diálogo hombre-máquina que tendría lugar en un marco real.

Elementos funcionales

El sistema consta de cuatro bloques funcionales, como vemos en la figura 1:

- PABX
- unidad de síntesis de voz
- unidad de reconocimiento de voz
- elemento de control del sistema.

Para facilitar la introducción en el sistema de nuevos componentes o técnicas en cuanto vayan apareciendo, se mantiene una separación funcional entre los cuatro subsistemas, que se comunican entre sí por enlaces convencionales RS 232. Las funciones principales de los subsistemas se detallan en los párrafos siguientes.

La PABX realiza todas las funciones normales: supervisión de líneas, detección de las cifras marcadas por teléfonos que utilicen la marcación normal, enviar llamada a los teléfonos y establecimiento de las

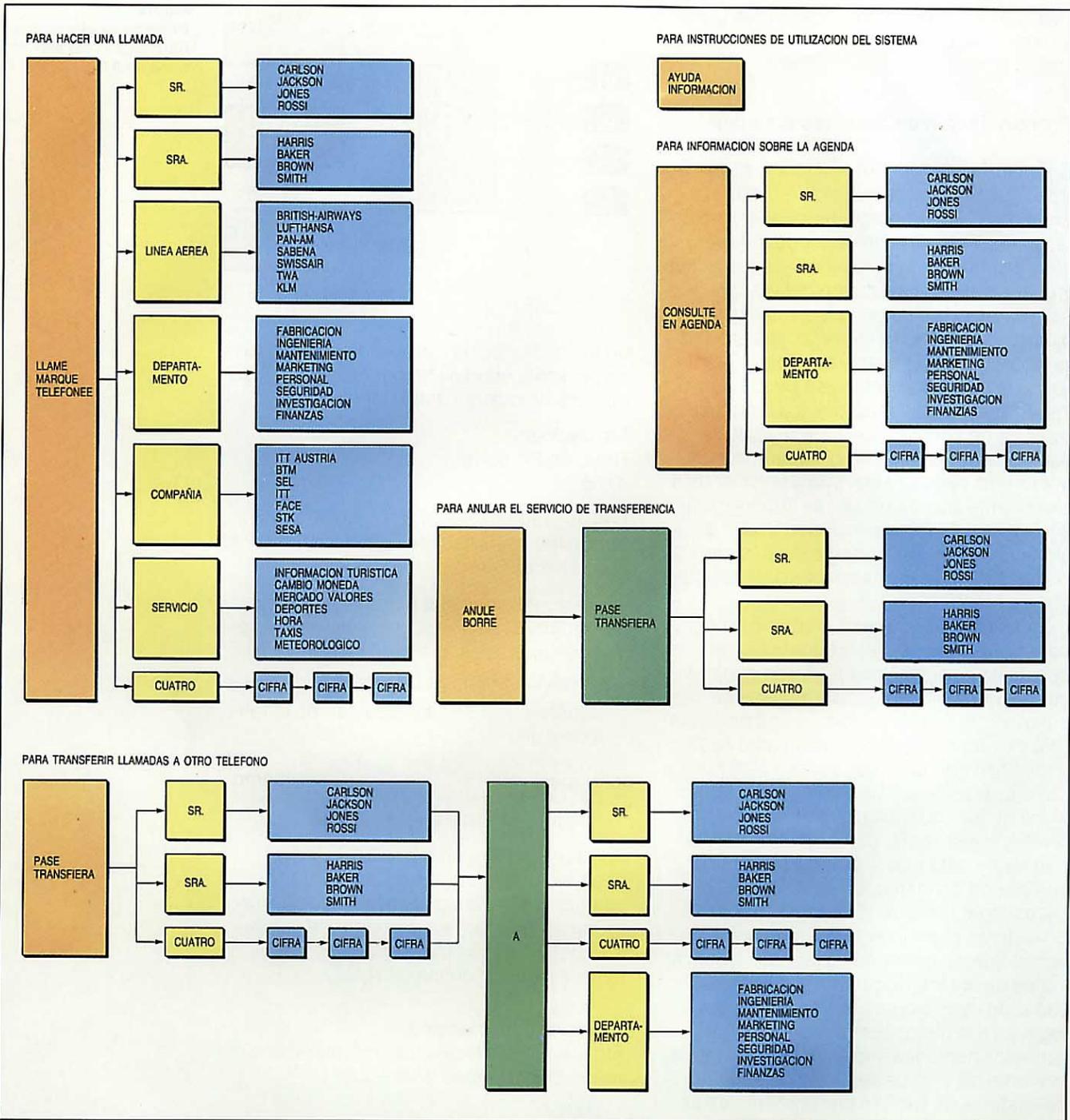


Figura 2
Vocabulario y sintaxis del sistema de reconocimiento de voz de ATC. Se han traducido del inglés las órdenes verbales originales para facilitar la comprensión al lector.

conexiones de voz. Además, la PABX se comunica con el elemento de control del sistema para proporcionar información sobre el estado y recibir las órdenes de control.

La unidad de síntesis de voz (procesador de análisis-síntesis de voz desarrollado por Standard Elektrik Lorenz, asociada alemana a ITT) genera mensajes sintetizados por LPC (codificación por predicción lineal), dirigida por el elemento de control del sistema. Esta unidad genera también los coeficientes LPC durante el desarrollo del sistema.

La unidad de reconocimiento, un CSR (reconocedor de voz continua) desarrollado por la División de Comunicaciones para la Defensa de ITT, detecta entradas orales y empareja sus plantillas al subconjunto de vocabulario especificado por el elemento de control.

El elemento de control del sistema realiza la coordinación de los otros subsistemas y todas las funciones de toma de decisión para la transacción con el usuario. Esta unidad analiza las palabras reconocidas basándose en reglas sintácticas, determina la confianza en la orden pronunciada y

decide la acción a ejecutar. Después crea un mensaje de salida apropiado a partir de una biblioteca de palabras y frases, y dirige el sintetizador de voz de acuerdo con ello. La agenda de nombres y números reside también en el elemento de control.

En la figura 2 aparece el vocabulario completo del sistema, que consta de 68 palabras.

Sistema de B&CC

Hasta ahora, las experiencias con sistemas de reconocimiento y síntesis de voz unidos al equipo de las PABX sólo han respondido a algunas de las preguntas sobre factores humanos. Se necesita probar continuamente en entornos de oficina reales para completar la parte ergonómica de estas tecnologías y asegurar así su éxito. Las pruebas de campo han de reunir los datos precisos para el desarrollo de un sistema acabado, ya que guiarse únicamente por la intuición o la percepción de los diseñadores podría ser desastroso.

Prueba de campo I sobre factores humanos (PCFH-I)

Se designó con tal nombre a la realización y prueba en Raleigh del sistema de B&CC (Business and Consumer Communications). La finalidad primordial de la PCFH-I fue recoger información acerca de los factores humanos que implica el uso de tecnología de la voz en las PABX. En la figura 3 se presentan los tres elementos principales de la experiencia:

- PABX Sistema 3100 de ITT
- multibus CSR
- placa de avisos con voz sintetizada desarrollada por B&CC.

Descripción del sistema

El multibus CSR está conectado al Sistema 3100 mediante una placa interfaz que extrae la voz digitalizada del bus del sistema de transmisión MIC y pasa los datos al CSR, donde es reconocida.

Los resultados de este reconocimiento se devuelven a través de un enlace RS232 a la placa interfaz, donde se traducen a dígitos, más cinco caracteres especiales, y se envían al controlador del Sistema 3100 por una conexión en paralelo con el receptor de teclado (Tel-Touch*) del sistema.

La programación de control del Sistema 3100 se modificó para permitir cambios en las rutinas de progresión de las llamadas, entre los que se incluye la interrupción de

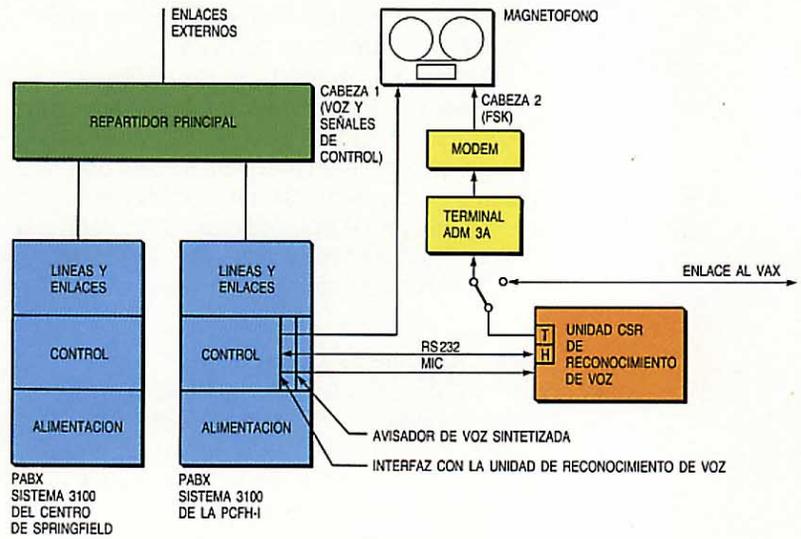


Figura 3
Soporte físico de las pruebas de campo sobre factores humanos según el sistema de B&CC.

los tonos normalmente utilizados (de ocupación, confirmación, etc.) con el fin de poder dar mensajes orales sintetizados al abonado. Todo esto se consiguió mediante órdenes del controlador del sistema a la placa de avisos con voz sintetizada.

El avisador es un sintetizador de voz de cuatro canales, compatible con el Sistema 3100, con memoria capaz de almacenar unas 180 palabras, aunque en la prueba sólo se utilizaron alrededor de 60, escogidas especialmente de manera que pudieran formar mensajes relacionados con el avance de la llamada.

Prueba de campo

La prueba de campo sobre factores humanos se llevó a cabo en Raleigh, en la oficina Springfield de B&CC. Para ella se seleccionaron veinte participantes, basándose en tamaño del grupo, configuraciones de llamada y acentos, y sus teléfonos fueron desconectados del sistema principal de comunicación del edificio (Sistema 3100), conectándolos al sistema de prueba; a su vez, éste actuó como intercomunicador conectado al Sistema 3100. Cada abonado disponía de una línea entrante directa "no activada por voz", con objeto de que no perdieran llamadas importantes en caso de fallo del sistema experimental.

En la prueba de campo, el reconocedor de voz utilizó plantillas de palabras aisladas no sensibles al locutor. Al pronunciar el abonado una palabra, el sistema respondía con un pitido para espaciar las órdenes habladas.

En la prueba se incluyeron la marcación oral y la activación de facilidades por órdenes habladas. El sistema permitía a cada uno de los veinte abonados llamar oralmente a los demás, así como a una lista específica de 17 nombres comunes o luga-

* Marca registrada del Sistema ITT

res exteriores, teniendo que marcar cualquier otro número de forma manual.

El sistema comprobaba las órdenes habladas de los abonados y, si se producían errores sintácticos, el controlador del sistema generaba un mensaje sintetizado que pedía al abonado: "Repita, por favor."

En la figura 4 aparecen el vocabulario y la sintaxis utilizados en la prueba, así como las facilidades realizadas en el sistema.

Recogida de datos

El estudio del B & CC se ha basado en el uso del sistema dentro de un ambiente de oficina real. Por haberse terminado la prueba muy recientemente, no se ha completado el análisis de todos los datos.

Para la prueba se instaló un magnetófono que grababa toda petición verbal de un abonado para acceder al sistema, mientras

que un segundo canal grababa la salida del reconocedor de voz, lo que permitió multitud de observaciones interesantes. Además de recoger ruidos típicos de oficina, como máquinas de escribir, conversaciones, etc., se pudo analizar la transacción oral que realmente ocurría, comprobando la precisión del reconocedor de voz.

Resultados iniciales

A lo largo de la prueba de campo, los 20 participantes completaron dos cuestionarios, cuyos resultados fueron tabulados. Hasta ahora sólo se ha hecho un análisis superficial de los datos, si bien el resultado de los cuestionarios ha aclarado algunos puntos.

En primer lugar, la mayoría de los abonados encontraron inadecuado, y en muchos casos intolerable, el espaciamiento por pitidos; hacerlo mediante confirmación inmediata de la palabra pronunciada será igualmente inaceptable, puesto que tiende a hacer más lenta la transacción. Nuevos ensayos pueden demostrar que sólo se aceptará la confirmación de frases completas, prefiriéndose que éstas sean lo más cortas posibles (p. ej., "conecte al dos-siete-cinco"). Este breve mensaje asegura al usuario que se le está conectando con el destino deseado, y le da la oportunidad de colgar si el mensaje indicara que se está llamando a un número equivocado.

El ochenta por ciento de los usuarios pidió la facilidad de programar uno o más números personales para la marcación oral; esto indica claramente que el sistema tendrá mayor utilidad si los usuarios pueden programar ellos mismos una lista personal de los números a los que llaman más a menudo. La longitud de tal lista puede ser variable, aunque todos los usuarios agradecerían, al menos, una posibilidad de marcación personal.

Varios abonados comentaron que el uso de sus aparatos de teclado multifunción para marcar y acceder a servicios programados era más rápido que la utilización de facilidades orales. Por consiguiente, la dotación de tales facilidades para un abonado particular deberá ser programable, a modo de "clase de servicio", puesto que no todos los abonados tendrán el mismo tipo de aparato. Los que utilicen aparatos normales de teclado o disco podrán acceder a las mismas facilidades del sistema que los equipados con teléfonos multifunción. En este aspecto es donde el reconocimiento de voz podrá elevar más la productividad.

El uso de la voz masculina para mensajes sintetizados fue apoyado por la mitad de los participantes y rechazado sólo por un 10% de ellos, mientras que al 40% restante o "no les importaba", o preferían ligeramente

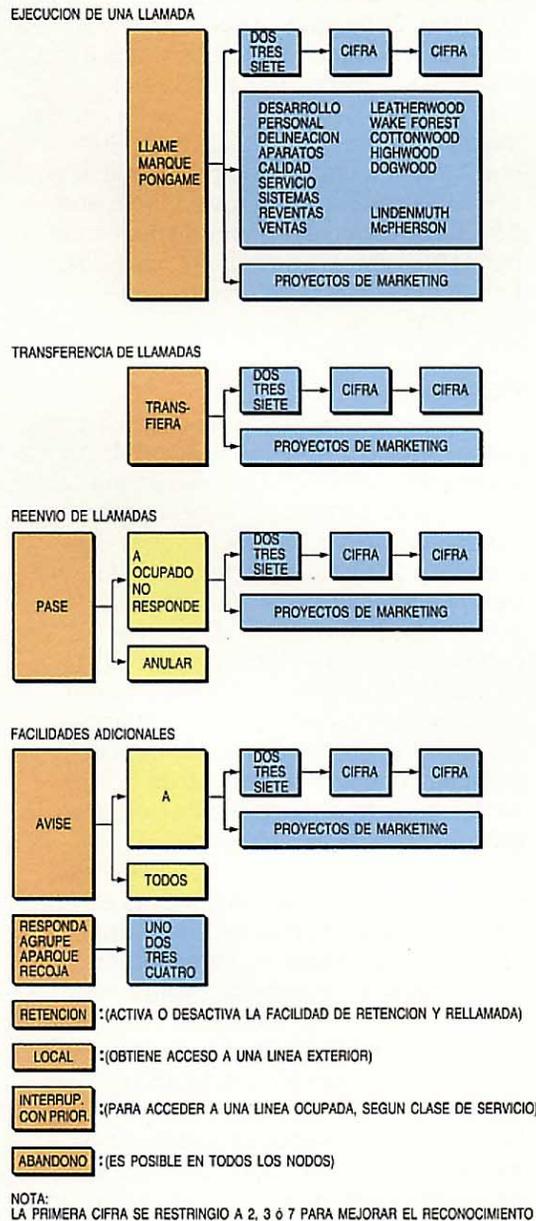


Figura 4
Vocabulario y sintaxis del sistema de B & CC para pruebas de campo sobre factores humanos (se han traducido las palabras inglesas para una mejor comprensión).

la voz masculina. Sin embargo, con la tecnología actual, la elección entre voz masculina o femenina para la confirmación puede ser programable por "clase de servicio". El 40% de los abonados recomendó acortar los mensajes, y el 60% restante no respondió a la pregunta sobre cambio de los mensajes. La confirmación a los usuarios tendrá que ser inteligible. La longitud de mensaje se mantendrá en un número mínimo de palabras (2 a 5), ya que frases más largas pueden causar irritación.

Aplicaciones

Las dos aplicaciones más prácticas de la PABX, para una experiencia inicial, son la marcación oral y los servicios activados por la voz. Estos dos usos del tratamiento de voz en la PABX constituyen una primera etapa necesaria en la evolución hacia otros nuevos servicios, y proporcionan también una base para explorar los factores humanos con el fin de realizar tales servicios de un modo agradable al usuario. A continuación se señalan las cuatro aplicaciones de reconocimiento y síntesis de voz que más probablemente se establecerán en las PABX.

Marcación oral

Permite a cualquier teléfono del sistema establecer una llamada sin más que pronunciar las cifras, nombres, lugares o departamentos deseados, tal como estén especificados en el vocabulario incorporado. Se espera que la marcación oral progrese hasta incluir listas de números personales, dependientes del locutor, aumentando con ello su utilidad. El vocabulario en la realización básica debe incluir cifras independientes del locutor, nombres de departamentos normales y varios "lugares comunes" a los que llamen la mayoría de los usuarios.

Facilidades activadas por la voz

Durante el progreso de una llamada, los usuarios podrán activar facilidades tales como transferencia, llamada tripartita, retención y espera en línea ocupada, sin más que pulsar una tecla o el gancho conmutador y nombrar luego la facilidad deseada. Hasta ahora, para lograr una conexión, los usuarios debían recordar o consultar el código de acceso, y después marcarlo. Dado que el usuario conoce el nombre de la facilidad que quiere utilizar, le será más cómodo y agradable acceder oralmente a ella. Al ir aumentando el número de facilidades, el usuario podrá utilizarlas todas sin estar limitado por las teclas de acceso que existan en su aparato.

Mantenimiento del sistema activado por la voz

Lo mismo en las centrales que en las PABX, las pruebas de diagnósticos (localización de fallos) se realizan normalmente *in situ*, por un ingeniero de mantenimiento que inicia secuencias automáticas de prueba, bien utilizando conmutadores o bien un dispositivo de entrada/salida especialmente diseñado.

El mantenimiento del sistema por órdenes orales se podría iniciar a distancia por medio de un teléfono, a través de una línea equipada con reconocimiento de voz que accedería directamente al equipo de mantenimiento rutinario del sistema. Para garantizar la seguridad podría interponerse un identificador de voz, ya que normalmente el mantenimiento es labor de un grupo reducido de ingenieros especialmente entrenados.

Una vez obtenido el acceso, el personal adecuado iniciaría las rutinas de mantenimiento para probar los circuitos específicos, pudiéndose emplear la síntesis de voz para dar los resultados de la prueba al que llama.

Central de mensajes telefónicos

En este tipo de aplicación, la voz es detectada, codificada, identificada y almacenada en formato ASCII (código normalizado americano para intercambio de información), pudiendo luego recuperarse en forma de voz sintetizada, ser entregada a un terminal a modo de carta, o a una impresora de líneas para su distribución.

Se está investigando esta facilidad para futuros sistemas de comunicaciones, en los que incluiría una biblioteca de usuario de 5.000 a 15.000 palabras, lo cual implica fuertes exigencias en memoria y capacidad de proceso.

La central de mensajes telefónicos necesitará importantes avances, tanto en el soporte físico como en el lógico, para ser económicamente asequible entre 1990 y 1995.

Conclusiones

El reconocimiento de voz, utilizado inicialmente como alternativa a la marcación manual, no merece la pena a menos que se realicen fiablemente extensos vocabularios. Para explotar plenamente la tecnología, el usuario debe poder acceder por diálogo oral a todas las facilidades de la PABX. Cuando sirva como medio de acceso a servicios, lo cual requiere vocabularios menores, el reconocimiento de voz será claramente ventajoso.

Los mensajes sintetizados deben de sonar naturales y dar suficiente información al usuario ocasional, aunque no deben ser tan prolijos que hagan impacientarse al usuario habitual.

La precisión actual del reconocimiento de voz es buena, siempre que el usuario conozca las reglas básicas establecidas, tales como el vocabulario, la sintaxis y la necesidad de pronunciar cada palabra separada y distintamente. La mayoría de los errores en el reconocimiento provienen de órdenes pronunciadas incorrectamente. En lugar de pedir mayor colaboración al usuario, lo cual se opondría al objetivo primario de hacer agradable el uso del sistema, es preferible aumentar las posibilidades del sistema de reconocimiento para que admita una sintaxis más flexible en las órdenes habladas.

Los resultados obtenidos con el equipo experimental de ITT indican que es viable

integrar el tratamiento de voz en una PABX, y así se llevará a cabo en sistemas futuros, sin que los diseñadores desconozcan los obstáculos a superar para poder conseguir tales PABX.

Hoshang Mulla recibió los grados de BSEE, MSEE y PhD en la Universidad de Michigan. Desde 1977 hasta 1982 estuvo en Motorola Inc, donde trabajó en proceso digital de señal y de voz. En 1982, el Dr. Mulla se incorporó al Advanced Technology Center de ITT, y allí actualmente se dedica a proceso de voz y a la aplicación del proceso digital de señales a sistemas de telecomunicación.

John F. Vaughan nació en 1941. Entró en la División CSD de ITT Kellogg en 1963 para trabajar en comunicaciones militares, y luego fue transferido a la división comercial de la compañía en 1965. Desde 1968, ha participado en diversos proyectos sobre PABX. La posición actual del Sr. Vaughan como miembro principal del staff técnico, es la de director del proyecto de investigación en la aplicación de B&CC de síntesis y reconocimiento de voz en una PABX.

Marcador telefónico activado por la voz

Un prototipo de marcador telefónico oral, basado en refinadas técnicas de reconocimiento de la voz, proporciona a los abonados un interfaz agradable con el sistema telefónico. El marcador reconoce nombres o dígitos pronunciados y marca luego, automáticamente, la línea requerida.

M. Immendörfer

Centro de Investigación de
Standard Elektrik Lorenz AG, Stuttgart,
República Federal de Alemania

Introducción

El acceso a los sistemas de conmutación telefónica se hacía en un principio por comunicación oral entre personas. El usuario descolgaba su teléfono y solicitaba a la operadora la persona con la que deseaba hablar, tras lo cual la telefonista encaminaba la llamada hacia la línea requerida. Este procedimiento natural se ha sustituido por sistemas de conmutación automática que permiten a un abonado establecer conexión con otro marcando un número, posiblemente compuesto de muchas cifras.

Sin embargo, la gente recuerda más fácilmente nombres que números, por lo

cual podría tener muchas ventajas un sistema que permitiera establecer llamadas por comunicación oral, como en los primeros años del siglo. Recientes avances en la investigación sobre el reconocimiento de voz y la microelectrónica han puesto al alcance las tecnologías necesarias para el desarrollo de sistemas de marcación oral, capaces de comprender órdenes pronunciadas de forma natural.

En la práctica, el equipo de reconocimiento de voz se puede localizar bien en la central o bien en el teléfono del abonado, según sea independiente o dependiente del locutor. Los equipos para estos dos sistemas básicos se diferencian bastante en lo que respecta al tamaño del vocabulario, la generación de patrones de referencia y el coste.

El Centro de Investigación de SEL ha diseñado y desarrollado un prototipo de marcador telefónico oral aplicable al usuario final; dicho marcador transforma el nombre del abonado llamado, claramente pronunciado ante el microteléfono, en la marcación automática del número de dicho abonado. Varias características agradables para el usuario aseguran que el marcador oral ofrece un modo fácil de establecer una conexión telefónica.

Requisitos de un marcador controlado por voz

Se han considerado dos tipos posibles de entrada para marcar mediante el reconocimiento de voz en un terminal de usuario final:

- entrada por orden verbal de los nombres de abonado que han de traducirse a números de llamada almacenados en una agenda telefónica electrónica
- entrada verbal de los números de teléfono, sustituyendo a la marcación normal por teclado o disco.

Prototipo del marcador telefónico oral desarrollado por SEL.



La principal ventaja de entrar pronunciando el nombre del abonado es que se tiene acceso directo a un número de abonado completo, mientras que la entrada verbal de cifras sigue exigiendo el recordar los números, y ello no mejora básicamente el servicio de marcación telefónica. Un buen sistema de marcación oral, agradable al usuario, sería el que permitiera la entrada tanto por nombres como por números.

Para dar una idea de los problemas y posibilidades del equipo de reconocimiento de voz, se resumen a continuación los requisitos básicos de un sistema práctico de marcación oral, en el área del usuario final:

- Los usuarios deberán poder definir el vocabulario de aplicación (esto es, generar las plantillas vocales individuales de abonado a las que se pueda acceder mediante entrada oral).
- El tamaño del vocabulario será de 30 a 50 nombres por usuario.
- El mismo terminal podrá ser utilizado por distintos usuarios; el tamaño del vocabulario crecerá en proporción al número de usuarios potenciales.
- Será también posible marcar por disco o teclado normal.

Un examen de la tecnología actual sobre reconocimiento de la voz demuestra que sólo un sistema dependiente del locutor puede ofrecer características tales como la generación, por parte del usuario, de un vocabulario de aplicación individual o el procesamiento de cien o más expresiones orales en tiempo real. Los dispositivos hoy existentes para reconocimiento de voz se basan en algoritmos que comparan la expresión desconocida con plantillas de referencia almacenadas, y seleccionan entre ellas la que más se le asemeja.

La generación de plantillas de referencia independientes del locutor implica la reco-



Figura 1 Interconexión de un marcador telefónico oral con el aparato telefónico del usuario y la red telefónica.

gida y procesamiento de una gran base de datos con expresiones pronunciadas por distintos locutores, para lo cual es preciso recurrir al fabricante del equipo. Además de las restricciones en la generación de plantillas, el número de expresiones que puede procesar un sistema independiente del locutor es un orden de magnitud menor que el correspondiente a los dispositivos que dependen del locutor. De ahí que el vocabulario de un sistema desvinculado del locutor tenga que ser demasiado limitado (diez a veinte palabras) para la marcación telefónica oral.

En consecuencia, se han desarrollado circuitos para el reconocimiento, dependiente del locutor, de palabras aisladas (separadas por pausas). El algoritmo de reconocimiento se basa en un conjunto de características digitales de los parámetros de codificación por predicción lineal, y en la comparación de patrones mediante un entramado dinámico del tiempo*. Los circuitos utilizados en el marcador oral para reconocimiento son capaces de procesar más de cien expresiones vocales en tiempo real.

Concepto del sistema

La figura 1 muestra cómo se incorpora el marcador oral a la red telefónica. Este dispositivo se puede añadir a teléfonos normales, por simple conexión a los hilos a y b de la línea. Se utiliza tanto en redes telefónicas públicas como privadas.

* Se denomina así un algoritmo que admite variaciones en el tiempo de los diferentes sonidos (fonemas, diptongos) que componen una palabra. Esto es necesario porque nunca se pronuncia dos veces la misma palabra exactamente del mismo modo.

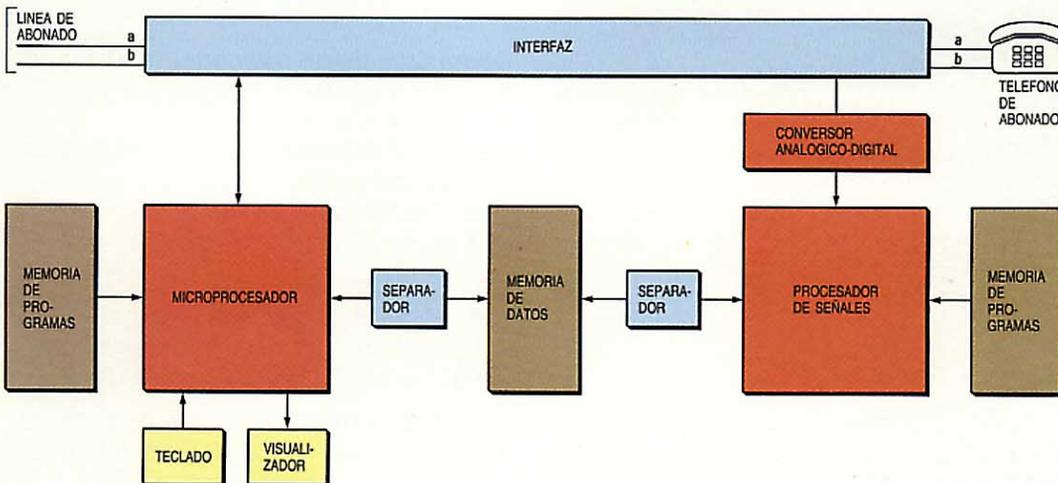


Figura 2 Diagrama de bloques del marcador oral desarrollado por SEL.

La configuración básica del marcador oral (Fig. 2) consiste en una unidad de reconocimiento de voz, una unidad de control y una memoria de datos para la agenda electrónica de teléfonos y el almacenamiento de las plantillas de referencia. El reconocimiento de la voz se realiza esencialmente por un circuito integrado procesador rápido de señales; un microprocesador estándar de 16 bits ejecuta partes del algoritmo menos críticas en tiempo, que necesitan una lógica de decisión compleja. El microprocesador se utiliza sobre todo para el control global del sistema, incluyendo la supervisión del estado de la línea telefónica, el tratamiento de la entrada por teclado y la generación de la salida al visualizador. Se puede acceder a la memoria de datos tanto por el microprocesador como por el procesador de señales. El programa para el procesador de señales, que procesa la parte principal del algoritmo de reconocimiento de voz, y el programa para el microprocesador se localizan en dos memorias programables de sólo lectura (PROM) separadas.

El teclado alfanumérico permite al usuario introducir en la memoria de la agenda los nombres de los abonados y sus correspondientes números de teléfono. Cuando se utiliza la marcación oral, el visualizador ayuda al usuario, confirmando su entrada verbal al presentar el nombre y su correspondiente número de teléfono o, en el caso de marcación verbal del número, los dígitos reconocidos.

El marcador oral se conecta a la línea telefónica por medio de un interfaz cuyas principales funciones son las de conmutar el micrófono del teléfono a la unidad de reconocimiento de voz cuando el abonado ha escogido la marcación verbal, y transmitir las señales de marcación a la línea telefónica una vez reconocida la señal de voz. Durante la fase de conversación, el micrófono vuelve a conmutarse a la línea telefónica, siendo ésta observada constantemente por el interfaz y la unidad de control del microprocesador.

En la figura 3 se puede apreciar la división de la memoria de datos en dos partes: una para almacenar las plantillas acústicas de referencia, y otra para la información de la agenda electrónica de teléfonos. Una vez que la unidad de reconocimiento de voz ha asignado la entrada oral a una plantilla de referencia válida, aparece en pantalla el nombre del abonado reconocido junto con su número de teléfono, el cual será utilizado por el microprocesador para llamar a dicho abonado.

El prototipo entero del marcador oral se aloja en una caja, sobre la cual puede colocarse el teléfono. Además del teclado alfanumérico normal, se equipan varias teclas

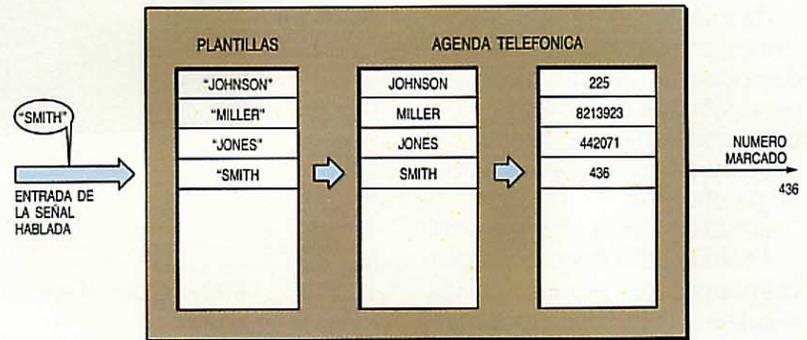
para funciones tales como la activación y el adiestramiento del sistema de reconocimiento de voz. El visualizador de cristal líquido tiene dos líneas de 32 caracteres cada una. Una fuente de alimentación local separada proporciona los 10 W que necesita el marcador.

Características del marcador oral

Marcando mediante la pronunciación del nombre de un abonado, se elimina la necesidad de buscar su número en una guía telefónica y teclearlo después. El usuario almacena la guía telefónica en la memoria del sistema de marcación oral, y la activa con su propia voz.

Se han preparado tres diferentes modali-

Figura 3
Organización de la memoria de datos que almacena las plantillas de reconocimiento de voz y la agenda de teléfonos.



dades de marcación en el prototipo, a elección del usuario:

- activada por la voz
- por teclado o por disco
- por repertorio telefónico, característica adicional no controlada por voz.

Una vez descolgado el microteléfono y activada la unidad de reconocimiento al pulsar la tecla apropiada, la señal de voz de entrada se reconoce y traduce a un código secuencial de marcación. Dicha señal de entrada consiste en expresiones verbales (nombres de personas, compañías, ciudades) que pueden ser definidas individualmente por cada usuario, o en la pronunciación de las cifras que componen el número deseado. En la memoria de datos del marcador se pueden almacenar hasta 128 nombres de abonados y sus correspondientes números de teléfono (máximo de 16 cifras). Diferentes usuarios pueden utilizar el mismo dispositivo si éste ha sido previamente adiestrado para reconocer su vocabulario. El usuario no necesita identificarse, ya que el sistema explora la memoria entera de plantillas de referencia buscando la locución pronunciada.

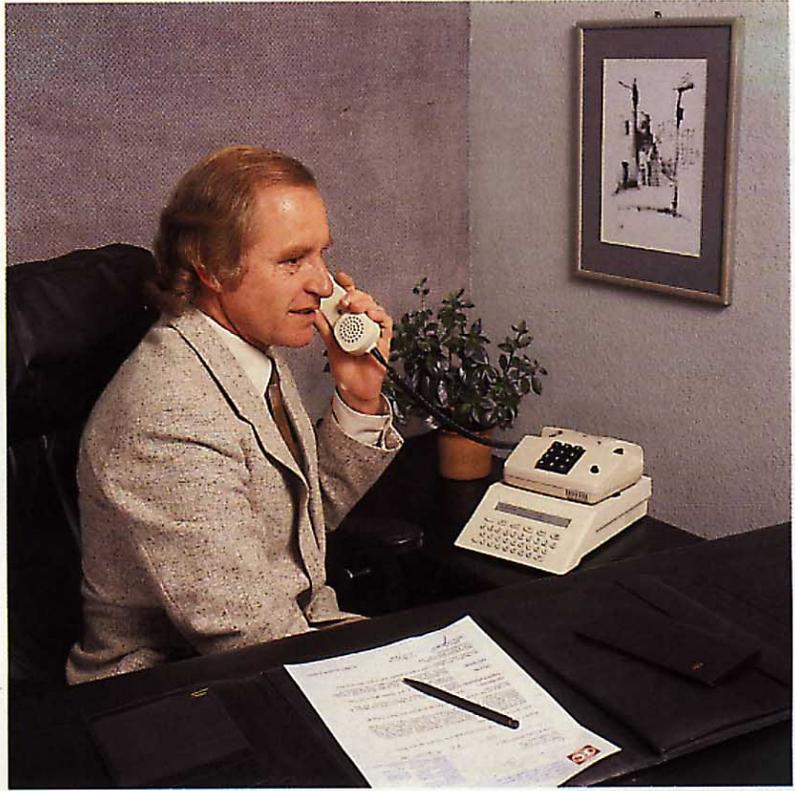
En el caso de secuencias numéricas de llamada largas y complejas, es posible una entrada oral combinada de nombres y dígitos. Por ejemplo, se podría acceder a una compañía pronunciando un nombre, al cual se añadirían las cifras dadas verbalmente del abonado específico deseado. Sin embargo, si a este abonado se le llama con frecuencia, convendría más generar una sola orden para el número entero.

Se han incluido otras posibilidades controladas por la voz en el programa de aplicación para que el usuario tenga un acceso agradable al teléfono de marcación oral. Así, por ejemplo, se puede volver a llamar al último número, pronunciando sólo la orden de rellamada. Además, si se conecta el equipo marcador oral a una centralita, se pueden activar verbalmente ciertas facilidades de la misma que normalmente sólo responden a códigos especiales (p. ej., transferencia de llamadas).

Otra manera posible de funcionar es la de marcar con el disco o teclado convencional, del modo habitual. Esto puede reservarse para números no incluidos en la agenda, o para casos en que el teléfono sea utilizado por alguien no familiarizado con el sistema de reconocimiento o cuya voz no esté todavía registrada en el sistema.

Además de la marcación controlada oralmente o por el teclado estándar, puede utilizarse el prototipo como marcador de repertorio telefónico. En este caso, se puede efectuar la búsqueda de un nombre de abonado en la agenda electrónica del teléfono tecleando las letras iniciales; dichos nombres están almacenados por orden alfabético en la agenda, de modo que en el visualizador aparece el nombre más próximo a las letras tecleadas. Partiendo de este punto, puede hallarse el nombre deseado muy rápidamente, avanzando o retrocediendo nombres. A continuación, se inicia la llamada pulsando un botón en el teclado. La marcación por repertorio es especialmente útil si el usuario está acatarrado o afónico y, por lo tanto, no puede esperar resultados satisfactorios del sistema de reconocimiento de voz.

Para poder trabajar con los diferentes sistemas de conmutación telefónica, el prototipo del marcador oral contiene un generador de impulsos, además de un generador de multifrecuencia, pudiendo así conectarse tanto a un receptor de impulsos decádicos como a un receptor de multifrecuencia. En el caso de llamadas de salida vía centralita, una unidad especial de detección del tono de marcar hace que se tenga en cuenta la señalización del enlace. Si fallara la potencia, la batería de reserva evitaría la pérdida de datos en la agenda telefónica.



Marcador oral utilizado en un ambiente de oficina típico.

Manejo del sistema y adiestramiento acústico

El contenido de la agenda electrónica de teléfonos se puede presentar por orden alfabético en el visualizador, utilizando las teclas de función. La búsqueda de determinados nombres de abonado por sus letras iniciales simplifica el acceso a la agenda, especialmente en el caso de vocabularios extensos, cuando se marca mediante repertorio.

Se pueden añadir fácilmente nombres de abonados a la agenda, siendo asimismo muy sencillo el adiestramiento acústico para generar nuevas plantillas de referencia. En primer lugar, se teclean los nombres de los abonados y sus números de teléfono correspondientes. Seguidamente, el sistema solicita al usuario que le enseñe a reconocer los nombres introducidos. Para cada locución emitida, la unidad de reconocimiento de voz genera plantillas específicas del usuario, que sirven de referencia en el transcurso de la marcación oral.

De forma similar, se utiliza el teclado para efectuar modificaciones en la agenda (p. ej., suprimir un abonado o cambiar algún número de teléfono). Cuando se utiliza por primera vez el marcador oral, se necesita un adiestramiento inicial del sistema, y en tal caso éste pide al usuario que pronuncie todos los dígitos y unas pocas órdenes de

control, tales como "rellamar" y "error". Estos patrones se podrán luego utilizar para marcar por cifras y controlar oralmente el sistema.

Se ha diseñado un diálogo hombre-máquina que facilita el adiestramiento acústico y simplifica la activación oral del sistema. Mediante el visualizador, el usuario conoce los estados más importantes del sistema y la operación que debe ejecutar. Durante la marcación oral, cuando el usuario ha activado la unidad de reconocimiento de voz, el sistema le pide que pronuncie ante el micrófono el nombre del abonado de destino. Aparece luego en el visualizador el nombre del abonado reconocido y el correspondiente número telefónico almacenado, realizándose la marcación inmediatamente. Si en lugar del nombre se pronuncian los números, son las cifras reconocidas las que se visualizan. La última cifra pronunciada se puede borrar mediante una orden verbal de error, lo cual permite modificar adecuadamente una secuencia numérica incorrecta. Los errores en la secuencia pueden proceder tanto de la unidad de reconocimiento como del usuario. Si el sistema no reconoce una pronunciación, pedirá al usuario que repita la entrada.

Si la misma expresión verbal fuera rechazada varias veces, podría ser necesario readiestrar el correspondiente patrón de referencia. Para ello se accionará una tecla especial que marcará el patrón malo en la memoria, reemplazándose luego éste por un nuevo patrón de entrada pronunciado por el usuario.

Para controlar el volumen de las señales de entrada verbales y la velocidad de la pronunciación, el visualizador avisa a los usuarios inexpertos si sobrepasan ciertos valores límite.

Conclusiones

Los ensayos iniciales demuestran que el marcador telefónico oral tiene un gran valor práctico, beneficiando sobre todo a personas que precisan llamar frecuentemente a un gran número de abonados diferentes. La tecnología del reconocimiento de voz ofrece las ventajas de la comunicación oral entre los seres humanos y el equipo telefónico. El prototipo de marcador desarrollado por SEL permite al usuario llamar a otro abonado con la simple pronunciación de su nombre, siendo este nombre reconocido por el sistema, que automáticamente marca el número correspondiente. Puede también servir para reemplazar el disco o el teclado convencional por una entrada oral de los dígitos.

El uso general de la tecnología moderna de reconocimiento de la voz, combinado con las técnicas más recientes de microprocesadores y procesadores de señales, sitúa este marcador oral a la cabeza de los dispositivos controlados por la voz que ofrecen ventajas reales al usuario. En un futuro próximo, el uso de nuevos circuitos integrados procesadores de señales, de coste económico y de una programación adecuada, harán posible la aplicación del reconocimiento de voz a la marcación oral, consiguiendo fabricar estos dispositivos con un coste comparable al de los marcadores de repertorio convencionales.

M. Immendörfer nació en Stuttgart, Alemania, en 1941. Recibió el título de Licenciado en Física de la Universidad Técnica de dicha ciudad en 1968, y el grado de Doctor en 1973, por su trabajo sobre la relajación acústica de polímeros. Ingresó, en 1973, en el Centro de Investigación de SEL como ingeniero de planificación de sistemas. Desde 1977, el Dr. Immendörfer dirige el desarrollo del proceso digital de la voz en dicho Centro de Investigación.

Codificación de señales de vídeo

Hay que codificar en digital las señales de vídeo y reducir la velocidad binaria para la transmisión por una futura red digital de servicios integrados de banda ancha.

Existen para ello varias técnicas, como la modulación por impulsos codificados, la modulación por impulsos codificados diferencial y la codificación por transformada. Se ha desarrollado una familia de codecs de vídeo que operan a la velocidad binaria de las vías de transmisión.

G. Bostelmann

P. Pirsch

Centro de Investigación de
Standard Elektrik Lorenz AG, Stuttgart,
República Federal de Alemania

Introducción

Los progresos tecnológicos en transmisión por fibra óptica e integración a muy grande escala conducen a servicios nuevos y mejorados que utilizan redes de banda ancha. Existen ya por todo el mundo redes que distribuyen emisiones de televisión y enlaces internacionales vía satélite para el intercambio de programas de televisión; nuevos servicios, como la videoconferencia y el videoteléfono, están en fase de introducción o de planificación. Su realización requiere técnicas de codificación que adapten la señal de la fuente a la velocidad de la línea, con el fin de conseguir una transmisión eficaz de estas señales por enlaces digitales.

La salida de un sistema de imágenes es normalmente una señal analógica, la cual ha de ser transformada por un codificador a valores discretos y después a símbolos de código digital, de tal forma que la imagen sea representada por el mínimo número posible de tales símbolos; para ello se ha de eliminar la información redundante e irrelevante. La redundante se puede reconstruir totalmente en el receptor a partir de las propiedades estadísticas de la señal. La información irrelevante no significa nada para el receptor, y cuando éste es una persona, su percepción visual es quien la determina. Por tanto, el diseño de un codec de imagen depende de la descripción estadística de la señal de origen y de las características del ojo humano.

Codificación de la fuente

La teoría de la información enseña que la salida de una fuente discreta se mide por el número de decisiones binarias. Si cada símbolo se selecciona entre K símbolos posibles, la entropía máxima H_0 de la fuente es $\log_2 K$, expresada en bitios. No se necesita más de $\log_2 K$ bitios por símbolo para representar la señal enteramente.

La información no redundante en cada símbolo depende de la probabilidad del mismo, y si ésta es alta la información será escasa; cuando la probabilidad es uno, el símbolo no lleva información y por tanto se puede estimar totalmente en el receptor, siendo innecesaria su transmisión. La información no redundante por símbolo $I(u)$ viene dada por $-\log_2 P(u)$ bitios para todos los símbolos u con probabilidad $P(u)$.

La información media de los K símbolos posibles, denominada entropía de la fuente¹, es calculable como valor esperado

$$H(U) = \sum_u P(u) I(u) = - \sum_u P(u) \log_2 P(u).$$

Una señal de la fuente codificada carece de redundancia si cada símbolo se codifica del modo óptimo respecto a su información no redundante y la velocidad de transmisión resultante esta dada por la entropía H . En el caso de las señales de vídeo, las muestras adyacentes tienen una influencia estadística esencial y hay que tener en cuenta la estadística condicional. Por tanto, la información de una fuente de imagen será la entropía condicional:

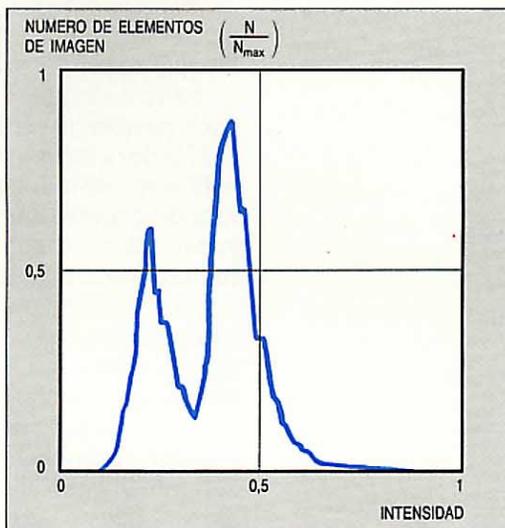
$$H(U_0|U_1 \dots U_N) = - \sum_{u_0} \dots \sum_{u_N} P(u_0, \dots, u_N) \log_2 P(u_0|u_1 \dots u_N)$$

siendo $P(u_0|u_1 \dots u_N)$ la probabilidad de que la muestra u_0 siga la secuencia $u_1 \dots u_N$.

La redundancia R , dada por la diferencia $H_0 - H(u_0|u_1 \dots u_n)$, indica cuántos bits por símbolo pueden ahorrarse con una codificación óptima, es decir, cuando cada muestra u_0 está codificada conforme a su contenido de información. Huffman ha publicado un algoritmo para construir códigos de mínima redundancia², en los que la longitud de palabra se aproxima al negativo del logaritmo de la probabilidad, y por tanto, la longitud media de la palabra de código se acerca a la entropía H .

Cuando las señales de vídeo se exploran y muestrean de acuerdo con las normas usuales en televisión, la función de probabilidad dista mucho de ser uniforme y depende de la escena real. La figura 1 muestra el ejemplo de una imagen de vídeo típica e indica que la entropía de las señales de vídeo será inferior al valor máximo. Las medidas de entropía de imágenes con $K = 256$ niveles de amplitud ($H_0 = 8$ bits) dan una entropía incondicional de 7,3 bits por pixel (elemento de la imagen). La entropía condicional, considerando únicamente el pixel contiguo a la izquierda, es de 4,6 bi-

Figura 1
Intensidades de
luminancia de una
imagen típica³.



tios por elemento, lo cual indica una fuerte correlación entre los pixels adyacentes.

Distorsión de señal

Además de reducir la redundancia, es posible eliminar los elementos de señal no significativos y, por tanto, decrementar aún más la velocidad de transmisión. La cota

inferior está marcada por la función distorsión dependiente de dicha velocidad⁴. Si el símbolo u de la fuente se cambia al símbolo v de salida del codificador, y la distorsión de u y v está dada por la función $d(u, v)$, entonces la distorsión media D será:

$$D = \sum_u \sum_v P(u) P(v|u) d(u, v).$$

Esta operación cambia la distribución de probabilidad $P(u)$ a $P(v)$, y la información a la salida del codificador relativa a la información de la fuente vendrá dada por la media mutua (fijada):

$$T(U, V) = H(U) - H(U|V)$$

en la que U y V son conjuntos de símbolos a la entrada y salida del codificador. La entropía condicional $H(U|V)$ describe la información del símbolo u de la fuente dado el símbolo v recibido, y dependerá de la distorsión introducida por el codificador y por el canal de transmisión.

La función distorsión/velocidad es la cota inferior de la información mutua:

$$R(D)^* = \min T(U, V)$$

donde D^* es el valor máximo de todas las distorsiones medias D .

Si se conoce la función de distorsión $d(u, v)$, se puede determinar la velocidad binaria de un codificador de fuente óptimo mediante la función distorsión/velocidad. No obstante, queda todavía por encontrar una función de distorsión consistente con las propiedades visuales del ojo humano. Incluso estando determinada la función distorsión/velocidad, ella no da por sí misma la codificación óptima de la imagen.

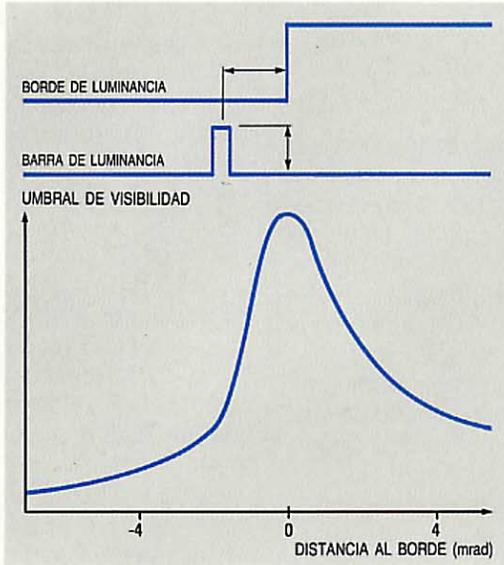
Propiedades de la visión humana

Sólo se consideran aquí algunas propiedades básicas de la visión que afectan a la codificación de las imágenes. En la referencia 5 se da descripción detallada.

La visión humana del color se puede modelar combinando tres señales: el rojo R , el verde G y el azul B . Ello significa que cada color puede especificarse como mezcla de tres colores primarios. En la codificación de televisión, estas señales R , G y B se combinan linealmente para producir una señal Y de luminancia y dos señales de crominancia $CR = R - Y$ y $CB = B - Y$. El brillo del color está descrito por la señal de luminancia, mientras que el matiz y la saturación lo están por las de crominancia.

La sensibilidad a frecuencias espaciales es importante para la codificación de imágenes.

Figura 2
Umbral de visibilidad para una barra superpuesta en los alrededores de un borde de luminancia.



nes, y determina la frecuencia máxima que un observador es capaz de distinguir. El ángulo con que el ojo abarca la imagen representada es un parámetro esencial que indica el grado de detalle apreciable en dicha imagen, por lo cual las frecuencias espaciales se expresan en ciclos por grado. La máxima sensibilidad a retículas sinusoidales se da hacia los 3 ciclos por grado. El ojo trabaja como un filtro espacial paso bajo, con sensibilidad a la luminancia inapreciable a los 40 ciclos por grado, correspondientes a una frecuencia de vídeo de unos 5 MHz según la norma G de la televisión europea y a una distancia de visión de seis veces la altura de la imagen. La pérdida de sensibilidad a la crominancia ocurre alrededor de los 15 ciclos por grado.

La sensibilidad al contraste afecta también a la codificación. La mínima diferencia de luminancia apreciable ΔL en un área pequeña depende de la luminancia del contorno. La relación $\Delta L/L$ (fracción de Weber) resulta ser casi constante para un amplio margen de brillo

$$\frac{\Delta L}{L} = k \quad k = 0,01 \dots 0,02.$$

La mínima diferencia discernible (umbral de visibilidad) se determina para un fondo de luminancia constante, incrementándose en áreas sujetas a cambios espaciales y temporales de señal. Se produce asimismo un notable aumento del umbral de visibilidad

junto a los bordes de la imagen. La figura 2 representa el umbral de visibilidad en función de la distancia al borde, mostrando el efecto de máscara que es importante para la codificación en MIC diferencial, donde se cuantifican directamente las variaciones de amplitud de la señal.

Este ejemplo demuestra que los efectos visuales son demasiado complejos para ser representados por una función de distorsión analítica. Para usos prácticos, la calidad de imagen dada por sistemas de codificación no se valora con medidas analíticas de la distorsión, siendo recomendables las pruebas con observadores humanos para enjuiciar la calidad de la imagen, según convenios internacionales de calidad, degradación o escalas de comparación⁶.

Técnicas de codificación

Estas técnicas deberán proporcionar una presentación digital de las señales de vídeo con requisitos mínimos en cuanto a velocidad de transmisión y capacidad de almacenamiento. Para lograrlo hay dos métodos básicos; reducir la redundancia estadística, o cuantificar con el mínimo número de niveles compatible con la calidad de imagen especificada.

Modulación por impulsos codificados

En un sistema MIC básico de codificación de imágenes (Fig. 3), se muestrea y cuantifica la señal continua de vídeo, asignando luego a cada muestra cuantificada un grupo de bits de longitud constante, denominado palabra de código.

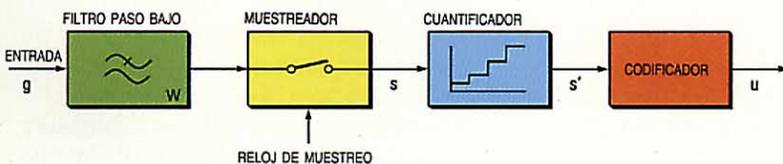
Siendo la imagen bidimensional por naturaleza, el muestreo ideal requiere una función espacial compuesta de un conjunto infinito de funciones delta de Dirac ordenadas en mallas con separación X, Y . En una secuencia de imágenes, como la televisión, cada muestra o pixel varía con el tiempo y por tanto tiene que ser muestreado en un intervalo de tiempo T . Ello da lugar a un modelo tridimensional de muestreo, cuyo proceso está descrito por el producto de la señal de imagen analógica $g(x, y, t)$ y el conjunto de deltas de Dirac. La imagen muestreada viene dada por:

$$s(x, y, t) = g(x, y, t) \text{ rep}_{x,y,t} \delta(x, y, t).$$

La notación "rep" indica la repetición de la delta de Dirac a lo largo de cada dimensión x, y, t con un espaciamiento X, Y, T . Después de aplicar la transformada de Fourier⁷, el espectro de la imagen muestreada será:

$$S(f_x, f_y, f_t) = \frac{1}{XYT} \text{ rep}_{\frac{1}{X}, \frac{1}{Y}, \frac{1}{T}} G(f_x, f_y, f_t).$$

Figura 3
Diagrama de bloques de un codificador MIC.



Esto significa que el espectro de la señal muestreada es una versión a escala del espectro de la imagen analógica $G(\cdot)$ repetido a intervalos de $1/X$, $1/Y$, $1/T$ en el espacio tridimensional de frecuencias f_x, f_y, f_t . Para evitar el solapamiento de los espectros repetidos se requiere la limitación de banda de $g(\cdot)$ de tal modo que:

$$W_x < \frac{1}{2X}; W_y < \frac{1}{2Y}; W_t < \frac{1}{2T}.$$

La imagen analógica se puede reconstruir a partir de la versión muestreada utilizando un filtro paso bajo tridimensional con bandas atenuadas de anchuras W_x, W_y, W_t .

En los sistemas prácticos, las muestras se obtienen por un método de barrido de la imagen. Una cámara explora continuamente las líneas de una secuencia de imágenes y produce a la salida una señal unidimensional g_{video} con un ancho de banda W_{video} . Esta señal se muestrea a intervalos al menos de $1/2 W_{\text{video}}$. Para señales de televisión según la norma G europea, W_{video} vale 5 MHz.

Una vez muestreado, cada pixel pasa a ser discreto al aplicarle un cuantificador, el cual transforma la variable analógica s en otra discreta $Q(s)$, generalmente mediante una función en escalera (Fig. 4) según reglas de cuantificación dadas por un conjunto de niveles de decisión $d_k (k=0, \dots, K)$ y niveles representativos $r_k (k=0, \dots, K-1)$. A cada amplitud s se le asigna un nivel r_k si $d_k \leq s < d_{k+1}$. Con el fin de conseguir el mínimo número de niveles de cuantificación K para una calidad de imagen deseada, se deben elegir para r_k y d_k valores óptimos.

Las condiciones analíticas para el diseño de un cuantificador se pueden deducir fácilmente con un criterio de error cuadrático medio mínimo³. Se ha comprobado que las medidas analíticas de error no coinciden suficientemente con los errores apreciados por el ojo humano. Se puede diseñar mejor un cuantificador teniendo en cuenta la visibilidad de las distorsiones. En el caso del MIC son críticos los pequeños cambios en la señal, cuya cuantificación grosera ocasionará contornos artificiales (Fig. 5). Por ello el método adoptado debería intentar mantener los contornos cuantificados por debajo del umbral de visibilidad.

Con los umbrales de visibilidad dados por la ley de Weber, los niveles de decisión y de representatividad no deberían espaciarse igualmente, lo cual puede lograrse mediante un preproceso casi logarítmico. En el caso particular de las señales de vídeo, esto ya lo hace la función de preproceso interno de la cámara con el fin de corregir la característica no lineal del tubo de rayos catódicos (la llamada precorrección gamma).

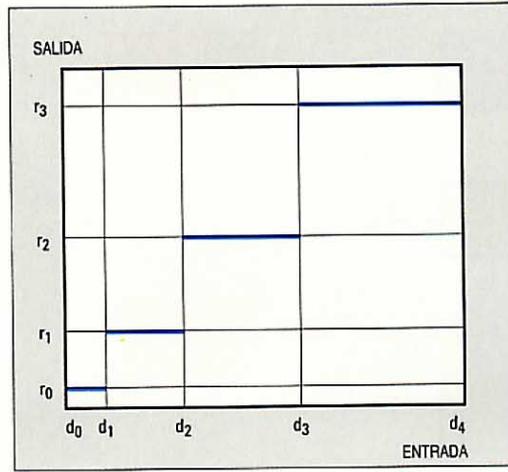


Figura 4 Principio operativo de un cuantificador, donde la transformación de una variable analógica en variable discreta produce una función en "escalera".

Las medidas de visibilidad del ruido de cuantificación del MIC han demostrado que se requieren 256 niveles (palabras de

Figura 5 Efecto de contorno por el uso de la cuantificación MIC grosera: cuantificación con 8 bits (arriba) y cuantificación con 3 bits (abajo).



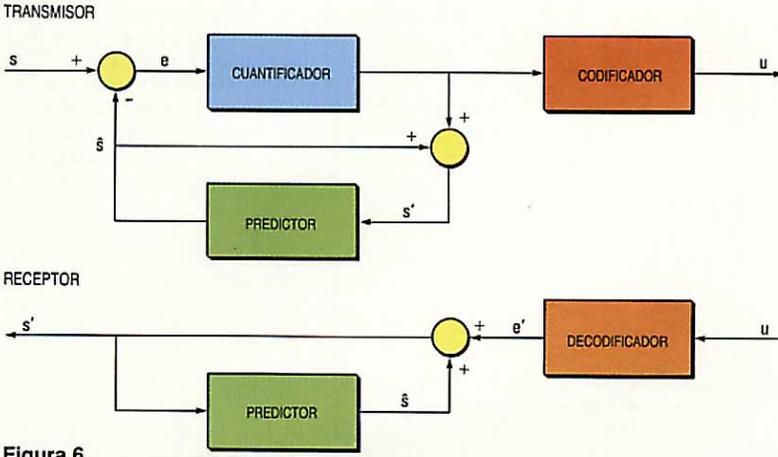


Figura 6 Diagrama de bloques de un sistema MICD.

Figura 7 Error de predicción de un sistema MICD que utiliza predictor bidimensional con todos los coeficientes positivos: imagen original (arriba) y magnitud del error de predicción (abajo).



código de 8 bits) en los componentes de la señal de vídeo, a fin de asegurar una alta calidad de la imagen. Utilizando las frecuencias de muestreo fijadas por el CCIR para estudios de televisión digital⁸, la codificación MIC básica exige una velocidad de transmisión de 216 Mbit s⁻¹ en las tres componentes de vídeo. Esta elevada velocidad puede reducirse mediante técnicas de codificación más refinadas que explotan las propiedades de la visión humana y el carácter estadístico de la señal de vídeo.

Modulación de impulsos codificados diferencial (MICD)

La relación estadística entre pixels contiguos y la menor visibilidad de la distorsión en la proximidad de los bordes han inducido a proponer una transmisión diferencial basada en métodos de predicción. La técnica que más se utiliza es la MICD³. En este sistema básico (Fig. 6), a partir de la información previamente transmitida, se hace una predicción \hat{s} de la muestra s que ha de codificarse. El error de predicción e obtenido por sustracción del valor predicho \hat{s} se cuantifica en un conjunto de niveles, representados por palabras de código de longitud constante o variable, los cuales se envían al codificador de canal para su transmisión. En el receptor, se reconstruye la muestra s , menos el error de cuantificación, añadiendo el error de predicción cuantificado e' al valor predicho \hat{s} . Para asegurar valores de predicción idénticos en el transmisor y en el receptor, la predicción en ambos lados se basa en las muestras s' reconstruidas.

En la literatura, el sistema MICD está mejor explicado sobre bases estadísticas³; aquí, sin embargo, lo examinamos desde supuestos psicofísicos. La predicción se comporta perfectamente para zonas de la imagen donde la señal cambia poco, pero los cambios bruscos (en luminancia o crominancia) no admiten una predicción tan buena. Por tanto, el error de predicción es un indicador de actividad, que se toma como medida de los cambios de señal. De acuerdo con el efecto de enmascaramiento, el umbral de visibilidad aumenta en las zonas de elevada actividad. Se permite, pues, una cuantificación no uniforme del error de predicción, aplicándose una cuantificación grosera a errores de predicción grandes. Debería diseñarse un predictor con estas directrices, para que el error de predicción diera una buena medida de la actividad. La figura 7 muestra que un predictor lineal bidimensional con todos los coeficientes positivos se comporta así para cambios espaciales de señal. El error de predicción es proporcional a la pendiente y a la diferencia de amplitud de los bordes.

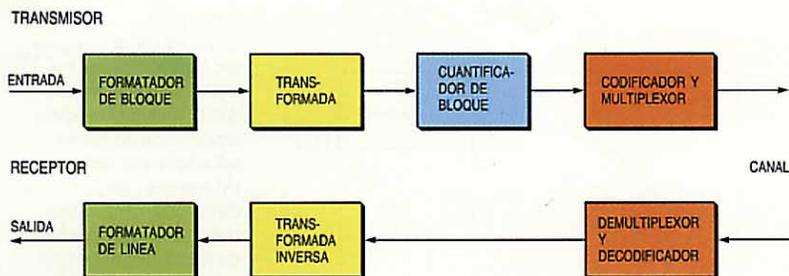


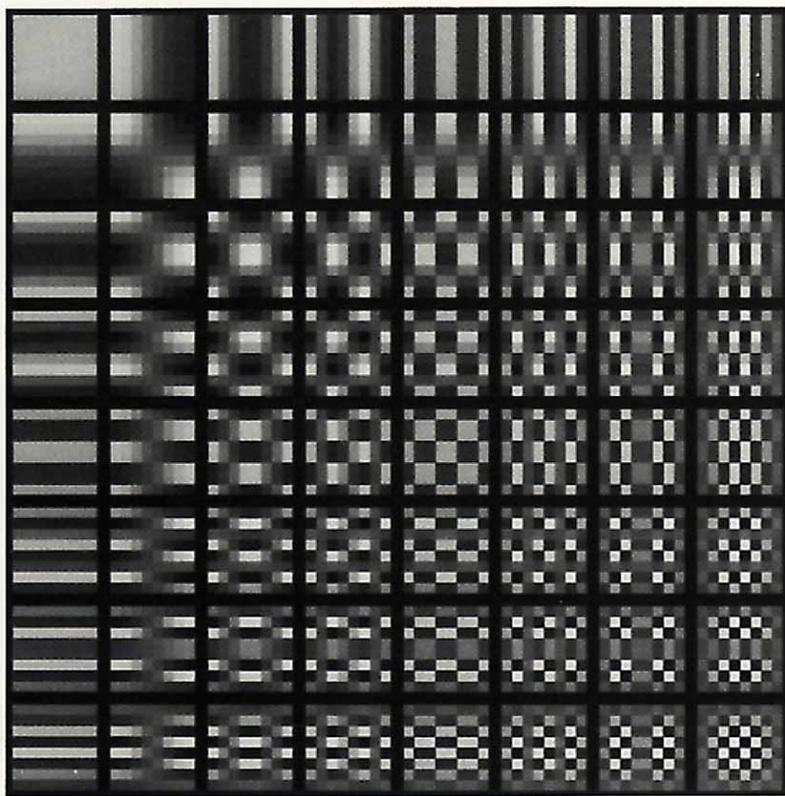
Figura 8
Diagrama de bloques de un sistema de codificación por transformada.

El diseño de un cuantificador con un número de niveles mínimo es fácilmente realizable sabiendo la relación entre el umbral de visibilidad del error de cuantificación y el error de predicción. Se ha medido esta relación efectuando pruebas subjetivas⁹ en las condiciones de visión especificadas por el CCIR⁶. Utilizando predictores fijos bidimensionales para imágenes naturales (no de prueba), se requieren al menos 14 niveles de luminancia y 8 en cada componente de crominancia para que los errores de cuantificación se mantengan bajo el umbral de visibilidad.

Los sistemas elementales MICD no adaptativos de este tipo pueden reducir la velocidad de transmisión en un factor de dos a tres con respecto al MIC. Los MICD más complejos, con predicción adaptativa, cuantificación adaptativa y palabras de código de longitud variable logran factores de reducción mucho mayores.

En una secuencia de televisión, la mayoría de los pixels permanecen prácticamente

Figura 9
Imagen básica de una transformación de coseno discreto bidimensional con bloques de dimensión 8×8 .



invariables de una trama a la siguiente, y en ello se basa la codificación intertrama, según la cual una imagen se divide en áreas variables de una trama a otra y en otras que no cambian. Para las partes que varían, es muy eficaz la predicción interior de la trama, mientras que las áreas inalteradas se reconstruyen a partir de la trama anterior. Un sistema de rellenado condicional, como el indicado, debe transmitir información de direccionamiento que indique al receptor la localización de las áreas cambiantes.

Se han desarrollado algunas modificaciones del rellenado condicional básico³. Se pueden alcanzar grandes factores de reducción por codificación compensada dinámicamente, en la cual se estima el desplazamiento de una trama a otra de un objeto móvil para considerarlo en la predicción intertrama¹⁰.

Codificación por transformada

En la codificación por transformada un bloque de pixels se sustituye, a través de una transformación lineal, por un grupo equivalente de coeficientes, con el fin de poder eliminar más fácilmente la redundancia estadística y visual en los datos. Para la codificación de imagen se han utilizado la transformada discreta de Fourier, la del coseno discreta, la de Hadamard-Walsh y la de Karhunen-Loeve¹¹. La transformación da un conjunto de coeficientes relativamente incorrelados a partir de un bloque de pixels correlados, siendo luego cuantificados individualmente y codificados para la transmisión estos coeficientes (Fig. 8).

La codificación por transformada puede interpretarse como representación de un bloque de $N \times N$ pixels mediante una suma ponderada de imágenes elementales con las dimensiones del bloque. Cada imagen elemental es una subimagen de dimensiones $N \times N$, con un patrón de señal de vídeo específico. Estas imágenes son ortogonales entre sí, y los factores de ponderación son los coeficientes transmitidos. En la figura 9 se muestran las imágenes elementales de una transformada de coseno discreta con bloques de 8×8 pixels. Una imagen elemental representa el valor medio de un bloque; las otras describen las estructuras específicas de un bloque. En base al número de cambios de signo a lo largo de ambos ejes, se pueden asignar diferentes frecuencias espaciales a cada imagen elemental. Dada la menor sensibilidad del ojo humano a las altas frecuencias espaciales, los coeficientes que representan imágenes elementales con partes de alta frecuencia pueden cuantificarse groseramente. Se aplica una técnica de muestreo zonal para los sistemas de baja velocidad binaria: sólo se seleccionan para transmisión los coefi-

cientes de una zona predeterminada (usualmente los de frecuencias espaciales bajas). Los coeficientes transmitidos se cuantifican según una asignación de bitios elegida para ser proporcional a la varianza de los coeficientes¹¹.

El código de transformación y el MICD bidimensional requieren ambos alrededor de 4 bitios por pixel para tener una elevada calidad de imagen. La realización del soporte físico de un codificador por transformada es mucho más compleja. Para sistemas de baja velocidad binaria, este tipo de codificación tiene algunas ventajas, entre ellas la flexibilidad de elegir una velocidad binaria especificada mediante asignación de bitios a los coeficientes.

Codecs para transmisión de señal de vídeo

Normas para estudios de televisión y transmisión

En 1981 el CCIR emitió una Recomendación⁸ introduciendo la codificación digital por separado de las componentes de luminancia y crominancia en los estudios de radiodifusión de televisión, en lugar de los sistemas de televisión en color existentes, PAL, SECAM y NTSC. Ello aumentó la inmunidad ante ruido, distorsiones lineales y no lineales, e interferencias. Los parámetros principales son:

- codificación separada de los componentes de la señal de vídeo (Y, CR, CB)
- frecuencias de muestreo de 13,5 MHz para la luminancia y de 6,75 MHz para la crominancia
- cuantificación uniforme utilizando 8 bitios por muestra
- velocidad de transmisión de 216 Mbit s⁻¹.

Se han especificado para los estudios velocidades de muestreo digital mayores que el doble del ancho de banda, con objeto de dejar espacio para el proceso de señal, que se utilizará más en los sistemas futuros. La frecuencia de muestreo de la crominancia se elige lo bastante alta para mejorar el comportamiento del regulador cromático.

Además del nivel de calidad para estudios, están en discusión otros niveles, como uno auxiliar y otro para recogida de mensajes electrónicos. Hay que especificar también una norma de distribución común para el área de abonados, regulando los nuevos servicios de distribución de televisión digital, videoconferencia y videoteléfono.

Se requieren técnicas sencillas de transcodificación entre diferentes normas. Por esta razón, la relación entre las frecuencias de muestreo de normas de codificación

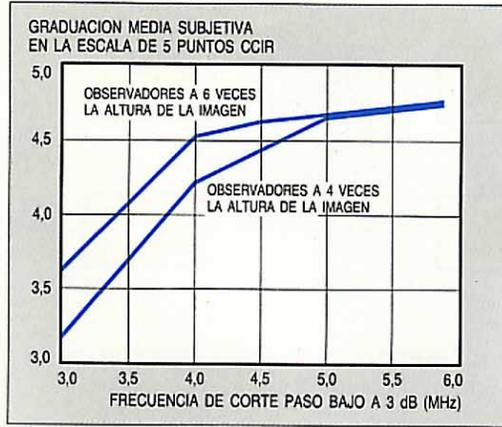


Figura 10
Calidad de la imagen en función de la frecuencia paso bajo a 3 dB. Valores medios obtenidos en siete laboratorios sobre una escala de 5 puntos del CCIR, utilizando una imagen de prueba de 625 líneas¹².

distintas debe tener factores racionales. La nueva norma de distribución debería cumplir las exigencias sobre calidad de la imagen (nivel 4 a 5 de la escala de calidad⁶) y permitir la transmisión con la jerarquía de niveles recomendados por el CCITT. Las velocidades estándar aplicables a la transmisión de señales de vídeo con calidad de radiodifusión son 139 Mbit s⁻¹ y 34 Mbit s⁻¹. Además se han presentado propuestas de una velocidad de 69 Mbit s⁻¹, por ser relativamente fácil de realizar.

En el receptor, dos parámetros son muy relevantes para la calidad de la imagen: el ancho de banda de las componentes de señal y la resolución de amplitud por muestra. La Unión Europea de Radiodifusión ha efectuado pruebas subjetivas de calidad¹² para determinar el valor de estos parámetros, cuyos resultados se exponen en la figura 10.

Se acepta generalmente un valor de 4,5 (en la escala de 5 puntos) para calidad de la televisión de estudio. Por tanto, basta un ancho de banda de 4,5 MHz en la componente de luminancia a la distancia de visión normal (seis veces la altura de la imagen). Valores similares para las señales de crominancia hacen requerir un ancho de banda de 1,5 a 2 MHz. Estos resultados indican que la velocidad de muestreo puede ser inferior a la de las normas de estudio.

Se necesitan parámetros de codificación para la norma de distribución que permitan la codificación MIC de vídeo con velocidades de transmisión de 139 Mbit s⁻¹. Se ha propuesto reducir a la mitad la frecuencia de muestreo de la señal de crominancia. Utilizando los intervalos blancos horizontales de las señales de vídeo, se puede alcanzar una velocidad binaria de 135 Mbit s⁻¹. No obstante, es útil tener una norma de codificación que sirva también de base para los codecs de menor velocidad, tales como los de 69 o 34 Mbit s⁻¹. Para evitar frecuentes cambios en la velocidad de muestreo, se propone la siguiente norma:

- frecuencias de muestreo de 10,125 MHz para la luminancia y de 3,375 MHz para la crominancia
- cuantificación con 8 bits por muestra
- velocidad de transmisión de 135 Mbit s⁻¹.

Como se expone en la figura 11, estas velocidades de muestreo se utilizan como normas comunes para las velocidades de transmisión de 140, 70 y 34 Mbit s⁻¹. Únicamente se requieren cambios en la velocidad de muestreo (por factores de 3/4 y 1/2) en la salida del estudio de televisión. Los filtros transcodificadores apropiados no se describen aquí¹³. Para los servicios de videoconferencia y videoteléfono, la digitalización se realiza a frecuencias de muestreo reducidas, eliminando la necesidad de los filtros transcodificadores.

Codecs de vídeo para el área de abonados

El codificador de vídeo tiene que digitalizar las señales de vídeo y las de sonido que las acompañan. En caso necesario, hay métodos de codificación de la fuente que reducen la velocidad binaria. La señal de televisión, la de sonido y los datos adicionales se han de multiplexar en un tren de bits común. El decodificador realiza la operación inversa.

Un codec de vídeo para velocidad de 139 Mbit s⁻¹ sólo necesita codificación MIC de vídeo (135 Mbit s⁻¹) utilizando las velocidades de muestreo de la norma de distribución antes señalada. La codificación de un canal de sonido estereofónico requiere 1 Mbit s⁻¹ (16 bits por muestra, 32 kHz de frecuencia de muestreo). Esto deja alrededor de 3 Mbit s⁻¹ para la sincronización de trama y datos adicionales. El uso de estos canales añadidos depende del tipo de servicio. En general, además del vídeo, el abonado debería disponer al menos de un canal de RDSI. En la videoconferencia, los canales adicionales podrían servir para transmitir gráficos, texto u otros datos. En la distribución de televisión interactiva, se desea también el sonido de radiodifusión y el videotex. Habría, pues, que prever la asignación flexible de dichos canales.

Los principales costes de equipo físico para este codec de vídeo corresponden al multiplexor, demultiplexor, generación y distribución del reloj y convertidores analógico/digital. Se dispondrá de circuitos VLSI de diseño a medida, que permitirán realizar el codec de vídeo en unos pocos circuitos integrados solamente.

Para una velocidad de transmisión de 69 Mbit s⁻¹, el codec debe reducir la velocidad MIC de vídeo aproximadamente a la mitad. Esto se puede lograr mediante MICD no adaptativo, así como por codificación por transformada; se utilizará el MICD por ser

más sencillo de realizar. Teniendo en cuenta la calidad de imagen requerida, se utiliza un predictor lineal bidimensional y un cuantificador fijo con 16 niveles para cada componente (luminancia y crominancia). La predicción sólo con coeficientes positivos es ventajosa en cuanto a la calidad de la imagen y a posibles efectos de inestabilidad. Para simplificar, los factores de ponderación del predictor se restringen a potencias de dos. El esquema de predicción utilizado para codificar las señales de luminancia y crominancia es:

$$\hat{s} = \frac{1}{2} s'_1 + \frac{1}{4} (s'_3 + s'_4)$$

donde s'₁ es el pixel precedente, s'₃ es el de la línea anterior justo encima del pixel actual, y s'₄ es el de dicha línea anterior delante de s'₃. Los cuantificadores para luminancia y crominancia se deducen modificando características ya publicadas⁹. Seguidamente se enumeran los niveles representativos (valores de salida de ± 255 para señales de entrada de 8 bits):

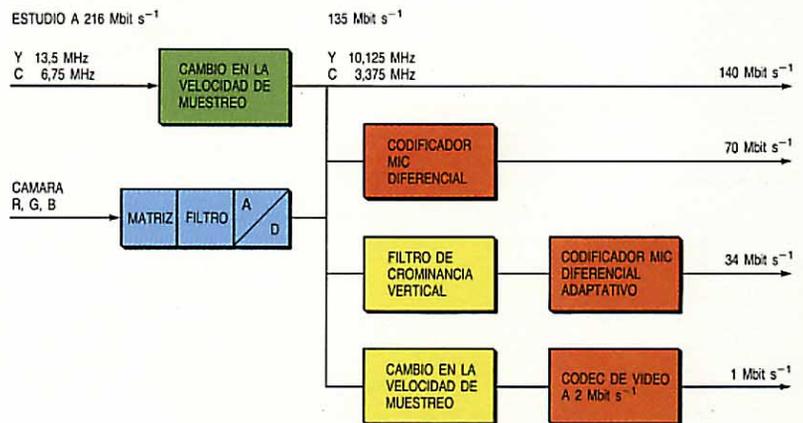
Luminancia: -75, -60, -47, -36, -25, -16, -9, -2, 3, 10, 17, 26, 37, 48, 61, 76.

Crominancia: -64, -49, -36, -25, -16, -9, -4, -1, 2, 5, 10, 17, 26, 37, 50, 65.

Los niveles de decisión están en medio de dos niveles representativos adyacentes.

Se ha investigado la realización de codecs MICD no adaptativos en circuitos integrados individuales¹⁴. Debido a restricciones de tiempo en el bucle recursivo del codificador MICD, se necesita tecnología CMOS de 2 μm¹⁵. El codec MICD de luminancia requiere unos 25.000 transistores, y 19.000 el codec MICD de crominancia. Un diseño a medida permite integrar del orden de 80.000 transistores en una pastilla, y por tanto deberían poderse realizar codecs MICD en un solo circuito integrado. Aunque

Figura 11
Familia de codecs con una frecuencia de muestreo normalizada común.



69 Mbit s⁻¹ no sea una velocidad de transmisión típica, tiene la notable ventaja de que utilizando una pastilla más que el codec de 139 Mbit s⁻¹, se reduce a la mitad la velocidad binaria.

Es preciso dividir por cuatro la velocidad binaria para conseguir una velocidad de transmisión de 34 Mbit s⁻¹. Hay que utilizar tanto intervalos blancos horizontales como verticales, así como transmisión MICD adaptativa. Además, hay que reducir la resolución vertical de crominancia transmitiendo una de cada dos líneas, lo cual es posible por ser la resolución vertical superior a la horizontal en la crominancia. Debe utilizarse un filtro digital para la crominancia vertical, a fin de eliminar el solapamiento causado por el submuestreo vertical. La calidad de imagen deseada la consigue un sistema MICD con predicción adaptativa entre tramas y dentro de la trama, y codificación con palabras de longitud variable. Por tanto, la predicción entre tramas de las tres componentes exige una memoria de trama con capacidad total del orden de 3 Mbit. Además, la codificación con palabras de longitud variable necesita una memoria tampón. En suma, la realización de un codec de 34 Mbit s⁻¹ será mucho más compleja que la del de 69 Mbit s⁻¹, y no tan económica, por lo que este codec debería utilizarse sólo en transmisiones de larga distancia y no en el terminal de abonado.

Conclusiones

La codificación se apoya tanto en bases estadísticas (teoría de la información) como psicofísicas. Las técnicas básicas de codificación de ondas, como la MIC, MICD y codificación por transformada pueden todas aplicarse a las señales de vídeo, y jugarán un papel importante en los servicios futuros de transmisión de imágenes a través de redes integradas de banda ancha. Se dispone ya de las tecnologías necesarias para la realización de un codec con mínimo número de circuitos VLSI, si bien son de esperar mayores adelantos. Así, los codecs de vídeo han llegado a una fase en la que pueden dejar de ser objeto de investigación para convertirse en componente clave de las redes de banda ancha.

Referencias

- 1 C. E. Shannon: The Mathematical Theory of Communication: *The University of Illinois Press*, 1964, Urbana.
- 2 R. G. Gallager: Information Theory and Reliable Communication: *John Wiley Inc*, 1968, Nueva York.
- 3 A. N. Netravali y J. O. Limb: Picture Coding: A Review: *Proceedings of the Institute of Electrical and Electronics Engineers*, 1980, volumen 68, n° 3.
- 4 C. E. Shannon: Coding Theorem for a Discrete Source with a Fidelity Criterion: *Information and Decision Processes*, McGraw-Hill, Nueva York, 1960.
- 5 T. M. Cornsweet: Visual Perception: *Academic Press*, 1971, Nueva York.
- 6 Recomendación 500-2 del CCIR: Método de evaluación subjetiva de las imágenes de televisión: *XV Asamblea Plenaria*, 1982, Ginebra, volumen XI.
- 7 D. E. Pearson: Transmission and Display of Pictorial Information: *Pentech Press*, 1975, Londres.
- 8 Recomendación 601 del CCIR: Parámetros de codificación de televisión digital.
- 9 P. Pirsch: Design of DPCM Quantizers for Video Signals Using Subjective Tests: *Institute of Electrical and Electronics Engineers Transactions on Communications*, julio 1981, volumen COM-29, n° 7, págs. 990–1000.
- 10 T. Ishiguro y K. Linuma: Television Bandwidth Compression Transmission by Motion Compensated Interframe Coding: *Institute of Electrical and Electronics Engineers Communication Magazine*, noviembre 1982, págs. 24–30.
- 11 W. K. Pratt: Digital Image Processing: *John Wiley Inc*, 1978, Nueva York.
- 12 Informe 962 del CCIR: Filtrado, muestreo y multiplaje de las señales de componentes de color para sistemas de modulación digital.
- 13 P. Pirsch y M. Bierling: Changing the Sampling Rate of Video Signals by Rational Factors: *Conference Record EUSIPCO*, 1983, págs. 171–174.
- 14 P. Pirsch: Design of a DPCM Codec for VLSI Realization in CMOS Technology: *Proceedings of the Institute of Electrical and Electronics Engineers*, marzo 1985.
- 15 P. Pirsch: Codec de vídeo para comunicaciones en banda ancha: *Comunicaciones Eléctricas*, 1984, volumen 58, n° 4, págs. 447–449.

Gert Bostelmann nació en Fallingbostal, Alemania, en 1944. Estudió telecomunicaciones en Braunschweig desde 1966 hasta 1972, y se graduó Dr. Ingeniero por la Universidad de Hannover en 1977. En el mismo año ingresó en Standard Elektrik Lorenz AG para trabajar en técnicas de transmisión avanzadas. Desde 1979 hasta 1982 el Dr. Bostelmann fue adjunto del director técnico de SEL, y actualmente dirige la división de sistemas del centro de investigación de SEL.

Peter Pirsch nació en 1942 en Hannover, Alemania. En dicha ciudad obtuvo el grado de ingeniero por la Fachhochschule en 1966, y los de Dipl.-Ing. y Dr.-Ing. por la Universidad en 1973 y 1979, respectivamente. Trabajó en el departamento de televisión de Telefunken entre 1966 y 1973, después se incorporó a la Universidad de Hannover como adjunto de investigación y luego ingeniero superior. De 1979 a 1981 pasó un período sabático en el departamento de investigación de comunicaciones visuales de los laboratorios Bell. El Dr. Pirsch entró en el centro de investigación de SEL en 1983, dirigiendo el departamento de proceso digital de señales.

Filtros integrados para señales de vídeo

Se ha desarrollado una nueva técnica de diseño de filtros digitales para procesar señales de vídeo en buenas condiciones económicas. Por ser repetitiva su estructura, los filtros son especialmente adecuados para realizarse en una sola pastilla VLSI, por ejemplo, en tecnología CMOS de $2\ \mu\text{m}$.

A. Ali

Centro de Investigación de
Standard Elektrik Lorenz AG, Stuttgart,
República Federal de Alemania

Introducción

La digitalización, extendida tanto al área de abonado como a redes de larga distancia, y la introducción gradual de redes digitales de comunicación de gran anchura de banda han abierto el camino para la transmisión digital de señales de vídeo.

El éxito de las pruebas de campo del BIGFON (red óptica integrada de banda ancha sobre fibra de vidrio), realizadas por SEL (Standard Elektrik Lorenz) en Alemania¹, ha demostrado las ventajas que tales redes de banda ancha ofrecerán a los abonados. Entre ellas se incluyen un número y variedad de servicios mucho mayor y una superior calidad operativa (mayor relación señal-ruido, menores diafonías, etc.). La mejora de la calidad se ha logrado utilizando tecnología digital para procesar y transmitir las señales, si bien a costa de aumentar el ancho de banda. La tecnología de la fibra óptica posibilita tanto el admitir esta mayor anchura de banda como el ampliar el número de servicios que el abonado puede recibir simultáneamente (televisión, radio-difusión estereofónica, videoteléfono, télex, teletex).

En la red BIGFON, los servicios de distribución (programas de televisión y radio) y los servicios conmutados (p. ej., el videoteléfono) se transmiten todos digitalmente desde una central al terminal del abonado. Todas las señales de televisión y radio se digitalizan en la central antes de la transmisión.

Una de las facilidades que deberá proporcionar una red digital de servicios integrados de banda ancha en el futuro, será el proceso y la transmisión digital de señales de vídeo desde una central hasta los terminales de abonado, a través de enlaces de fibra óptica.

Señales de vídeo

En general, existen dos tipos diferentes de señales de vídeo en la central: analógicas y digitales.

Señales digitales

Se transmiten señales digitales de vídeo a la central desde un estudio de televisión, por ejemplo. El CCIR ha normalizado la codificación de los componentes de las señales de televisión². La señal de luminancia se muestrea a 13,5 MHz, mientras que cada una de las dos señales de crominancia se muestrea a 6,75 MHz. Utilizando MIC de 8 bits por cada muestra, resulta para vídeo una velocidad de datos total de $216\ \text{Mbit s}^{-1}$.

Las señales digitales de vídeo se transmitirán en bucles de abonado cuya estructura se procurará conformar a la jerarquía del MIC, con el fin de asegurar su compatibilidad con la red de enlaces. Una velocidad binaria de $139.264\ \text{kbit s}^{-1}$ resulta económica para transmitir al abonado un canal básico de banda ancha; este canal debería también poder transportar otros tipos de información de datos, sincronización y señalización.

Todavía no se ha decidido la capacidad de canal que debe asignarse sólo a la señal de vídeo en el bucle de abonado. No obstante, con independencia de la norma que se adopte, parece ineludible la conversión de la velocidad binaria de vídeo de $216\ \text{Mbit s}^{-1}$ a $135\ \text{Mbit s}^{-1}$ o menos. La conversión habrá de realizarse mediante filtros digitales³ para asegurar que no se producen errores por solapamientos.

Señales analógicas

Las señales de las fuentes analógicas de los programas de televisión se deben codifi-

car digitalmente antes de la transmisión. Hay que elegir una frecuencia de muestreo adecuada, junto con una técnica de conversión analógico-digital (p. ej., MIC). Según un conocido teorema, la frecuencia de muestreo debe ser al menos el doble de la máxima componente frecuencial de la señal de vídeo (frecuencia de Nyquist). La experiencia muestra que únicamente influyen aquellas componentes frecuenciales cuya amplitud sea mayor de -40 dB con respecto al nivel de baja frecuencia de la señal de vídeo. Por lo tanto, para eliminar errores de solapamiento hay que limitar la banda de la señal de vídeo antes de muestrearla. Hasta ahora, se han utilizado generalmente filtros analógicos con sus inherentes desventajas (volumen, coste elevado, imprecisión, envejecimiento, ajuste). Además, no son aptos para realizarse en tecnología VLSI.

Los filtros digitales se pueden utilizar como filtros anti-solape. Como ejemplo, la señal analógica de luminancia se hace pasar por un filtro analógico sencillo (sólo con resistencias y capacidades) que proporciona los 40 dB de atenuación necesarios a una frecuencia relativamente elevada ($33,75$ MHz). Un conversor analógico-digital muestrea luego la salida del filtro a $67,5$ MHz y entrega muestras de 8 bits a esa velocidad. La señal pasa después a través de un filtro digital que reduce el contenido frecuencial de la señal de vídeo a fin de evitar errores por solapamiento, para lo cual desecha cuatro de cada cinco muestras consecutivas (entresacado), alcanzando así la velocidad de muestreo de $13,5$ MHz normalizada por el CCIR. Los filtros digitales pueden considerarse como conversores de la velocidad de muestreo.

Otra utilización posible de los filtros digitales es la conversión digital-analógico en el terminal de abonado. La señal digital recibida se debe convertir a formato analógico antes de pasar al monitor de televisión, necesitando además un filtro paso bajo que reconstruya completamente la señal analógica resultante de dicha conversión. Puede introducirse un filtro digital antes del conversor digital-analógico para conformar el espectro de frecuencia de la señal de tal manera que sea suficiente un filtro analógico sencillo (resistencias y capacidades) a la salida del referido conversor.

Estrategia de diseño

SEL ha elaborado una completa metodología de diseño destinada al desarrollo de filtros digitales para las aplicaciones anteriores. Ante todo, sólo se han considerado estructuras no recursivas de fase lineal, las

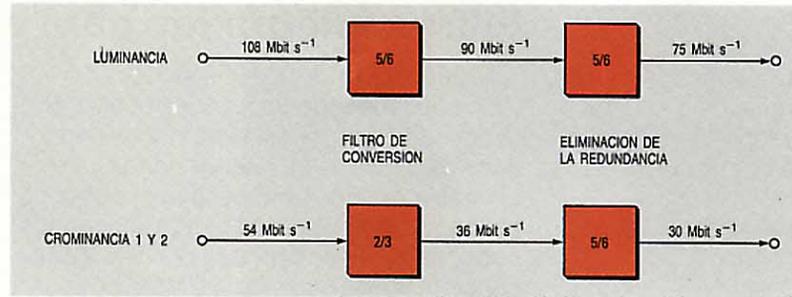


Figura 1
Esquema para la reducción de la velocidad de datos de la señal de vídeo desde 216 hasta 135 Mbit s^{-1} .

cuales no contienen multiplicadores y por ello no presentan problemas de velocidad ni de complejos circuitos. Únicamente se permiten simples operaciones de desplazamiento (equivalentes a multiplicar por potencias de 2), realizables mediante programación cableada sin necesidad de añadir soporte físico alguno.

En segundo lugar, se diseña el filtro no recursivo elemental utilizando programación entera mixta⁴, método que ha demostrado ser más eficiente que otros conocidos para el diseño de filtros cuyos coeficientes son potencias de 2 (o sumas de potencias de 2)⁴.

En tercer lugar, lo que interesa es el diseño de filtros no recursivos elementales como bloque básico. Esto es, el filtro básico no satisface por sí mismo los exigentes requisitos de atenuación de amplitud cuyo fin es suprimir los efectos debidos al solapamiento. La operación de filtrado restante se puede conseguir colocando en cascada bloques de filtros de primero o segundo orden elementales con las características requeridas (eliminación de banda o función correctora de paso alto). En caso necesario, el bloque de filtro original se puede conectar en cascada consigo mismo con objeto de aumentar la atenuación.

Por último, algunas de las secciones en cascada no necesitan operar a la velocidad de muestreo más alta (la de interpolación). Un método eficaz sería llevar a cabo la interpolación y el entresacado en diversas etapas, realizando en cada una de ellas algunas operaciones de filtrado.

Esta estrategia de diseño da lugar a estructuras de filtro sencillas, que pueden trabajar a alta velocidad. Tales estructuras son repetitivas y por tanto adecuadas para realizarse en VLSI, dada la sencillez de su diseño lógico, trazado y pruebas.

Filtro de conversión de velocidad de datos

Los programas de televisión se transmiten desde un estudio hasta la central a 216 Mbit s^{-1} . Esta velocidad binaria repre-

senta una señal de luminancia de 108 Mbit s⁻¹ (frecuencia de muestreo de 13,5 MHz con 8 bits por muestra) y dos señales de crominancia, cada una muestreada a 6,75 MHz con 8 bits por muestra. A continuación la central distribuye los programas a los abonados sobre cable de fibra óptica por canales de 139.264 kbit s⁻¹.

Aunque hay muchos esquemas posibles de conversión de velocidad de datos, el aquí considerado requiere una relación de conversión de 5/6 para la luminancia y de 2/3 para la crominancia, como se indica en la figura 1. Por tanto, la nueva frecuencia de muestreo para la señal de luminancia es 11,25 MHz (13,5 MHz × 5/6), resultando una velocidad binaria de 90 Mbit s⁻¹. Después de eliminar la información redundante en cada línea de televisión (otro factor adicional de 5/6), la velocidad binaria de luminancia queda reducida a 75 Mbit s⁻¹.

En el caso de las señales de crominancia, la nueva frecuencia de muestreo es 4,5 MHz (6,75 MHz × 2/3), lo que da una velocidad de 36 Mbit s⁻¹. Al eliminar la información redundante en las líneas de televisión se reduce aún más la velocidad binaria de crominancia, quedando en 30 Mbit s⁻¹. Todo ello da una velocidad binaria final de 135 Mbit s⁻¹ para la señal digital de vídeo.

Los filtros digitales son necesarios para ambas relaciones de conversión, aunque por razones de brevedad nos limitaremos a considerar el diseño del filtro conversor de la señal de luminancia.

Proceso de diseño

SEL ha utilizado un proceso de diseño compuesto de tres etapas. Primeramente, la velocidad de muestreo de la señal de entrada se multiplica por un factor de cinco al introducir cuatro ceros entre cada dos-muestras adyacentes. A continuación, la señal resultante se pasa por una combinación de filtros paso bajo $H(z)$, según indica la figura 2. Finalmente, la velocidad de muestreo de la señal a la salida del filtro se divide por seis, suprimiendo cinco de cada seis muestras consecutivas.

La combinación de filtros consiste en una cascada de dos bloques de filtros básicos, cada uno con una función de transferencia $B(z)$, y una combinación de cuatro secciones de primer orden cuya función de transferencia global es $F(z)$. Por tanto, $H(z) = B(z) * B(z) * F(z)$.

El bloque de filtro básico $B(z)$ es un filtro no recursivo de fase lineal⁴. Su respuesta en amplitud, que no cumple las especificaciones, se muestra en la figura 3. El filtro es de orden 17; dos de los coeficientes son cero, seis requieren únicamente operaciones de desplazamiento, cuatro tienen una magnitud de uno, y sólo seis coeficientes

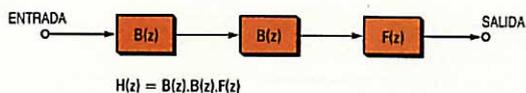


Figura 2
Combinación de filtros digitales $H(z)$.

implican operaciones de desplazamiento y adición. Se han modificado los coeficientes del filtro de modo que no se necesiten operaciones de adición (todos son potencias enteras de dos):

$$\begin{aligned}
 a_0 = a_{17} &= 2 & a_1 = a_{16} &= 0 \\
 a_2 = a_{15} = a_3 = a_{14} &= -2 \\
 a_4 = a_{13} &= -1 & a_5 = a_{12} &= 1 \\
 a_6 = a_{11} &= 4 \\
 a_7 = a_{10} = a_8 = a_9 &= 8.
 \end{aligned}$$

Seguidamente, el filtro básico se conecta en cascada consigo mismo y con la siguiente combinación en cascada $F(z)$ de secciones de primer orden:

$$\begin{aligned}
 F(z) &= [(1+z^2)/2] * [(1+z^3)/2] \\
 &* [(1+z^4)/2] * [(9-z^5)/8]
 \end{aligned}$$

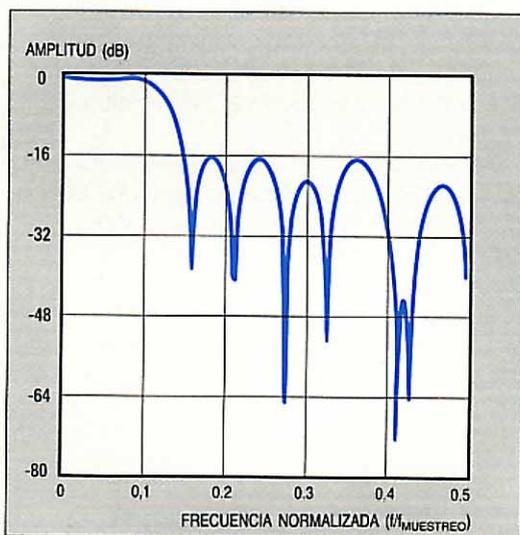


Figura 3
Respuesta en amplitud del filtro original $B(z)$.

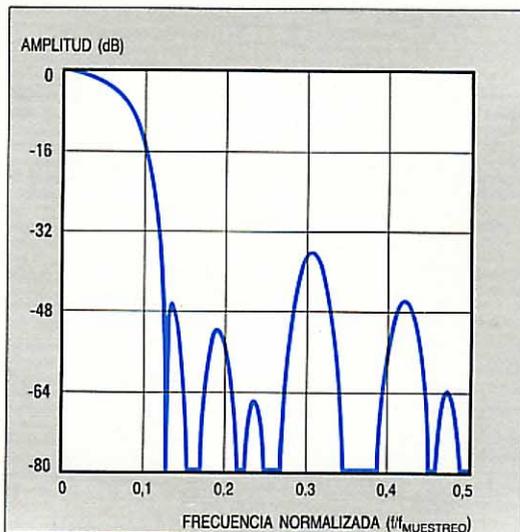


Figura 4
Respuesta en amplitud del filtro combinado $H(z)$.

Figura 5 (izquierda) Espectro de amplitud de la señal de vídeo de entrada.

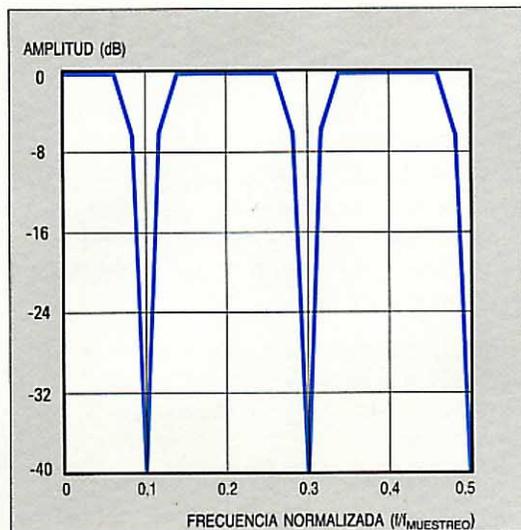
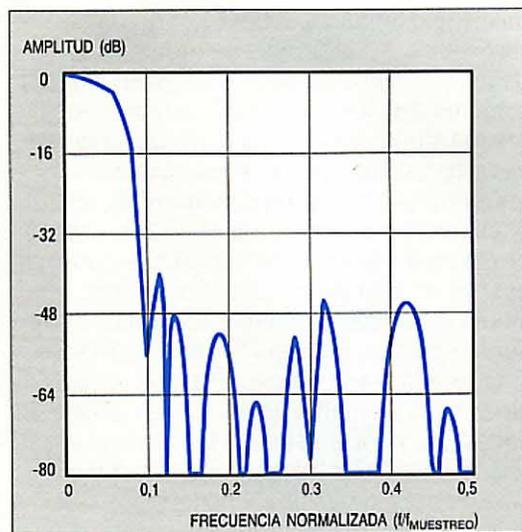


Figura 6 (derecha) Espectro de amplitud de la señal de vídeo filtrada.



en la que

$$z = \exp(-j\omega T)$$

$$T = (1/67,5) \mu s = 14,815 \text{ ns.}$$

Las tres primeras secciones de $F(z)$ agudizan la respuesta a frecuencias, mientras que la cuarta sección corrige algo la amplitud en la banda de paso. La respuesta del filtro total combinado $H(z)$ se muestra en la figura 4.

Para apreciar cómo afecta este filtro a la señal de vídeo, el espectro de dicha señal (Fig. 5), que es repetitivo como consecuencia de quintuplicarse la velocidad de muestreo, se multiplica por la función de transferencia del filtro $H(z)$, dando como resultado la señal de vídeo filtrada de la figura 6. El espectro de señal tiene 3,4 dB de atenuación a 4 MHz, y una atenuación mínima de 40 dB en la banda eliminada, a partir de

6,4 MHz. A la nueva frecuencia de Nyquist de 5,625 MHz ($6,75 \times 5/6$ MHz), la atenuación es mayor de 15,5 dB.

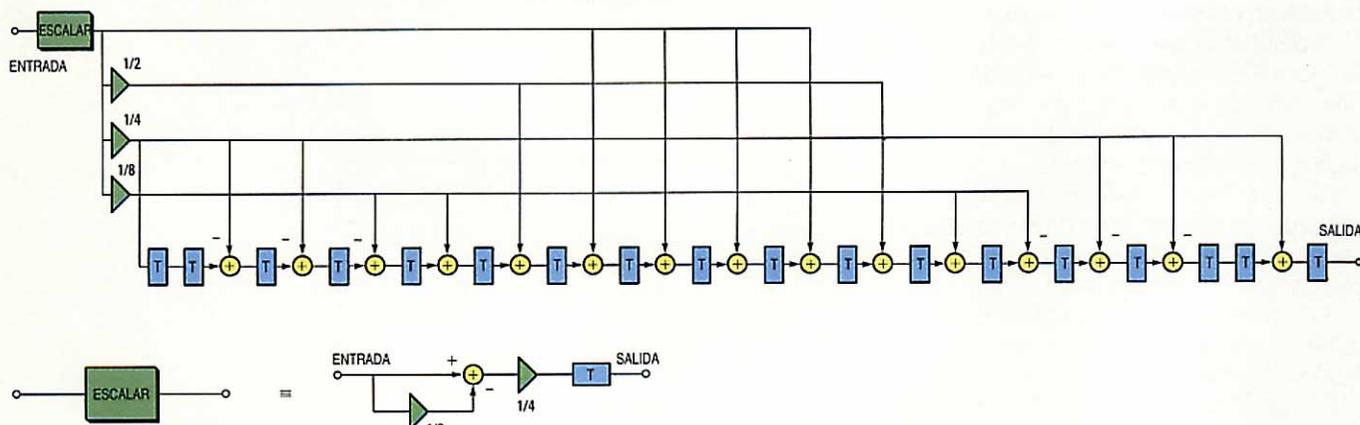
En la figura 7 se muestra la estructura del filtro. La entrada del filtro $B(z)$ debe reducirse de escala para asegurar una ganancia unidad en el origen. Puede ser de interés mencionar que si se hubiera utilizado el algoritmo de permutación de Remez para diseñar un filtro no recursivo de fase lineal con las mismas especificaciones, dicho filtro habría tenido un orden mínimo de 23 (con multiplicadores de precisión infinita).

Consideraciones generales de diseño

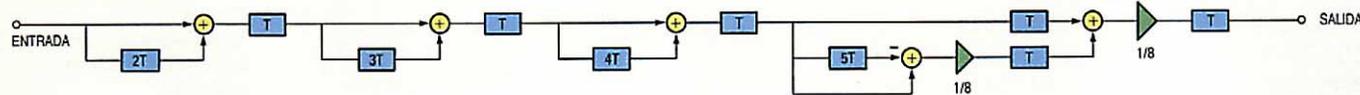
El escalamiento del bloque de filtro $B(z)$ implica un restador, y consigue un factor de escala de 0,21875 en lugar del valor de 0,2222 requerido. Este error (1,56%) puede despreciarse. La operación de escala-

Figura 7 Realización del soporte físico de las secciones de filtro $B(z)$ y $F(z)$.

(a) ESTRUCTURA PARA $B(z)$



(b) ESTRUCTURA PARA $F(z)$



miento del primer bloque $B(z)$ se puede realizar en la señal de vídeo original (antes de introducir los ceros) a la velocidad reducida de 13,5 MHz.

Los sumadores de $B(z)$ operan a la elevada velocidad que puede alcanzarse mediante la tecnología CMOS de $2\ \mu\text{m}$. No obstante, debería poderse utilizar el método de "guardar el acarreo", o cualquier otra estructura sistólica rápida, con el fin de incrementar aún más la velocidad de los sumadores.

La sección en cascada $F(z)$ comprende cinco sumadores que operan a 67,5 MHz cada uno, lo cual se asegura introduciendo retardos adicionales entre las secciones.

Como la sección paso alto de $F(z)$ sólo tiene un elemento de retardo de $5T$, se puede utilizar como primera etapa de filtrado, actuando sobre la señal de vídeo antes del aumento de la frecuencia de muestreo. En este caso operaría a una velocidad menor (dos sumas en 74 ns), con el consiguiente ahorro de dos registros de desplazamiento.

Las operaciones de escalamiento se pueden incluir en los sumadores asociados.

Si se desea, se puede corregir aún más la característica de amplitud de la señal de vídeo en la banda de paso mediante utilización de una simple combinación RC después del conversor digital-analógico.

Si bien nos hemos concentrado en la respuesta en amplitud y en asegurar que el filtro tiene fase lineal, se podrían investigar otros parámetros (p. ej., impulso de $2T$) de la señal de vídeo para garantizar que se cumplen otras especificaciones del CCIR.

Conclusiones

SEL ha desarrollado una nueva estrategia de diseño de filtros digitales sencillos para aplicaciones de proceso de vídeo. En concreto, los filtros son adecuados para la

conversión de la velocidad de datos de las señales digitales de vídeo. Estos filtros de fase lineal no requieren multiplicadores. Los coeficientes se reemplazan únicamente por operaciones de desplazamiento, consiguiendo estructuras de filtros digitales simples cuyo soporte físico es relativamente sencillo y que pueden trabajar a alta velocidad. A pesar de su simplicidad, las estructuras resultantes realizan operaciones rigurosas de filtrado que imponen estrictas especificaciones y velocidades de operación en el filtro.

La estructura total, formada por sencillas secciones en cascada de filtro de bajo orden, es repetitiva y en consecuencia particularmente adecuada para realizarse en tecnología VLSI (p. ej., con proceso CMOS de $2\ \mu\text{m}$).

Referencias

- 1 H. Ohnsorge: BIGFON und seine Nutzungsmöglichkeiten: *Proceedings of the Telematica 84*. 1984, parte 2, VDE-Verlag, Stuttgart, Alemania.
- 2 Parámetros de codificación de televisión digital: *Recomendación 601 del CCIR*.
- 3 P. Pirsch y M. Bierling: Changing the Sampling Rate of Video Signals by Rational Factors: *Proceedings of the European Signal Processing Conference*, 1983, Erlangen, Alemania, págs. 171–174.
- 4 Y. C. Lim: Efficient Special Purpose Linear Programming for FIR Filter Design: *Institute of Electrical and Electronics Engineers Transactions on Acoustics, Speech and Signal Processing*, volumen ASSP-31, nº 4, agosto 1983, págs. 963–968.

A. Ali fue inscrito miembro del City and Guilds del London Institute, y se graduó BSc y PhD por el Imperial College of Science and Technology de Londres, en 1976 y 1980, respectivamente. En 1980 consiguió la beca Alexander-von-Humboldt, que invirtió en la Universidad Bochum de Alemania. Desde 1981 pertenece al Centro de Investigación de SEL, donde inicialmente trabajó en estudios básicos relacionados con sistemas RDSI de banda ancha. Más recientemente, el Dr. Ali ha dirigido un grupo de proceso de señales encargado de la concepción y diseño del equipo electrónico para sistemas RDSI de banda ancha utilizando tecnología VLSI.

Nuevo enfoque de la conformación de haz adaptativa

Una nueva técnica de control supera el problema de falta de convergencia inherente a los algoritmos convencionales de conformación de haz adaptativa. Es apropiada para realizarse como un conjunto sistólico.

P. J. Hargrave

C. R. Ward

Standard Telecommunication Laboratories Limited, Harlow, Inglaterra

Introducción

El objetivo fundamental de un conformador de haz es combinar las salidas de los elementos de una red de antenas para producir un patrón de haz de campo lejano que optimice en un sentido determinado la recepción de una señal deseada. Esto puede realizarse sin más que elegir los pesos, en amplitud y fase, con que se combinan las salidas de los elementos para producir un haz dirigido hacia la posición de esa señal. Además, el haz puede configurarse de modo que sus características en los lóbulos laterales reduzcan las interferencias interceptadas por la red. Es también

posible explotar las técnicas de conformación de haz adaptativa para obtener nulos en las orientaciones específicas de las fuentes de interferencia.

En la figura 1 se muestra la combinación más sencilla de una red de antenas y un conformador de haz espacial. La red de ponderación se ajusta mediante un controlador de patrones de haz, que generalmente deduce los pesos aplicados de la información sobre la señal deseada introducida de antemano. El concepto puede ampliarse para que el controlador de haz sea capaz de responder adaptativamente al entorno electromagnético, proporcionándole entradas procedentes de los elementos de antena, y, en algunos algoritmos, muestras de la salida instantánea de la red de ponderación.

Es mejor configurar el controlador de patrones de haz con circuitos digitales para superar los errores prácticos inherentes a las realizaciones analógicas, y también para permitir una mayor flexibilidad en la selección del algoritmo de control. Si se utiliza un algoritmo de control de "bucle cerrado", los desplazamientos y la no linealidad en la red de ponderación se corrigen automáticamente, haciendo posible adoptar una realización analógica del conformador de haz¹.

Figura 1
Disposición sencilla de conformador de haz.

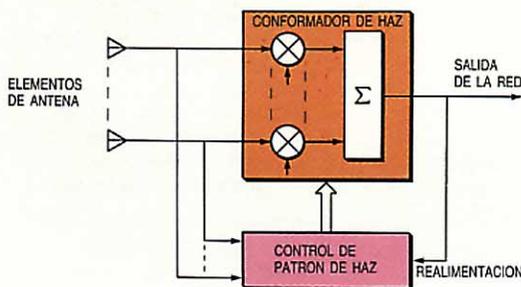
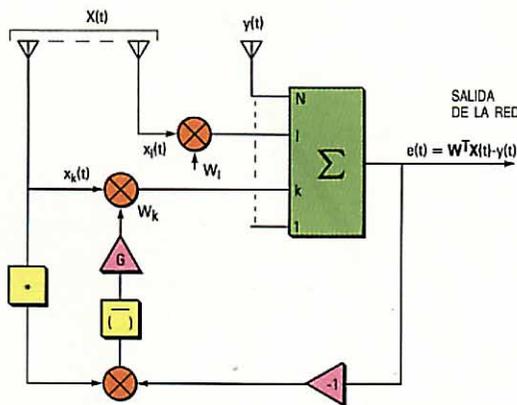


Figura 2
Procesador adaptativo de bucle cerrado
* conjugada compleja
(-) error esperado.



Técnicas de bucle de correlación

La figura 2 muestra un típico proceso de control de bucle cerrado. El peso aplicado a la salida de uno de los elementos se fija en la unidad; los otros pesos se actualizan estimando las correlaciones entre la salida sumada y las señales de los elementos individuales. El algoritmo trata de minimizar la potencia en la salida del conformador de haz, con la limitación impuesta por el peso

fijo aplicado a uno de los elementos, y utiliza un procedimiento iterativo, durante el cual los pesos variables convergen a los valores deseados por el camino de máxima pendiente.

Expresada en formato discreto en el tiempo, la ecuación de actualización adaptativa de pesos es:

$$\mathbf{W}(k+1) = \mathbf{W}(k) - \mu \mathbf{G}(k)$$

$\mathbf{W}(k)$ es un vector complejo que describe los pesos aplicados en la actualización k -ésima, y $\mathbf{G}(k)$ es el gradiente de la potencia de salida de la red respecto a los cambios en el vector de ponderación después de dicha actualización. El factor μ es el factor de ganancia de la actualización. El gradiente se mide por el proceso de correlación antes mencionado, y se representa matemáticamente por el producto de la matriz de covarianza \mathbf{M} y el vector de ponderación actual $\mathbf{W}(k)$. Las componentes de \mathbf{M} representan las correlaciones entre las salidas de todos los posibles pares de elementos de la red.

El vector de ponderación, en régimen permanente, obtenido por el procedimiento anterior viene dado por \mathbf{W} , que es proporcional a $\mathbf{M}^{-1}\mathbf{C}$, siendo \mathbf{C} un vector columna con todos los elementos iguales a cero, excepto el correspondiente al elemento de peso fijo, que es la unidad². Esta ponderación produce un patrón de haz de campo lejano con nulos dirigidos a todas las fuentes de señal, hasta un número máximo alcanzable igual al número de elementos de la red menos uno. En la figura 3 se muestra un ejemplo de la forma de dicho patrón. Este algoritmo adaptativo de orientación por nulos constituye una potente técnica de rechazo de interferencia, con tal que la adaptación se haga solamente cuando se sepa que está ausente la señal deseada, para evitar que ésta se anule junto a cualquier interferencia.

El proceso de control puede, sin embargo, ser modificado para permitir la adaptación en presencia de la señal deseada, para lo cual se ha de incluir un conocimiento *a priori* sobre tal señal y deducir el vector de ponderación correspondiente al máximo de la relación señal deseada-ruido³. Dicho vector en régimen permanente viene entonces dado por \mathbf{W} que es proporcional a $\mathbf{M}^{-1}\mathbf{S}^*$, donde \mathbf{S} es un vector columna de exponenciales complejas con factores de fase que caracterizan un frente de ondas que llega a los elementos de la red desde la dirección de la señal deseada. Además de producir nulos dirigidos a las interferencias, el haz resultante tendrá también un lóbulo principal orientado hacia la señal deseada. En la figura 4 se muestra un ejemplo de este tipo de patrón de haz adaptativo.

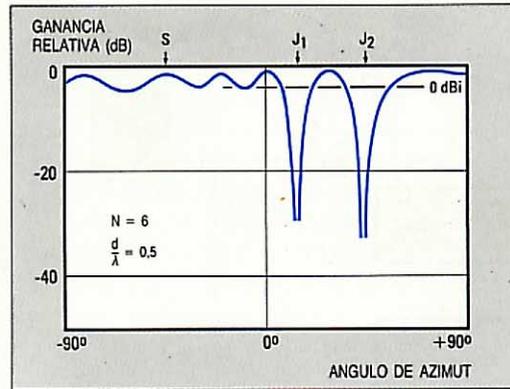


Figura 3
Patrón de haz de procesador adaptativo de pesos con variación discreta.

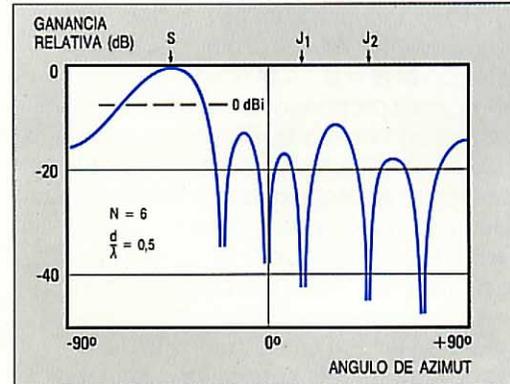


Figura 4
Patrón de haz adaptativo dirigido.

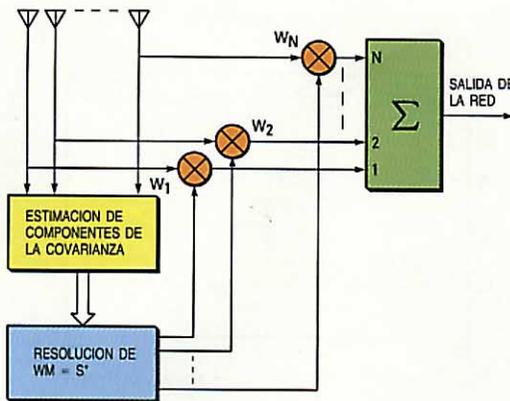
Solución directa por inversión de la matriz de muestras

La utilización del proceso de control de máxima pendiente entraña un serio problema, cual es la deficiente convergencia en ciertas condiciones adversas de señal. Este problema puede ser superado merced al proceso de inversión de la matriz de muestras⁴, que calcula los pesos necesarios directamente a partir de muestras de las ondas de salida de los elementos. En vez de iterar hacia la solución deseada, se obtiene un alto grado de compensación después de $2N$ muestras solamente (N es el número de elementos de la red). La calidad del vector de ponderación mejora cuando aumenta el número de muestras, y desaparece el problema de las condiciones adversas para la convergencia. Por calcularse dicho vector directamente de las salidas de los elementos, se denomina también a este procedimiento método de solución directa.

En la figura 5 se muestra una versión simplificada de una antena adaptativa que emplea esta técnica, apreciándose la complejidad del algoritmo resultante.

El procesador adaptativo tiene ahora tres subsistemas esenciales. El primero estima la matriz de covarianza que caracteriza el entorno a partir de un bloque de datos de entrada de la red de antenas. Esta matriz se

Figura 5 Diagrama de bloques simplificado del proceso de inversión de la matriz de muestras.



transfiere después a un procesador matemático de alta velocidad que resuelve la ecuación $MW = S^*$ ó $MW = C$, según convenga, para proporcionar la solución de ponderación deseada. Finalmente se aplica el peso al conformador de haz. Dado que ya no existen las propiedades autocorrectoras de un sistema de bucle cerrado, el conformador de haz se debe realizar con circuitos digitales para asegurar un buen comportamiento en compensación.

Aunque la técnica de inversión de la matriz de muestras supera los problemas de convergencia inherentes al algoritmo de máxima pendiente, introduce el nuevo requisito de una precisión aritmética extremadamente elevada durante los cálculos digitales. Esto sucede porque el cálculo de la matriz de covarianza de muestras lleva consigo un proceso relacionado con los cuadrados de los datos y, en muchas aplicaciones prácticas, resulta un conjunto de ecuaciones lineales muy pobremente condicionadas, que probablemente conduzcan a inestabilidades numéricas y degradación significativa de las prestaciones del sistema, a menos que los cálculos aritméticos sean muy precisos.

Además, el trabajo requerido por los cálculos, puede limitar seriamente la velocidad de actualización del vector de ponderación en caso de usar las convencionales arquitecturas secuenciales de proceso de

señal. Esto ha determinado que la primera generación de conformadores de haz adaptativo haya utilizado la técnica de bucle cerrado de máxima pendiente, y se haya empleado en aplicaciones que podían aceptar las características de convergencia correspondientes.

Como el problema de precisión surge en la estimación de la matriz de covarianza, es procedente averiguar si puede calcularse W directamente de los datos elementales instantáneos, sin necesidad de evaluar explícitamente M . Los recientes trabajos efectuados en Standard Telecommunication Laboratories dejan fuera de duda tal posibilidad, lo cual ha hecho reformular el algoritmo de solución directa en una forma que no requiere el cálculo explícito de la matriz de covarianza, y que por tanto es mucho menos sensible al uso de aritmética de precisión limitada.

En los apartados siguientes se estudia la aplicación de este algoritmo al problema de minimizar la potencia de salida con la restricción de un peso elemental fijo. Esto se puede transformar fácilmente para dar solución al problema de maximización de la relación señal-ruido mediante un adecuado proceso previo o posterior.

Algoritmo de dominio de datos

Según muestra la figura 6, se puede formar una matriz X con una sucesión de datos instantáneos procedentes de los elementos con peso variable, así como un vector columna y con los datos del elemento de peso fijo. El vector de ponderación deseado W es, entonces, el que hace mínima la longitud del vector de error e , definido por

$$e = XW - y.$$

En lugar de "elevar al cuadrado ambos miembros" de esta expresión, procedimiento que conduce al algoritmo de covarianza numéricamente sensible, pueden multiplicarse dichos miembros por una serie de transformaciones matriciales unitarias, las cuales corresponden a rotaciones de los ejes de coordenadas del espacio complejo en el que trabajamos. Tales rotaciones dejan necesariamente inalterada la longitud del vector e , y eligiéndolas convenientemente el miembro de la derecha de la ecuación se reduce a la siguiente forma triangular:

$$\begin{bmatrix} R \\ 0 \end{bmatrix} W = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

siendo R una matriz con ceros por debajo de la diagonal principal. Se ve claramente que la condición correspondiente a la longitud

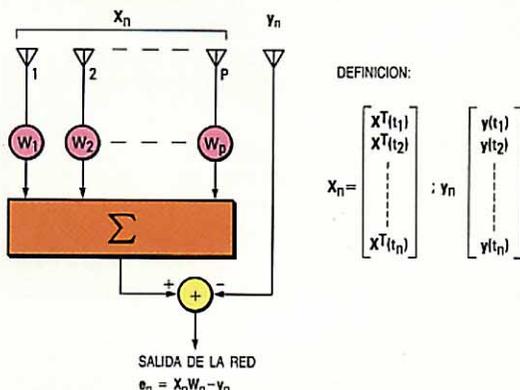


Figura 6 Notación del ordenamiento adaptativo.

mínima del vector \mathbf{e} es $\mathbf{RW} = \mathbf{b}_1$, de donde puede calcularse \mathbf{W} por un proceso de sustitución hacia atrás.

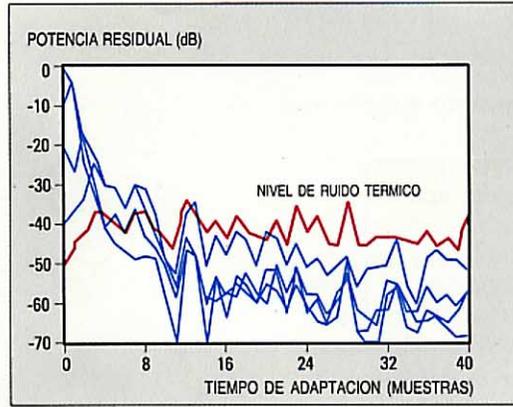
La solución directa al problema de antena adaptativa ha sido, pues, reformulada en base a \mathbf{X} , que simplificando es la raíz cuadrada de la matriz de covarianza \mathbf{M} . Ello tiene la ventaja de reducir la susceptibilidad al efecto de precisión aritmética finita. En la figura 7 se ilustra este punto, comparando las características de convergencia de los algoritmos de inversión de matriz de muestras y de dominio de datos para una situación donde existen cuatro fuentes de interferencias desiguales. En los cálculos con ambos algoritmos se utilizan longitudes de palabra idénticas. Se aprecia claramente que la técnica de inversión de la matriz de muestras está al borde de la inestabilidad, según indican las fluctuantes contribuciones de las interferencias a la potencia de salida del conformador de haz. Sin embargo, no hay indicios de problemas de esta naturaleza con el algoritmo de dominio de datos trabajando sobre idénticos datos de entrada.

Consideraciones sobre el proceso paralelo

Un aspecto interesante del algoritmo de dominio de datos, es que el conjunto triangular de ecuaciones considerado anteriormente puede ser actualizado de modo instantáneo, introduciendo cada nuevo vector de datos recibido por la red y aplicando una serie de transformaciones rotacionales que vuelvan a triangularizar las ecuaciones. Pueden seleccionarse estas rotaciones para operar por turno sobre el nuevo vector de datos y las filas sucesivas del actual conjunto triangular de ecuaciones. Si tales filas se almacenan en los procesadores rotacionales indicados en la figura 8, puede introducirse por arriba un nuevo vector de datos y aplicar una rotación al primer procesador para poner a cero el componente más a la izquierda del vector. El vector modificado, de longitud reducida, pasa seguidamente al segundo procesador rotacional. El procedimiento se repite hasta que todos los componentes menos uno del vector de datos hayan pasado a cero, y se haya actualizado eficazmente el conjunto triangular de ecuaciones.

Este procedimiento admite solapamientos puesto que, en cuanto el primer procesador ha completado su tarea, puede empezar a trabajar con el vector de datos siguiente. Por tanto, los procesadores rotacionales de la estructura mostrada en la figura 8, pueden operar independientemente sobre datos de instantes diferentes.

(a) INVERSION DE LA MATRIZ DE MUESTRAS



(b) DOMINIO DE DATOS

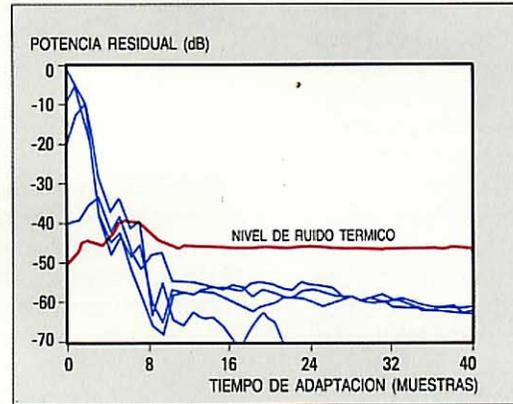


Figura 7 Resultados de simulación por ordenador, comparando los algoritmos de inversión de la matriz de muestras y de dominio de datos.

De particular interés es el significado del componente que resta del vector de datos transferido después de una actualización completa del sistema triangular. Sorprendentemente, representa una versión a escala de la salida deseada del conformador de haz que se hubiera obtenido de calcular el vector de ponderación por sustitución hacia atrás y aplicar esa solución al vector de datos. El factor de escala responde a los

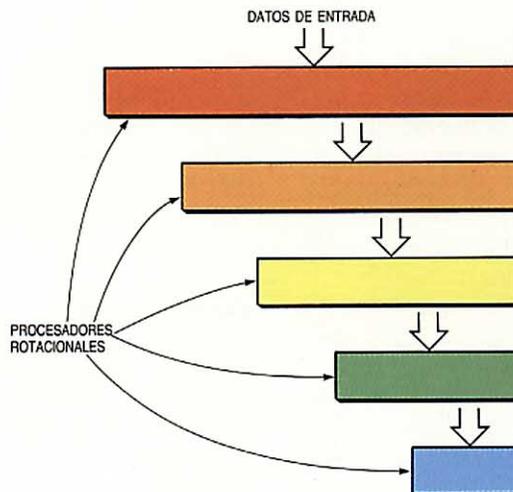


Figura 8 Descomposición triangular mediante procesadores rotacionales paralelos.

parámetros que caracterizan las rotaciones individuales aplicadas al vector de datos.

Realización en ordenamiento sistólico

Según se muestra en la figura 9, la red de procesadores rotacionales solapados puede realizarse mediante un conjunto sistólico triangular de nodos de proceso⁵. En la figura, cada ordenación horizontal de nodos realiza una transformación rotacional, con los parámetros de rotación calculados en los nodos circulares (o periféricos) y la rotación aplicada por los nodos cuadrados (o internos).

Los datos de entrada, en forma vectorial, se aplican al conjunto sistólico en secuencias de impulsos desfasados, y luego pasan por la red de nodos de proceso al ser aplicados sucesivos impulsos de reloj al conjunto. Por tanto, hay una acción regular de "latido" que define la operación del ordenamiento, y por la cual se utiliza el término sistólico (sístole es la contracción del corazón que impulsa la sangre y mantiene la circulación).

El ordenamiento triangular sistólico mostrado en la figura 9 actúa como un embudo que recoge por arriba el conjunto de datos en bruto y luego, aplicadas las necesarias transformaciones rotacionales, entrega el haz conformado resultante en la etapa final. En esta estructura, las uniones entre los nodos periféricos proporcionan acceso al factor de escala para el cálculo del haz conformado resultante.

Consideraciones prácticas

Cada nodo del conjunto sistólico de la figura 9 se puede realizar utilizando un avanzado circuito integrado de procesador digital de señal programable, con capacidad inmediata de coma flotante y facilidades de entrada/salida mejoradas. En la práctica, es de desear que los nodos sean programables, ya que con un solo diseño para el nodo podrían realizarse algoritmos diferentes y por tanto distintos ordenamientos sistólicos.

Las aplicaciones de una red de conformación de haz adaptativa en el campo de sistemas de comunicaciones van desde la transmisión de información de voz por radio en HF y VHF, con decenas de kilociclos de ancho de banda, hasta enlaces de datos de amplio espectro con bandas de hasta 10 MHz. Lo determinante es cumplir los requisitos del rango superior de estas aplicaciones, puesto que cada nodo de proceso debe realizar su función en un solo periodo de muestreo de Nyquist (aproximadamente de 50 a 100 ns) si el proceso adaptativo ha de operar "en línea".

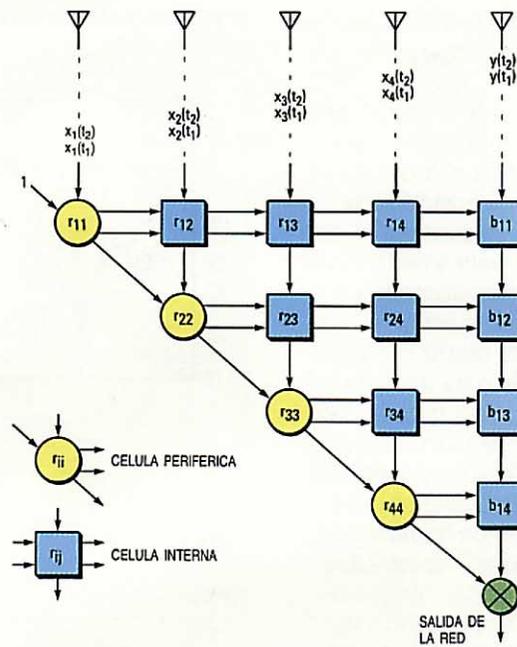


Figura 9
Ordenamiento sistólico para conformación de haz adaptativo.

La célula periférica es la que tiene el algoritmo de nodo más complicado de los dos existentes, puesto que requiere unas 10 operaciones en serie de coma flotante. Para la frecuencia de 10 MHz, límite superior postulado, hay que realizar cada operación de coma flotante en menos de 10 ns. Una complicación más de la célula periférica es la necesidad de dos funciones aritméticas no lineales: una raíz cuadrada y una reciprocidad. En un diseño práctico del nodo, estas operaciones pueden realizarse mediante una tabla de consulta.

Para aplicaciones a sistemas de comunicación, el tamaño de la red de antenas sería típicamente menor de diez elementos. Esta limitación práctica se debe generalmente a las características de complejidad del sistema y a la restricción que las dimensiones de la plataforma de la antena suelen imponer sobre el tamaño de dicha red. Considerando una red de seis elementos como representativa, el procesador sistólico asociado necesitaría 20 nodos de proceso. Tal sistema sólo sería práctico si cada célula de proceso se realizara en una pastilla o un grupo pequeño de ellas. La futura generación de antenas adaptativas que incorporen arquitecturas de proceso sistólico deberán por tanto explotar la tecnología VLSI en todas sus posibilidades.

Conclusiones

La realización en ordenamiento sistólico de un nuevo algoritmo de conformación adaptativa de haz, ofrece ventajas sobre las anteriores técnicas aplicadas al respecto. El algoritmo proporciona una óptima caracte-

rística de convergencia, y aún más importante es que ofrece una operación inherentemente estable utilizando precisión aritmética limitada con datos muy poco condicionados.

La adaptación rápida será un requisito previo esencial de muchas aplicaciones militares para obtener una mejor dinámica de las plataformas de antenas, hacer frente a artificiosas interferencias y admitir rápidos cambios de frecuencia en las formas de onda. El desarrollo de nodos de proceso de elevadas prestaciones con diseños VLSI, permitirá la realización práctica de arquitecturas paralelas de proceso en equipos compactos. Además, los circuitos VLSI proporcionarán capacidades de proceso acomodadas a las crecientes exigencias de comunicaciones de banda ancha seguras.

Agradecimientos

Este trabajo ha sido realizado con el apoyo del Procurement Executive del Ministerio de Defensa del Reino Unido. El artículo se publica con autorización de Standard Telephones and Cables plc, Londres.

Referencias

- 1 J. G. Searle y C. R. Ward: Multifunction Adaptive Processor for Small Antenna Arrays: *Institution of Electrical Engineers Proceedings*, volumen 130, partes F y H, n° 1, febrero 1983.
- 2 C. R. Ward y P. J. Hargrave: The Application of Sub-Optimal Control Methods to Adaptive Antennas for Airborne Communications Systems: *Institution of Electrical Engineers Conference Publication 195*, parte 1, 1981, págs. 174–178.
- 3 S. P. Applebaum: Adaptive Arrays: *Institute of Electrical and Electronics Engineers Transactions on Antennas and Propagation*, volumen AP-24, n° 5, septiembre 1976, págs. 585–598.
- 4 I. S. Reed, J. D. Mallett y L. E. Brennan: Rapid Convergence Rate in Adaptive Arrays: *Institute of Electrical and Electronics Engineers Transactions*, volumen AES-10, n° 6, noviembre 1974, págs. 853–863.
- 5 C. R. Ward, A. J. Robson, P. J. Hargrave y J. G. McWhirter: Application of a Systolic Array to Adaptive Beamforming: *Institution of Electrical Engineers Proceedings*, volumen 131, parte F, n° 6, octubre 1984.

Philip J. Hargrave se graduó BS en física por la Universidad de Bristol en 1971. Empezó investigaciones en radioastronomía en el Laboratorio Cavendish de Cambridge, y obtuvo el grado PhD en 1975. En 1977 ingresó en STL, trabajando en el diseño de receptores Navstar de elevadas prestaciones, estudio de sistemas adaptativos de antena, y modelos matemáticos de sistemas de defensa. Es miembro del IEE y socio de la Royal Astronomical Society. Actualmente dirige el laboratorio de sistemas adaptativos de STL.

Christopher R. Ward se graduó BS en ingeniería electrónica por el University College, Londres, 1975. Comenzó luego estudios de postgrado en dicho College y obtuvo el grado PhD en 1975 por sus investigaciones sobre técnicas de estrechamiento de redes mediante conformación de haz multiplicativa. Ingresó en STL en 1978, donde se ha dedicado a estudios teóricos y a simulación por ordenador de antenas adaptativas para receptores de navegación y comunicación en la banda L. El Dr. Ward es investigador principal del laboratorio de sistemas adaptativos de STL, y es miembro asociado del IEE.

Cálculo de la transformada discreta de Fourier en el procesador celular

El procesador celular es un ordenador del tipo de una sola instrucción para datos múltiples. Una aplicación es la evaluación de las transformadas discretas de Fourier mediante nuevos algoritmos que explotan el paralelismo en la arquitectura del procesador celular. Los tiempos de ejecución son mucho más cortos que para los procesadores matriciales de uso general.

J. M. Cotton

G. E. Masterson

ITT Advanced Technology Center, Shelton, Connecticut, Estados Unidos de América

Introducción

La TDF (transformada discreta de Fourier) juega un importante papel en el análisis, diseño y ejecución de algoritmos y sistemas de proceso digital de señales. Una razón de que el análisis de Fourier tenga tanta importancia es la existencia de muchos algoritmos eficaces para el cálculo de la TDF.

El procesador celular que se está desarrollando en el ITT Advanced Technology Center es una máquina de una sola instrucción para múltiples datos, adecuada para diversas aplicaciones. En esencia se trata de una matriz rectangular de elementos de cálculo binario que trabajan en cooperación

y pueden ejecutar simultáneamente una instrucción con múltiples operandos contenidos en la misma matriz. Potencialmente, esto da una ventaja en cuanto a capacidad de proceso sobre el ordenador convencional debida al número de operandos en la matriz, siempre que cada operando se utilice provechosamente en el cálculo. En otras palabras, la utilidad de una máquina de instrucción única para múltiples datos es máxima cuando está ejecutando un algoritmo con un alto grado de paralelismo. Pueden construirse algoritmos de cálculo de la TDF con el paralelismo y estructura deseados para que se ejecuten de modo eficiente en el procesador celular.

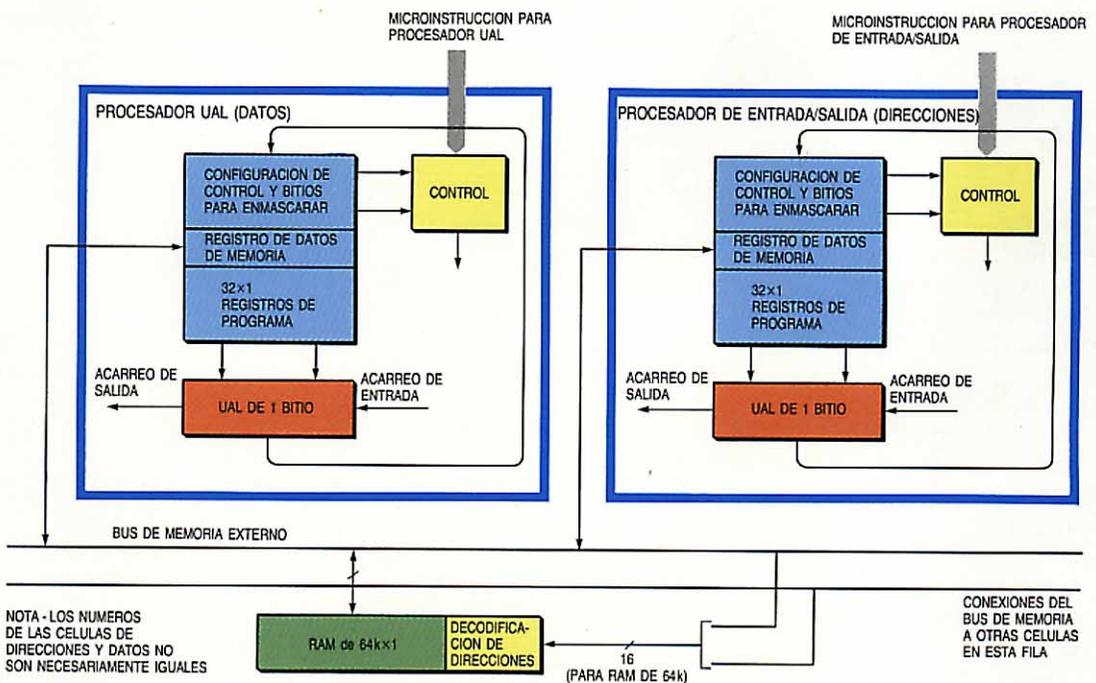


Figura 1
Diagrama de una célula UAL - unidad aritmética y lógica.

Características del procesador

El procesador celular ofrece un modo de proceso en el que un control secuencial único ejecuta simultáneamente el mismo cálculo sobre muchos elementos de datos. Cada célula es un procesador independiente de un bitio, con su propia unidad aritmética y lógica, un conjunto de 32 registros de 1 bitio, acceso directo a memoria, y memoria de datos de hasta 64 k-bitios. En la figura 1 se representa tal célula, apreciándose que el cálculo aritmético y la resolución de la dirección se pueden ejecutar en paralelo cuando convenga. El número de células de una matriz puede variar desde 256 hasta más de 16.000. La figura 2 muestra una matriz de 32 x 32 células y pone de relieve que todas las células obedecen simultáneamente a la misma instrucción.

La figura 3 muestra una fila de 16 células y sus interconexiones con las adyacentes y con un bloque de memoria de 64 k x 16 bitios, multiplexando las conexiones de memoria para datos y direcciones. Esta configuración es adecuada para aplicaciones de proceso de señal, aunque algunas situaciones especiales podrían exigir menos bitios de memoria por célula. Una fila completa de 32 bitios se puede direccionar como una sola unidad de memoria.

Las células están interconectadas en filas y columnas, con lo que las de una misma fila pueden procesar datos de múltiples bitios. El número de bitios en un elemento de datos se puede configurar dinámicamente: por ejemplo, un conjunto de 32 filas por 32 columnas puede trabajar con 1024 elementos lógicos separados, 256 números decimales codificados en binario, 128 pixels de 8 bitios, 32 elementos de datos de 32 bitios, o de cualquier otra manera, incluyendo formatos en coma flotante. Los elementos de datos no necesitan tener el mismo tamaño en todo el conjunto matricial: en particular, cuando se va a usar el conjunto para aplicaciones de bases de datos, los diferentes datos contenidos en un registro necesitarán márgenes de valores y, por tanto, tamaños distintos. Esta capacidad de adaptar el tamaño de los datos a su margen potencial de valores confiere gran eficacia al uso de las células y de la memoria asociada.

La identificación de los límites de un elemento de datos se logra poniendo códigos de control en las células apropiadas para determinar si corresponden al bitio más o menos significativo del elemento en cuestión. Es normal que los límites de los elementos ocupen las mismas posiciones de columna en cada fila, aunque ni siquiera esto es obligatorio.

Cuando se ha definido la posición de los

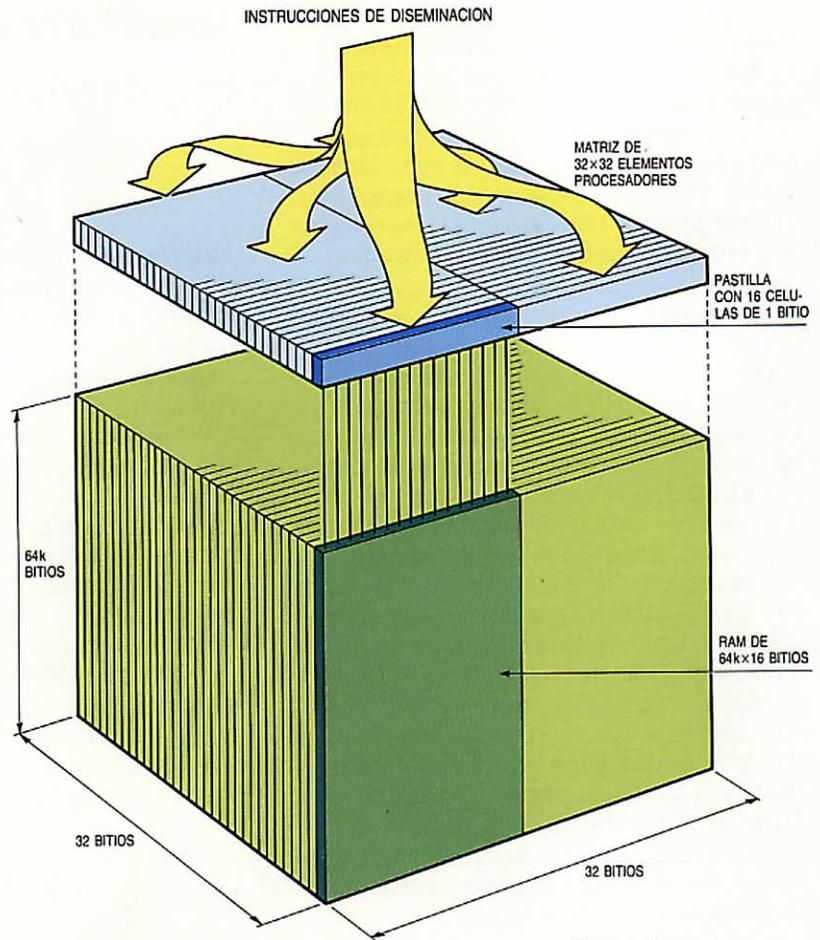
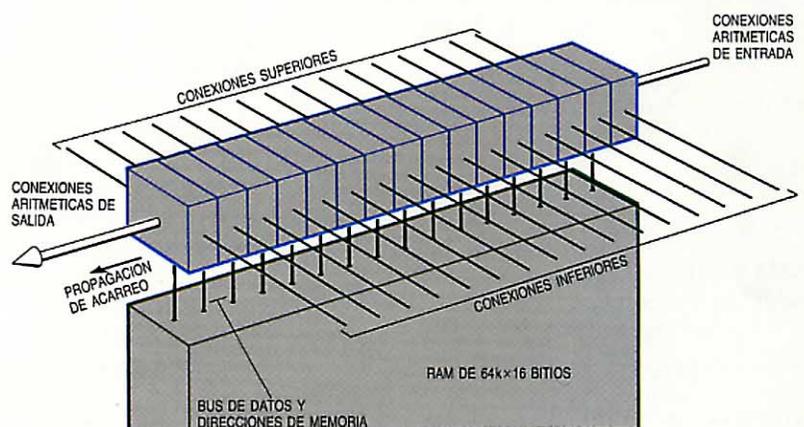


Figura 2
Configuración del sistema celular.

elementos de datos, y su tamaño, todos los contenidos de los registros de datos de las células comprendidas dentro de esa posición del elemento pueden haber realizado entre sí instrucciones aritméticas (o de desplazamiento, o booleanas), y tener todos los indicadores aritméticos normales (negativo, de estado [no-cero], desbordamiento, acarreo) activados en la célula del bitio más significativo. Una instrucción (de comparación) activa estos indicadores como si se hubiera ejecutado una sustracción entre los contenidos de dos registros, sin cambiarlos.

Figura 3
Fila de 16 células con memoria.



Aunque todas las células dentro de la matriz obedezcan simultáneamente a la misma instrucción (cada una a su manera, según la significación de su bitio), es posible inhibir o enmascarar una célula para que sólo ejecute las instrucciones condicionalmente. Esta máscara se puede poner o transferir desde otro bitio en la misma célula, o desde otra célula del elemento o fila de datos.

Los datos se pueden transferir, en dirección vertical, de una célula a la inmediata superior o inferior. Además, en cada columna hay un bus vertical mediante el cual se comunican todas las células de esa columna. En un instante cualquiera, sólo una célula puede transmitir al bus, mientras que las otras células de la columna reciben del mismo, sirviendo el mecanismo de máscara condicional para seleccionar cuál célula es la que va a transmitir.

Las direcciones para búsqueda de datos en las memorias de 64 k-bitios unidas a cada célula, se generan por filas, necesitando una sola dirección de 16 bitios para recoger nuevos datos de una determinada fila. Dicha dirección se posiciona en el extremo menos significativo de la fila. Al tener cada fila su propia dirección, la fila que esté activada por la máscara puede recoger sus propios datos de la memoria.

Transformada discreta de Fourier

La TDF se define como sigue:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} \tag{1}$$

donde

$$k = 0, 1, \dots, N-1$$

$$W_N = \exp(-2\pi j/N), \text{ y}$$

$X(k)$, $x(n)$ son, en general, números complejos.

Esta ecuación expresa la llamada transformada de análisis. La TDF inversa, o transformada de síntesis, viene dada por:

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn} \tag{2}$$

donde

$$n = 0, 1, \dots, N-1.$$

Esta pareja de transformadas tiene una relación matemática interesante por sí misma, sin darles un significado físico a $X(k)$ y $x(n)$.

Sin embargo, en proceso de señales, las $x(n)$ son, normalmente, secuencias de

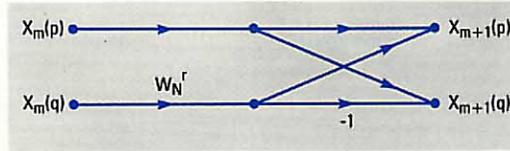


Figura 4
Grafo de mariposa.

duración finita, equiespaciadas en el tiempo, de amplitudes de la señal de interés, mientras que las $X(k)$ son las correspondientes frecuencias espectrales.

De la ecuación (1) se desprende que son necesarias N multiplicaciones complejas y $N-1$ sumas complejas para evaluar cada $X(k)$. Así, pues, la TDF completa requeriría del orden de N^2 multiplicaciones y sumas complejas. Esto puede suponer una enorme cantidad de cálculo para una N elevada. Afortunadamente, la exponencial compleja W_N^{kn} es simétrica y periódica, permitiendo notables economías en el algoritmo de cálculo. Como ejemplo, supongamos que el número de muestras N es par, en cuyo caso la secuencia se puede expresar como suma de las muestras pares más la suma de las muestras impares:

$$\begin{aligned} X(k) &= \sum_{r=0}^{(N/2)-1} x(2r) W_N^{2rk} + \sum_{r=0}^{(N/2)-1} x(2r+1) W_N^{(2r+1)k} \\ &= \sum_{r=0}^{(N/2)-1} x(2r) W_N^{2rk} + W_N^k \sum_{r=0}^{(N/2)-1} x(2r+1) W_N^{2rk} \end{aligned}$$

Dado que:

$$W_N^2 = \exp\left(-2j \frac{2\pi}{N}\right) = \exp\left(-j \frac{2\pi}{N/2}\right) = W_{N/2}$$

$$X(k) = \sum_{r=0}^{(N/2)-1} x(2r) W_{N/2}^{rk} + W_N^k \sum_{r=0}^{(N/2)-1} x(2r+1) W_{N/2}^{rk} \tag{3}$$

La secuencia, por tanto, se ha dividido en la suma de dos TDF de $N/2$ puntos. Para el caso importante en que N sea potencia de dos, este proceso se continúa hasta que $X(k)$ se exprese como suma ponderada de TDF de 2 puntos. Esta elección no es tan restrictiva como parece, ya que muchos problemas prácticos tienen la flexibilidad suficiente para poder alargar artificialmente una secuencia de señal con muestras de valor cero. El espectro resultante será entonces una versión interpolada del espectro no aumentado, y el cálculo se reduce aproximadamente, a $N/2 \log_2 N$ multiplicaciones y $N \log_2 N$ sumas. La descomposición de N , siendo N una potencia de dos o, de hecho, un producto arbitrario de enteros, se puede hacer de muchas maneras, que colectivamente se denominan transformadas rápidas de Fourier.

Construcción de una transformada rápida de Fourier

El resultado presentado en este artículo se obtuvo del análisis de una TDF de 64 puntos. Sin embargo, el método se explica utilizando una TDF de 8 puntos, ya que ello permite obtener los puntos importantes mediante diagramas relativamente directos. Las salidas de una TDF de dos puntos están relacionadas con sus entradas por ecuaciones de la forma:

$$X_{m+1}(p) = X_m(p) + W_N^r X_m(q) \quad (4)$$

$$X_{m+1}(q) = X_m(p) - W_N^r X_m(q) \quad (5)$$

Estas ecuaciones se pueden expresar convenientemente por el grafo de la figura 4, al que se le llama "mariposa". La descomposición sucesiva de una TDF de 8 puntos se puede ver en la figura 5. Este diagrama no es sólo expresión del algoritmo de la transformada rápida de Fourier, sino también el punto de partida para crear una estructura de cálculo que pueda ejecutarse eficientemente en el procesador celular. El primer punto a señalar en la figura 5 es que la secuencia de entrada $x(n)$ no está en un orden lineal simple, sino con el que corresponde a invertir sus bits. Así, por ejemplo, la muestra número seis, que en notación binaria es 110, está en la posición 011, o sea en la tercera posición, de la secuencia de entrada. Esto es consecuencia directa del modo de descomponerse la TDF en la transformada rápida de Fourier¹. Estas entradas se manipulan por parejas, de acuerdo con (4) y (5). En este cálculo de la mariposa, la exponencial compleja W_N^r se denomina "factor rotante". Como muestra la figura 5, después de cada etapa de cálculo con sus rotantes específicos, los operandos se reordenan en parejas antes de pasar a la nueva etapa con un conjunto diferente de factores rotantes.

Esto nos lleva al segundo punto importante. Si, al iniciarse los cálculos, los operandos están adecuadamente distribuidos en las filas del procesador celular, podrá entonces realizarse la primera etapa de cálculo. Sin embargo, en muchos casos, los operandos resultantes no estarán en las posiciones correctas, y por ello se requerirá una reordenación, o "barajado", entre las etapas de cálculo. Usando el grafo original de la figura 5 como modelo, se construye un segundo grafo en el que se han hecho los cambios apropiados entre las etapas de cálculo para posicionar correctamente los operandos en el procesador celular (Fig. 6).

Finalmente, se debe señalar que las ecuaciones (4) y (5) expresan relaciones entre cantidades complejas. Con los subí-

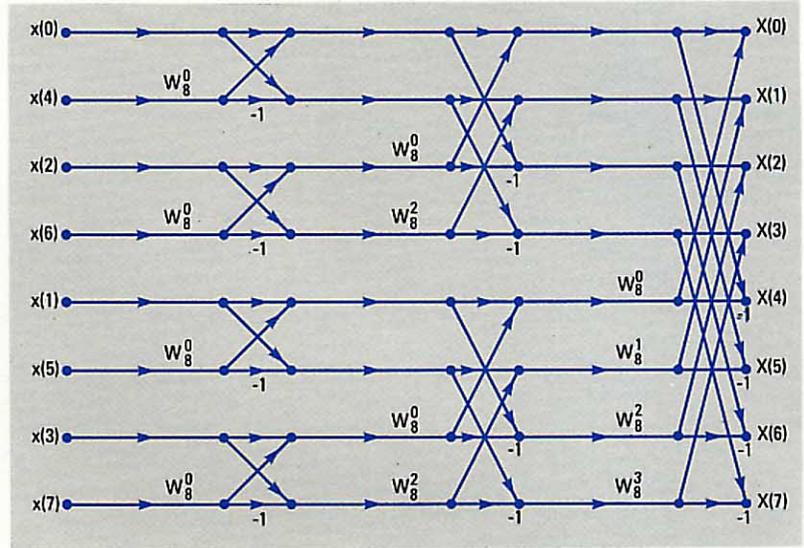


Figura 5 Grafo de una transformada discreta de Fourier de 8 puntos.

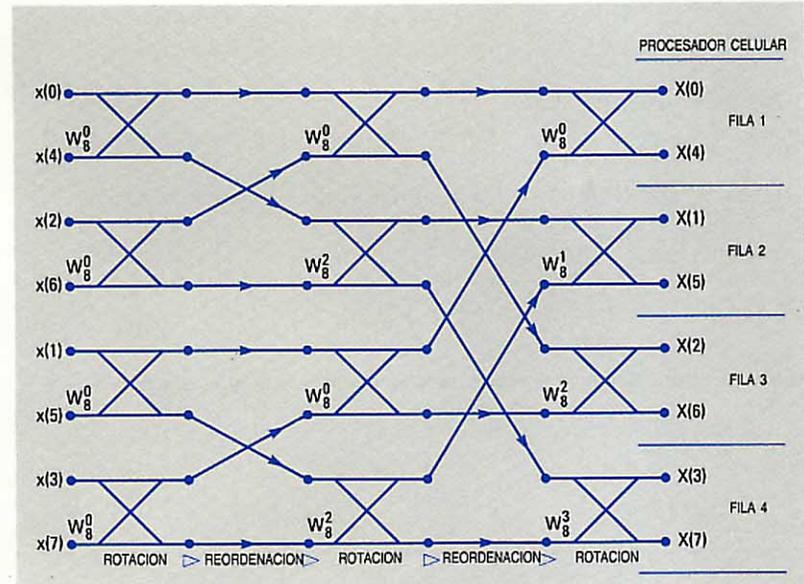


Figura 6 Grafo para el cálculo en un procesador celular.

indices R e I indicando las componentes real e imaginaria, la evaluación de una mariposa requiere cuatro cálculos:

$$X_{m+1}(p)_R = X_m(p)_R + [W_{NR}^r X_m(q)_R - W_{NI}^r X_m(q)_I] \quad (6)$$

$$X_{m+1}(p)_I = X_m(p)_I + [W_{NI}^r X_m(q)_R + W_{NR}^r X_m(q)_I] \quad (7)$$

$$X_{m+1}(q)_R = X_m(p)_R - [W_{NR}^r X_m(q)_R - W_{NI}^r X_m(q)_I] \quad (8)$$

$$X_{m+1}(q)_I = X_m(p)_I - [W_{NI}^r X_m(q)_R + W_{NR}^r X_m(q)_I] \quad (9)$$

Es posible simplificar algo lo anterior para valores concretos de r .

Para mayor sencillez, a $X_{m+1}(p)$ se le llama primer operando complejo, y a $X_{m+1}(q)$ segundo operando.

Tabla 1 — Tiempos de ejecución para TDF de distintos tamaños en un sistema 16 × 16

Número de puntos en la TDF	32	64	128	256	512	1024
Número de mariposas por etapa	16	32	64	128	256	512
Número de periodos T_1	1	2	4	8	16	32
Número de periodos T_2	—	2	8	24	64	160
Tiempo total de ejecución (μs)	50	120	280	640	1440	3200

El tiempo total de ejecución es la suma de los T_1 y los T_2 como se define en el apartado "Codificación del algoritmo".

Tabla 2 — Codificación y tiempos para la primera etapa de cálculo

Ciclos de máquina	Operación	Comentarios
2	V, LD R0, (I0) +	Valores reales de operandos de primera entrada en R0
2	V, LD R1, (I0) +	Valores imaginarios de operandos de primera entrada en R1
2	V, LD R2, (I0) +	Valores reales de operandos de segunda entrada en R2
2	V, LD R3, (I0) +	Valores imaginarios de operandos de segunda entrada en R3
1	V, LD R4, R0	Copia R0 en R4
1	V, LD R5, R1	Copia R1 en R5
1	V, ADD R0, R2	Valores reales de operandos de primera salida en R0
1	V, ADD R1, R3	Valores imaginarios de operandos de primera salida en R1
1	V, NEG R2, R2	
1	V, ADD R2, R4	Valores reales de operandos de segunda salida en R2
1	V, NEG R3, R3	
1	V, ADD R3, R5	Valores imaginarios de operandos de segunda salida en R3
16	TOTAL	

Para la primera etapa, todos los factores rotantes son iguales a la unidad.

Tabla 3— Codificación y tiempos para las siguientes etapas de cálculo

Para calcular ecuaciones de la forma $A = X + WY$, $B = X - WY$ donde $WY = (W_R Y_R - W_I Y_I) + j(W_I Y_R + W_R Y_I)$		
Ciclos de máquina	Operación	Comentarios
2	V, LD R30, (I2) +	Carga de factores rotantes reales
2	V, LD R31, (I2) +	Carga de factores rotantes imaginarios
1	V, LD R4, R2	
1	V, LD R5, R3	
18	V, MUL R4, R30	
18	V, MUL R5, R31	
1	V, SUB R4, R5	$(W_R Y_R - W_I Y_I)$ en R4
18	V, MUL R2, R31	
18	V, MUL R3, R30	
1	V, ADD R3, R2	$(W_I Y_R + W_R Y_I)$ en R3
1	V, LD R2, R0	
1	V, LD R5, R1	
1	V, ADD R0, R4	Valores reales de operandos de primera salida en R0
1	V, ADD R1, R3	Valores imaginarios de operandos de primera salida en R1
1	V, SUB R2, R4	Valores reales de operandos de segunda salida en R2
1	V, NEG R3, R3	
1	V, ADD R3, R5	Valores imaginarios de operandos de segunda salida en R3
87	TOTAL	

Análisis de TDF de 64 puntos y de tamaños superiores

Un sistema mínimo de procesadores celulares capaz de demostrar en la práctica las ventajas de esta máquina de instrucción única y datos múltiples, tendrá un tamaño 16×16 , y como mucho podrá evaluar una TDF de 32 puntos con operandos de 16 bits enteramente dentro del sistema celular. Sin embargo, se eligió una TDF de 64 puntos, ejecutándose en un sistema 16×16 para observar lo que sucede al evaluar una TDF de tamaño arbitrario. El proceso, que es general para TDF más grandes, consiste en cargar las 32 primeras muestras (en orden inverso de bits) en el sistema y proseguir el cálculo hasta la penúltima etapa, en cuyo momento los resultados se guardan en RAM (memoria de acceso aleatorio). Las 32 muestras restantes se procesan luego también hasta la penúltima etapa de cálculo, y se guardan. Finalmente, los datos de la RAM se vuelcan al sistema celular de acuerdo con los requisitos de la ordenación final, y se calculan los resultados.

Se han trazado diagramas equivalentes a las figuras 5 y 6 para una TDF de 64 puntos, obteniendo la conclusión de que el tiempo de ejecución es de $2T_1 + 2T_2$, ó $120 \mu s$, como se define en el siguiente apartado. Los resultados se extrapolan entonces a otros tamaños de TDF, según indica la tabla 1.

El método descrito aquí se presta al análisis de tiempos de ejecución de TDF en sistemas de tamaños distintos al de 16×16 . Así, si N es el número de parejas de operandos de 16 bits que se pueden manejar simultáneamente en un sistema celular (p.ej., 16 en un sistema 16×16), una TDF de $(2^{n+1} N)$ puntos se ejecutará en un tiempo:

$$2^n(T_1 + n T_2) \text{ para } n \geq 0.$$

Por ejemplo, una TDF de 1024 puntos aplicada a un sistema 32×32 ($N=64$ y $n=3$) se ejecuta en $8T_1 + 24T_2$, es decir, en $640 \mu s$.

Aunque el procesador celular no se diseñó específicamente para ejecutar TDF, su comportamiento supera al de procesadores matriciales más caros de uso general, como el SPS-21 que necesita 8,2 ms para ejecutar una TDF de 1024 puntos. Hay conjuntos de circuitos integrados especiales que pueden ejecutar TDF más aprisa que el procesador celular, pero son mucho menos flexibles y en su mayoría consumen demasiada energía.

Tabla 4 — Codificación y tiempos para la n -ésima reordenación dentro del conjunto

Ciclos de máquina	Operación	Comentarios
2	V, LD R31, (I1)+	Poner la máscara de desplazamiento hacia arriba
1	V, LD S0, R31	
2	V, LD R31, (I1)+	Poner la máscara de desplazamiento hacia abajo
1	V, LD S1, R31	
1	V, LD R4, R0	
1	V, LD R5, R1	
2^{n-1}	V, SHIFTU R4, # 2^{n-1}	Nuevos valores reales del segundo operando
1	V, LD [0] R2, R4	
2^{n-1}	V, SHIFTU R5, # 2^{n-1}	Nuevos valores imaginarios del segundo operando
1	V, LD [0] R3, R5	
1	V, LD R4, R2	Nuevos valores reales del primer operando
1	V, LD R5, R3	
2^{n-1}	V, SHIFTD R4, # 2^{n-1}	Nuevos valores reales del primer operando
1	V, LD [1] R0, R4	
2^{n-1}	V, SHIFTD R5, # 2^{n-1}	Nuevos valores imaginarios del primer operando
1	V, LD [1] R1, R5	
$14 + 2^{n+1}$	TOTAL	

Codificación del algoritmo

Los siguientes ejemplos se han codificado del modo descrito en la Ref^a. 2. Los factores rotantes de la primera etapa de cálculo son todos la unidad, con lo que el tiempo de cálculo es sólo de 16 ciclos de máquina, como se ve en el ejemplo de codificación de la tabla 2. Las etapas subsiguientes requieren calcular una expresión más compleja que ocupa 87 ciclos de máquina, según se indica en la tabla 3. Dichas etapas de cálculo deben ser reordenadas, lo cual para la etapa n -ésima, necesita $14 + 2^{n+1}$ ciclos de máquina (Tabla 4). Las cinco primeras etapas pueden desarrollarse sin tomar datos adicionales de la RAM, y ello requiere un tiempo T_1 de aproximadamente 50 μ s.

La evaluación de TDF con más de 32 puntos exige más de un recorrido a través del sistema celular para las cinco primeras etapas de cálculo; los resultados de cada paso se almacenan en RAM. Se vuelve luego a utilizar el sistema tantas veces como sea necesario para procesar los datos así almacenados. Cada paso a través del sistema consume aproximadamente un tiempo de 10 μ s, denominado T_2 .

Referencias

- 1 Alan V. Oppenheim y Ronald W. Schaffer: Digital Signal Processing: Prentice-Hall, Inc, 1975, pág. 298.
- 2 Robert W. Green: CAP Instruction Set Specification, Edición B: ATC Systems Research Division, mayo 1984.

John M. Cotton obtuvo su grado BSc en física especial en la Universidad de Londres en 1954, ingresando entonces en los Laboratorios de Electrónica Aplicada de GEC para trabajar en los primeros circuitos de aplicación de transistores y ordenadores digitales para control de aviones. Después de pasar por Computer Development Limited e International Computers Limited, en 1964 ingresó en Plessey Company, donde trabajó en diseño de ordenadores para telefonía. El Sr. Cotton ingresó en BNR (Canadá) en 1974 para trabajar en la familia de conmutación digital DMS, y dos años más tarde entró en el ITT Advanced Technology Center, siendo allí responsable de la arquitectura de la central digital Sistema 12 de ITT. En la actualidad, es director del grupo de arquitectura de sistemas avanzados en la División de Investigación de Sistemas.

Gregory E. Masterson se graduó BEng (Hons) en la Universidad de Liverpool, Inglaterra, 1949, e ingresó luego en el Laboratorio de Investigación Radar de GEC. Mientras trabajaba en la industria, consiguió el MEng por la Universidad de Liverpool, el MSc en matemáticas por la Universidad de Londres, y el PhD en ingeniería eléctrica por la Universidad de Pensilvania. Gran parte de su experiencia la ha conseguido en la industria de proceso de datos electrónicos. Cuando estuvo en Burroughs en 1958, fue el jefe de arquitecturas del D825. El Dr. Masterson entró en ITT como director técnico adjunto para ordenadores y sistemas en el staff técnico en 1974, y ahora es científico consultivo senior en el Advanced Technology Center.

Tendencias en procesadores digitales de señales programables de circuito integrado único

Los circuitos comerciales para proceso digital de señales difieren en un cierto número de características, por lo que hay que examinarlos con cuidado antes de seleccionar el que se destine a una aplicación concreta. En el futuro, los circuitos harán uso de la tecnología CMOS y de nuevas arquitecturas para mejorar su comportamiento.

K. N. Kneib

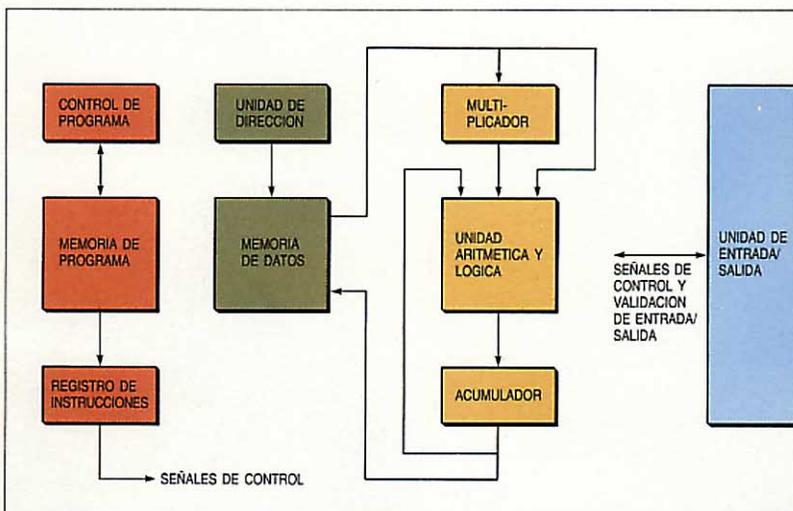
ITT Defense Communications Division,
San Diego, California, Estados Unidos
de América

Introducción

Desde el final de los 70, existen en el mercado de EE.UU. varios PDS (procesadores digitales de señales) programables, realizados en un solo circuito integrado. Los dispositivos introducidos durante este periodo ilustran las tendencias actuales en las arquitecturas y funciones especiales de los PDS, resultado del progreso continuo en la tecnología de circuitos integrados y la realización de nuevas aplicaciones para tales procesadores.

Los objetivos en el campo del proceso de señal se dirigen hacia el aumento del rendimiento y la reducción de la cantidad de circuitos; esto último disminuye la disipación y el tamaño del equipo. Tales objetivos son asequibles merced a la introducción de circuitos integrados de uso general para el proceso digital de señales.

Figura 1
Asociación de funciones en un procesador digital de señales con arquitectura basada en acumulador.



Historia del desarrollo

Históricamente, ha sido más frecuente el uso de técnicas analógicas en el proceso de señales, si bien las técnicas digitales de proceso de señal llevan varias décadas utilizándose para algunas aplicaciones. Con estas últimas se consigue mayor exactitud y estabilidad que con los circuitos analógicos y, además, la introducción de circuitos VLSI digitales a precios económicos, hace que se puedan utilizar en una gama de aplicaciones mucho más amplia. La precisión exigida en el campo del proceso y control numérico, determinan que una realización analógica resulte cara por los elevados costes de los componentes analógicos necesarios.

Una referencia para evaluar el tratamiento digital de señales es la realización de un producto escalar (es decir, la suma de productos) en menos de $0,5 \mu\text{s}$ por término, con una precisión en los datos de más de 16 bits (preferiblemente 24 ó 32 bits). El extenso proceso aritmético y la manipulación de datos se deben ejecutar en tiempo real.

La figura 1 es un diagrama de bloques de un procesador digital de señales: la disposición funcional optimiza el rendimiento del proceso de señal. Una unidad aritmética de datos, formada por un multiplicador, una unidad aritmética y lógica y un acumulador, ejecuta el intensivo proceso aritmético. Otras unidades aritméticas se dedican a calcular las direcciones de los datos y al control del programa. Las memorias de datos y de programa se encuentran en campos de dirección separados. Un registro de instrucciones, asociado con la memoria de programa, permite extraer las instrucciones antes de su ejecución, con lo que se aumenta el rendimiento.

Tabla 1 — Principales características de seis procesadores digitales de señales en circuito integrado único

	Procesador A Intel 2920	Procesador B AMI S28211	Procesador C NEC uPD7720	Procesador D TI TMS32010	Procesador E Fujitsu MB8764	Procesador F STC DSP 128
Fecha de introducción	1979	1984 (S2811-1980)	1981	1983	1984	1985
Tiempo de ciclo de máquina (ns)	400	300	250	200	100	320
Tecnología de fabricación	HMOS	3- μ m NMOS	3- μ m E/D NMOS	3- μ m SMOS	CMOS	3- μ m NMOS
Máxima disipación de potencia (W)	0,8	0,8	0,9	0,9	0,3 (100% ciclo de servicio)	1,3
Encapsulado	DIP 28 term.	DIP 28 term.	DIP 28 term.	DIP 40 term.	Enrejillado de 88 term.	Portador de circuito JEDEC de 68 contactos

E/D - empobrecido en electrones

DIP - encapsulado en doble fila

MOS - semiconductor metal-óxido

NMOS - MOS canal N

SMOS - NMOS escalado

CMOS - MOS complementario

HMOS - NMOS de alta prestación

VMOS - NMOS de canal V

JEDEC - Joint Engineering Design Electronics Council

Antes de finalizar los años 70, los sistemas de tratamiento digital de señales contenían muchos componentes con circuitos integrados, entre los que figuraban un multiplicador o multiplicador-acumulador, una unidad aritmético-lógica, una amplia memoria de programa, memoria de datos, registros o ficheros de registro, y controladores. Los avances en la tecnología han hecho posible incorporar muchas de estas funciones en un solo circuito integrado programable.

La selección de las funciones a integrar depende de la tecnología, de las aplicaciones previstas y del objetivo de coste del circuito. Los dispositivos de uso general en una sola pastilla VLSI se están utilizando ahora en muchas aplicaciones, como codificación de voz, módems, reconocedores de voz, analizadores de espectro y filtros digitales. El uso de PDS de una sola pastilla reduce el número de componentes, siendo adecuados por su programabilidad para un extenso campo de aplicaciones.

Procesadores digitales de señales disponibles comercialmente

En la tabla 1 se comparan algunas de las principales características de los PDS existentes en el mercado, cuyas fechas de introducción reflejan ciertas tendencias.

El procesador A, el primero que apareció es el único que tiene conversión analógica dentro del propio circuito integrado. El procesador B vino a continuación, pero nunca cumplió las especificaciones, debido a problemas asociados con la tecnología VMOS. En 1984 se tuvieron muestras del circuito que lo reemplazó, el S28211. Los

procesadores C y D han sido los PDS más populares en los dos últimos años y existen en cantidades suficientes para producción, ambos con suministradores alternativos.

El procesador E es el primer PDS realizado en CMOS. Consume sólo un tercio de la potencia de los otros procesadores, si bien por su gran superficie (93 mm²) podría tener menores rendimientos de producción, afectando negativamente a la disponibilidad e incrementando el coste del circuito.

Un nuevo circuito, del que se tienen muestras desde primeros de 1985, es el procesador F, un procesador de señal en tiempo real, integrado y conectable en cascada. El desarrollo se ha llevado a cabo en Standard Telecommunication Laboratories y se está fabricando en STC Semiconductors.

La duración de los ciclos de máquina ha ido disminuyendo a medida que pasa el tiempo, desde los 400 ns del procesador A a los 100 ns en el procesador E. Estos tiempos de ciclo se han aprovechado de los avances en tecnología de fabricación, entre ellos la mayor finura de pistas tanto en las tecnologías NMOS como en las CMOS. Sin embargo, la capacidad del procesador depende también del tamaño de la memoria, de las posibilidades de direccionamiento, de la cantidad de paralelismo y solapamiento en los procesos, y del repertorio de instrucciones.

Características de los procesadores digitales de señales

En la tabla 2 se resumen las características estructurales de los seis dispositivos. Las figuras 2 y 3 son diagramas de bloque de los procesadores F y E, respectivamente.

Tabla 2 — Características de la arquitectura de PDS programables en circuito único

	Procesador A Intel 2920	Procesador B AMI S28211	Procesador C NEC uPD7720	Procesador D TI TMS32010	Procesador E Fujitsu MB8764	Procesador F STC DSP128
Memoria						
Interna						
ROM de programa	192×24	512×18	512×23	1536×16	1024×24	0
RAM de datos	40×25	Dos (128×16)	128×16	144×16	256×16	Dos (256×16)
ROM de datos	24×25	128×16	510×13	En la ROM de programa	En la ROM de programa	No
Direccionamiento						
Indexado (incremento/decremento)	No	Sí	Sí	Sí	Sí	Sí
Indexado total (con registro)	No	No	No	No	Sí	Sí
Búsqueda en la tabla	No	Sí	Sí	Sí	Sí	Sí
Desplazamiento de datos (Z^{-1})	No	Sí	Sí	Sí	Sí	Sí
Acceso directo de la memoria externa						
Programa	No	No	No	Hasta 4K×16	Hasta 1K×24	Hasta 64 K×8
Datos	No	No	No	No	Hasta 1K×16	ROM de datos sólo en la ROM de programa
Aritmética						
Multiplicador						
E × E-S (bitios)	9×25–25 (sin multiplicador)	12×12–16	16×16–31	16×16–32	16×16–26	16×16–32
Tiempo para multiplicar-acumular en un bucle interno	aprox. 5 μs	300 ns	250 ns	400 ns	100 ns	320 ns
Unidad aritmético-lógica	Sí	Sólo unidad de suma/resta	Sí	Sí	Sí	Sí Unidades separadas
Acumulador						
Anchura (bitios)	28	16	Dos 16	32	26	35
Modo de saturación	Por circuito físico	Por circuito físico	Por programa	Por circuito físico	Por circuito físico	Por circuito físico y por programa
Entrada/salida						
E/S de datos en paralelo (bitios)	4 E/8 S	8	8	16	16	8
Interrupción por soporte físico	No	No	1	1	No	No
Analógicas en el circuito	9 bitios	No	No	No	No	No
Puerto en serie (frecuencia)	No	Sí (3 MHz)	Sí (200 kHz)	No	No	No
Acceso directo a memoria (transferencia de bloques)	No	Sí	Sí	No	Sí	Sí
Control por programa						
Longitud de la palabra de instrucción	24	18	23	16	24	8
Tamaños de la pila	No	1	4	4	2	8
Contador de bucles	No	Sí	No	Sí	Sí	Sí
Salto condicional	No	Sí	Sí	Sí	Sí	Sí

Arquitecturas

Estos PDS en un solo circuito están basados en la arquitectura Harvard, que es una modificación de la von Neumann. Las operaciones ejecutadas en una arquitectura von Neumann clásica son secuenciales, ya que los programas y los datos comparten el mismo bus; así, por ejemplo, la búsqueda

de instrucción, la búsqueda de datos, la operación con ellos y el almacenamiento de resultados, forman una secuencia repetitiva. Para aumentar la velocidad en el tratamiento digital de señales, se modifica esta arquitectura de modo que haya almacenamientos separados de programa y datos, así como buses independientes para trans-

Figura 2
Diagrama de bloques del procesador digital de señales DSP128 de STC (procesador F).

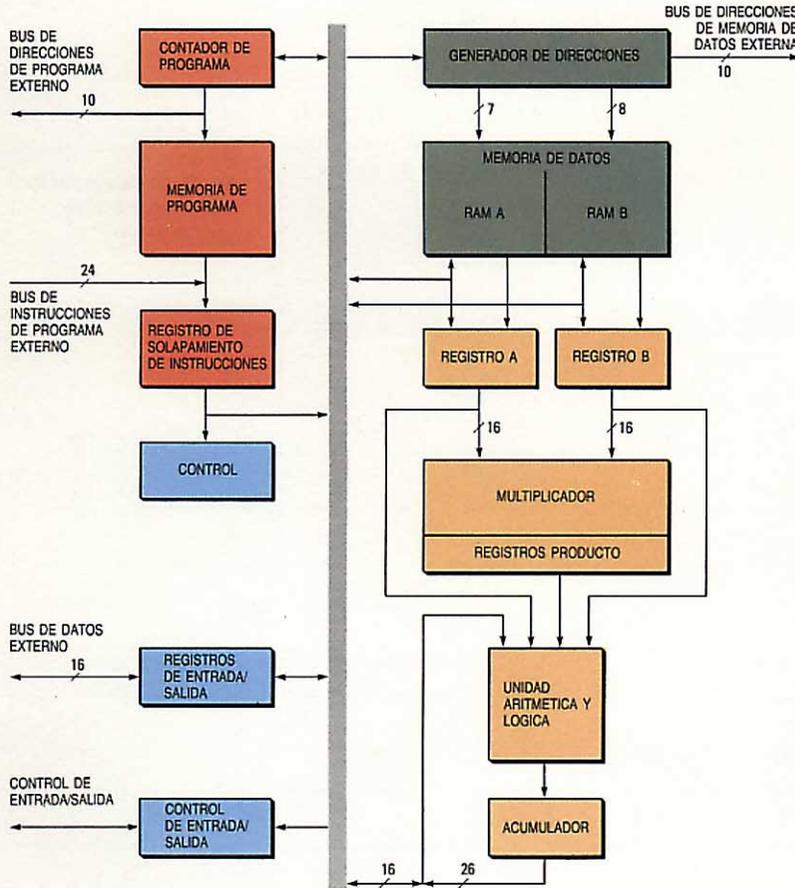
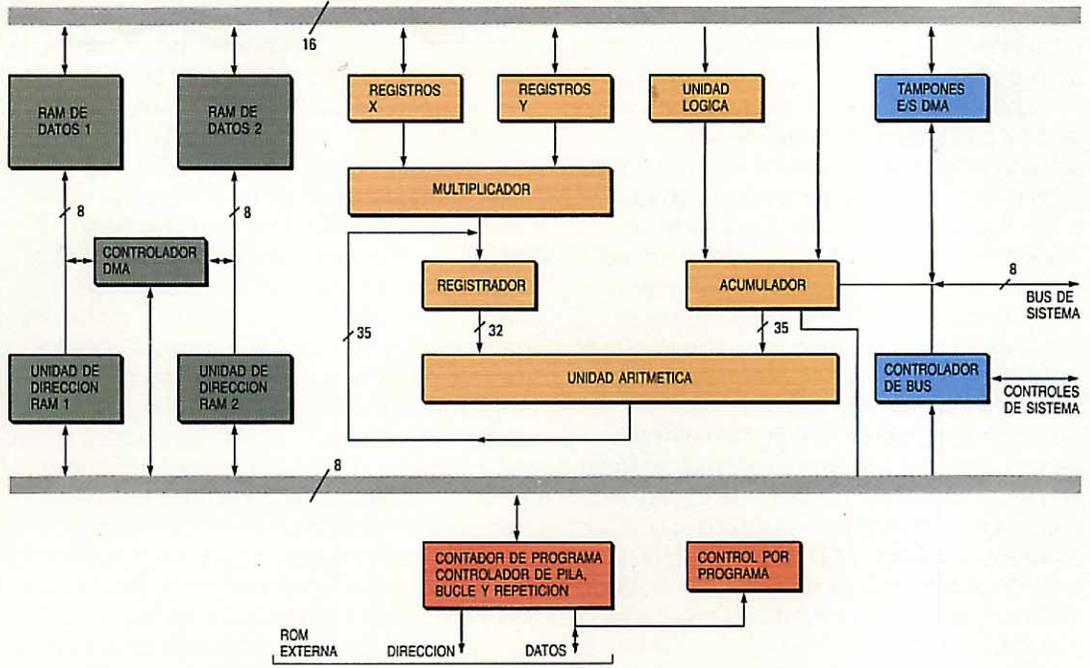


Figura 3
Diagrama de bloques del procesador digital de señales MB8764 de Fujitsu (procesador E).

ferencias y control de datos e instrucciones. Así, la búsqueda de programa puede ser simultánea a la ejecución de la instrucción anterior.

Puede haber más solapamientos dentro del proceso funcional del PDS, permitiéndole así ejecutar todo un conjunto de opera-

ciones con un resultado en cada ciclo de máquina. Sin embargo, la búsqueda de instrucciones, las operaciones con datos, y el almacenamiento de los mismos están gobernados por diferentes fases de tiempo de la secuencia del programa. El solapamiento tiene muchas ventajas para las funciones de proceso digital de señales, tales como el filtrado, en que la secuencia de programa es un bucle repetitivo. Una vez completado el solapamiento de procesos, se ejecuta una operación en cada ciclo de máquina. Sin embargo, una de tales operaciones aislada puede necesitar varios ciclos de máquina.

La figura 4 muestra un ejemplo para el caso de operaciones de multiplicación-acumulación de los procesadores D y E. El ejemplo compara una sola operación de ese tipo con la del bucle interno de un producto escalar. Se puede ver que el procesador D, cuya arquitectura tiene un solapamiento limitado, dobla la velocidad de salida en un bucle repetitivo, mientras que la arquitectura con elevado solapamiento del procesador E logra una mejora de cinco veces en dicha velocidad.

Se han introducido también etapas de paralelismo en las arquitecturas de los PDS de una sola pastilla, con el fin de incrementar aún más la velocidad, principalmente en las unidades aritméticas y en la zona de generación de direcciones.

El procesador F tiene tres buses internos separados que permiten realizar varias operaciones de proceso en paralelo, con un efecto neto en el aumento de rendimiento. Estos buses permiten ejecutar en paralelo operaciones de búsqueda de instrucción, movimiento de datos y de entrada/salida.

Memoria

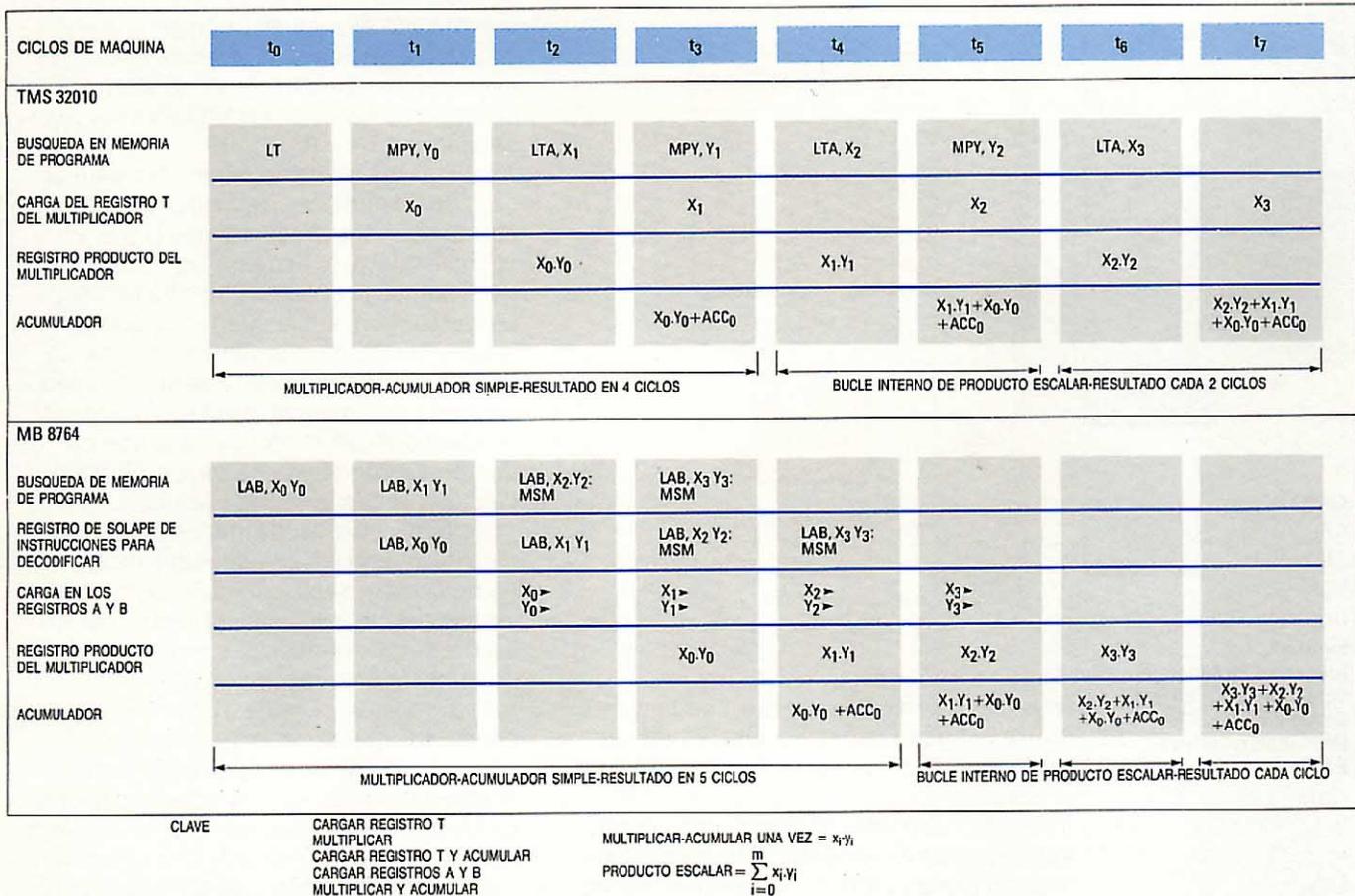
Los primeros circuitos tenían sólo memoria de programa interna en la pastilla, en la forma de ROM/EPROM, limitada a menos de 512 palabras de instrucción; no era posible utilizar una memoria externa. Para rutinas como las que se encuentran en análisis y síntesis de voz y módems de velocidad media a elevada, la longitud del programa que podía almacenar el dispositivo era menor que la del programa que podía ejecutar en tiempo real, con lo que el tamaño de la memoria de programa limitaba la capacidad del procesador.

Los procesadores D y E pueden trabajar con memoria de programa externa, además de tener, al menos, el doble de capacidad en palabras de instrucción internas que los primeros circuitos. El DSP128 de STC (procesador F) trabaja sólo con memoria externa y no tiene memoria de programa interna. El número de bits de la palabra de instrucción varía de 8 a 24. Operando con memoria externa, cuanto mayor sea el tamaño de la palabra, mayor será el número de circuitos de memoria externa necesarios. Además, siendo larga la palabra de instrucción, habrá operaciones que se ejecuten en un solo ciclo de máquina y que necesitarían dos ciclos con palabras de instrucción más cortas.

Al haberse reducido las dimensiones características, se integra más memoria de datos en el PDS, la cual consta de ROM y de RAM. La ROM de datos de los procesadores A, B y C sirve para almacenar los coeficientes utilizados en el proceso (p.ej., en funciones de filtrado). En las arquitecturas Harvard modificadas de los procesadores D, E y F no se necesita ROM de datos, ya que los coeficientes se pueden leer desde la memoria de programa. Esto facilita también la búsqueda en tablas, que suele formar parte de las rutinas de proceso digital de señales. En cuanto a la RAM para memoria de datos, puede llegar hasta 512 palabras. Sólo el procesador E puede acceder directamente a la memoria de datos externa, si bien el procesador D puede también hacerlo asignando la memoria de datos y el registro externo de direcciones a puertos E/S (entrada/salida), necesitando por lo tanto circuitos adicionales para decodificar el puerto E/S y el registro de direcciones.

La memoria de datos interna adopta diversas formas. El procesador A utiliza una RAM de tres puertos en la que se puede acceder al mismo tiempo a dos palabras de datos, localizadas en la misma dirección en dos bloques separados de la RAM. También se pueden escribir datos en una de estas

Figura 4
Ejemplos de ejecución de multiplicación-acumulación.



posiciones. STC ha adoptado también el diseño internacional de RAM, que consta de dos memorias independientes de 256 palabras. El procesador E proporciona dos RAM separadas de 128 palabras, que pueden asimismo utilizarse como RAM única de 256.

Muchos de los dispositivos tienen características de direccionamiento para memoria de datos que pueden ser activadas simultáneamente por instrucciones aritméticas, mejorando la capacidad de proceso al operarse en paralelo. La modificación de direcciones, además de la carga, almacenamiento, incremento y decremento de registros, se utiliza en rutinas de proceso de señal más complejas que las sencillas rutinas de filtrado. Esto lo puede hacer una unidad aritmética separada, dedicada a los registros de direcciones, que permite una total indexación de dichos registros. El circuito de STC incluye una unidad de direcciones RAM para cada uno de los bloques independientes de RAM internos, conteniendo cada unidad cuatro registros de direcciones base, un registro puntero y un registro de longitud vectorial, así como un sumador. Al refinarse las características de proceso de un PDS, este elemento se potenciará por la adición de sustractores y funciones lógicas.

Una función especial, utilizada en filtrado y en operaciones de correlación y convolución, es el desplazamiento de los datos en un lugar, dentro de un conjunto, para conseguir un retardo equivalente a una muestra. Esta función, la Z^{-1} , la poseen todos los PDS enumerados, con la excepción del procesador A.

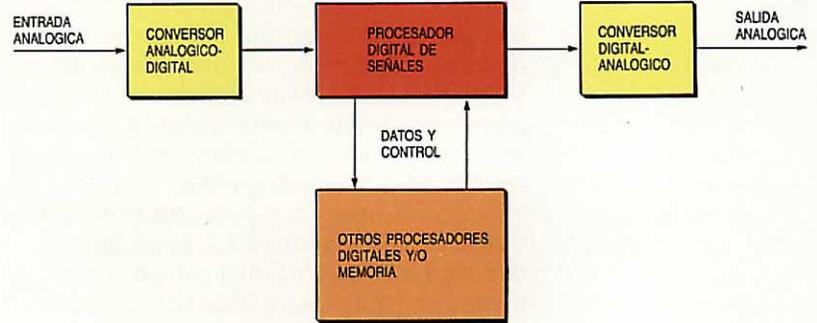
Aritmética

La unidad aritmética del PDS está formada por un multiplicador, una unidad aritmética y lógica (UAL), y un acumulador. Como el multiplicador del procesador A sigue el método de desplazamiento y suma, su tiempo de multiplicación y acumulación (aproximadamente $5 \mu s$) es lento, comparado con los multiplicadores discretos de los otros PDS que se ejecutan en uno o dos ciclos de máquina. Para tratamiento de señales es preferible un multiplicador de 16×16 bits que dé un producto de 32 bits, y sólo los procesadores F y D poseen un multiplicador global de doble precisión. En cuanto a la utilización del área de silicio, parece que se ha llegado a un compromiso entre el tamaño de la RAM de datos y la precisión del multiplicador en diseños de PDS en un solo circuito. También es muy deseable la aritmética de doble precisión en la UAL y en el acumulador. Salvo el procesador B, todos los PDS están provistos de UAL. Las operaciones booleanas admitidas

por la unidad lógica aportan algunas funciones de manipulación de bitio que constituyen una faceta importante de las operaciones de control.

En el tratamiento digital de señales, la excesiva sobrecarga de la aritmética puede producir grandes variaciones en el comportamiento del sistema debidas a cambios súbitos entre valores grandes y pequeños de los números. Para acometer de un modo

Figura 5
Una aplicación genérica de un sistema PDS.



digital este problema, característico de los sistemas analógicos, se han equipado protecciones de circuitos o de programación contra la sobrecarga o la aritmética de acumulaciones en modo de saturación. Sin embargo, es necesario que el programador comprenda el proceso físico involucrado y que sea prudente eligiendo la representación numérica a utilizar.

Entrada/salida

Hay que conceder importancia al tratamiento de la E/S. Todos los procesadores la incorporan en paralelo: desde los primeros, de 8 bits, hasta los de 16 bits, más modernos, con la excepción del PDS de STC. Una E/S en paralelo más amplia aumenta el número de terminales del circuito integrado, pero maneja con mayor eficacia que los buses más pequeños frecuencias de datos de E/S más elevadas.

Se pensó que la velocidad de E/S — para introducir nuevos datos a procesar o extraer los datos ya procesados — sería pequeña con respecto a la velocidad de los cálculos aritméticos internos. Esta teoría se basaba en la premisa de que el PDS tendría unos interfaces bastante primitivos con cualquier otra parte del sistema digital. Sin embargo, dicha velocidad de E/S se ha elevado más al añadir al circuito la nueva posibilidad de trabajar con memoria externa.

Ahora está claro que muchos sistemas en las aplicaciones de módems, análisis y síntesis de voz, y reconocimiento de voz necesitan utilizar los PDS en un entorno de multiproceso. Los requisitos de la aplicación se dividen de modo que las funciones de

proceso se distribuyan entre varios procesadores de un solo circuito, ya sean otros PDS, microprocesadores, y/o microordenadores, como se ve en la figura 5, que presenta el problema de diseño de mantener el movimiento de los datos y el control sin afectar mucho a las prestaciones del procesador de señal.

En los procesadores B y C hay unos puertos serie que primordialmente tratan una E/S analógica, pero que pueden utilizarse para comunicación de datos entre procesadores. Los otros procesadores también tienen controles para acceso directo a memoria en la transferencia de bloques de datos. El de STC ofrece mejores características de acceso directo a memoria entre el bus del sistema y las RAM internas, realizando el direccionamiento de bits invertidos necesario para las transformadas rápidas de Fourier de raíz 2. El acceso directo a la RAM está multiplexado en el tiempo con el acceso a RAM para operaciones internas. El procesador D tiene puertos de E/S para comunicación con otros dispositivos; sin embargo, las facilidades hasta ahora incorporadas a los PDS no permiten una comunicación entre procesadores como la que exigen los diseñadores de sistemas. Sólo con gran cantidad de circuitos externos y/o desaprovechando el tiempo de proceso, pueden utilizarse estos dispositivos en sistemas multiproceso.

La solución a este problema requiere a la vez unas velocidades de datos más elevadas en E/S, junto con señales de control para el interfaz y sincronización de los múltiples procesadores, a fin de interferir lo mínimo posible sobre los resultados del proceso de señal. El DSP 128 de STC ha abordado este problema, incorporando un bus de sistema con algunas señales de control que permiten al circuito funcionar como procesador maestro o esclavo y entenderse con dispositivos más lentos. Estas operaciones de E/S se hacen en paralelo con las operaciones internas.

La mayoría de los sistemas de proceso de señal están dirigidos por sucesos que ocurren en tiempo real, y a menudo fuera del PDS, exigiéndole a éste que responda a dichos sucesos. Las interrupciones por circuitos físicos proporcionan este medio de señalización, más eficaz para un procesador aritmético intensivo que la exploración secuencial, capaz de perturbar los cálculos en curso. Los procesadores C y D utilizan una única línea de interrupción. Sin embargo, las arquitecturas de procesadores múltiples suelen requerir más de una interrupción y por ello, desde la perspectiva del sistema, sería deseable tener diversas líneas de interrupción en los futuros PDS para reducir los circuitos externos.

El único PDS que incorpora funciones analógicas en la propia pastilla es el procesador A. Parece que el criterio general es el de ejecutar externamente la conversión analógico-digital, y ello se debe a que la combinación de circuitos analógicos y digitales en un mismo circuito integrado es un avance tecnológico reciente. Otra razón es la diferencia en las especificaciones impuestas a los conversores analógico-digital, como el número de bits (8, 12 ó 16), margen dinámico y otras características distintas según la aplicación.

Se ha estimado que el área total del circuito integrado de un PDS tendría que duplicarse para ofrecer un interfaz analógico con facilidades equiparables a las del PDS. Por tanto, para un procesador de señal programable de uso general, parece preferible utilizar un conversor analógico-digital externo.

Control por programas

El conjunto óptimo de instrucciones para un PDS programable debería minimizar el número de instrucciones y la complejidad de las operaciones. Por supuesto, el objetivo es permitir una elevada velocidad de cálculo y de ejecución de programa.

La ejecución no es necesariamente más rápida por ser larga la palabra de instrucción; ésta se ha concebido para poder ejecutar varias operaciones en paralelo dentro de una instrucción. Hay dos tipos de instrucción posibles: en un extremo, que cada bitio del código de operación de la palabra controle directamente una señal necesaria en el procesador y, en el otro, que cada palabra de instrucción implique una operación particular. El primero conduce a códigos de operación más largos, como los de los procesadores A y C, mientras que los procesadores D y F emplean el segundo criterio.

Un ejemplo para el procesador A es la instrucción "LTD", asociada con la ejecución de un producto escalar, la cual almacena en el acumulador la suma de la última multiplicación y del acumulador, carga el multiplicador con un nuevo valor y ejecuta la operación Z^{-1} en la memoria de datos. El procesador realiza operaciones especializadas que mejoran el rendimiento del sistema de proceso de señal, pero esta técnica exige mayor trabajo al diseñador del circuito, ya que un conjunto de instrucciones verticales requiere más decodificación que si fuera horizontal.

En el tratamiento digital de señales, una gran parte del tiempo de ejecución de las rutinas de un algoritmo se gasta en el código para el bucle interno. El paralelismo funcional en el circuito físico permite que se optimice el control del programa. Los conta-

dores de bucle y las instrucciones de repetición, como las que tienen los procesadores F y B, mejoran la ejecución de los bucles internos.

Se incorporan pilas de programas para poder utilizar eficazmente dichos programas y manejar bien las interrupciones. Con la capacidad de expansión a memoria externa, hay que aumentar el tamaño de la pila para que a los programadores les resulte ventajosa la programación estructurada, mediante subrutinas anidadas. La ausencia del proceso condicional es otra limitación, ya superada en los nuevos procesadores por STC y otros. Cuando las palabras de instrucción son cortas, estas funciones se describen en dos palabras, dado que no es crítico el tiempo de ejecución de este tipo de instrucciones.

Herramientas de desarrollo

Como en todos los dispositivos programables, los sistemas para desarrollo de equipos y programas son importantes cuando se utiliza un circuito en una aplicación. Se han introducido herramientas soporte de la programación que se pueden ejecutar en diversos ordenadores personales, minordenadores u ordenadores principales, con diferentes sistemas operativos como VMS, MS-DOS y Unix.

Los programas de soporte incluyen un ensamblador y montador, y un simulador. El ensamblador y el montador producen módulos de programa en una forma adecuada para la carga en el simulador o en el procesador físico. El simulador permite desarrollar y depurar el código fuera del procesador, y existe un emulador que da medios para depuración del circuito electrónico dentro del sistema y en tiempo real. De igual modo, los dispositivos que trabajan con memoria externa permiten al diseñador construir su propio sistema, utilizando memoria de programa RAM para el desarro-

llo, y reemplazándola después por una ROM, EPROM o PROM para la fase de producción.

Tendencias futuras

Los avances tecnológicos en diseño y técnicas de fabricación beneficiarán a las próximas generaciones de PDS en un solo circuito integrado. En el futuro, CMOS será la tecnología de proceso de los circuitos con elevadas prestaciones. Como se ha mencionado antes, en los diseños actuales se prevé mejorar las facilidades de direccionamiento de datos, aumentar el espacio de memoria, y potenciar el multiproceso de entrada/salida.

Son también muy probables los avances en las arquitecturas básicas de PDS en circuito integrado único. La arquitectura utilizada en los últimos PDS, con un solo acumulador, cederá paso a un fichero de registro, o algo similar, que sea adecuado para aplicaciones de redes de procesadores. A este concepto se vinculará la tendencia a utilizar aritmética en coma flotante de precisión total.

Cambios como los descritos permitirán expandir enormemente las aplicaciones de los procesadores digitales de señales realizados en circuito integrado único.

Kristine N. Kneib se graduó BSc en matemáticas y BSc en física en la Universidad de Queensland, Australia, en 1968 y 1971, respectivamente. En 1976 obtuvo un PhD en física aplicada en la Universidad de California, San Diego. En 1979 ingresó en ITT Defense Communications Division, donde ha impulsado técnicamente las áreas de proceso avanzado y futuros terminales de banda estrecha con secrafonía. La Dra. Kneib ha contribuido de un modo destacado a la definición y desarrollo de arquitecturas y a la selección de tecnologías. Ella dirige técnicamente por parte de ITT el desarrollo conjunto de una segunda generación de procesadores digitales de señales.

Procesador digital de señales para aplicaciones de telecomunicación

El analizador de señales de prueba es un procesador digital de señales orientado a la búsqueda de averías, programable remotamente para ser utilizado en la central digital Sistema 12. Su arquitectura de aplicación general y su conjunto de instrucciones orientadas al usuario han demostrado también su eficacia en muchas otras aplicaciones.

S. Das
S. Jayanthi
F. A. Middleton
S. G. Morton

ITT Advanced Technology Center, Shelton, Connecticut, Estados Unidos de América

Introducción

Por su estructura de diseño flexible y modular, el Sistema 12 permite la evolución de la red telefónica hacia un sistema enteramente integrado de distribución de información, capaz de tratar simultáneamente voz y datos. Siguiendo esta orientación, se ha desarrollado un sistema de prueba de líneas y enlaces con parecida flexibilidad, acomodable a los avances tecnológicos.

La prueba de líneas y enlaces en una central digital requiere nuevos procedimientos, ya que no es posible establecer una conexión metálica a través de la central con una posición de prueba, sobre todo si ésta es remota. En el Sistema 12 se

resuelve el problema creando una original arquitectura de pruebas.

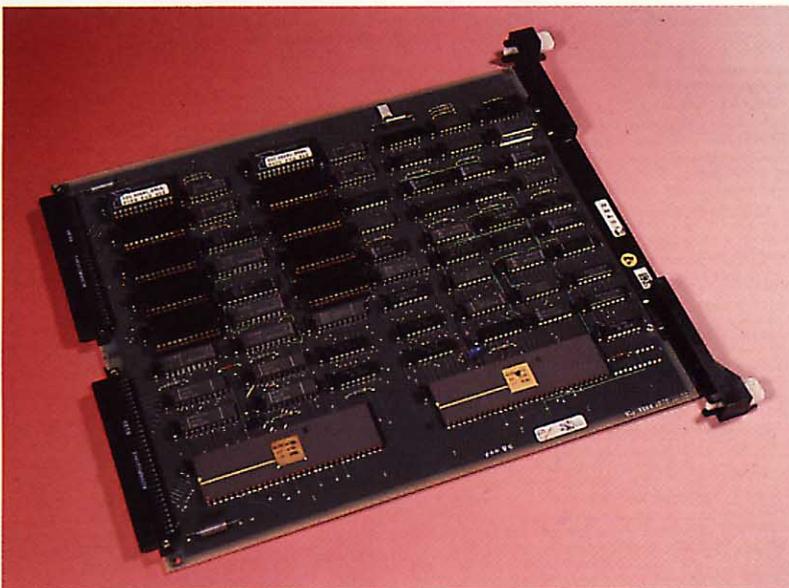
El sistema de pruebas presenta tres importantes características. Primera: puede probar líneas y enlaces no sólo en una central sino también en unidades remotas. Segunda: se aplica al mantenimiento de la planta exterior, además de los circuitos de línea y enlace. Por último, hay un circuito de línea de reserva en cada bastidor.

El equipo de pruebas del Sistema 12 se divide en dos bloques diferentes: la unidad de acceso a prueba (UAP) y el analizador de señal de prueba (ASP). La primera, en esencia analógica, da acceso metálico a la planta exterior, y suele equipar una placa impresa por cada bastidor de línea o de enlace, aunque el número de UAP puede variar en función de las necesidades locales. El ASP, independiente del número de UAP, utiliza técnicas de proceso digital de señales para generar los tonos y obtener el filtraje. Su función está centralizada, bien en la central local o bien en un centro de mantenimiento, y se conecta a las UAP mediante un enlace MIC típico.

La UAP ofrece el medio de acceso a la línea para la realización de pruebas hacia el interior o hacia el exterior, mas no ejecuta medición alguna. Además, la unidad proporciona una precisa conversión analógico-digital. Ha sido objetivo importante del diseño simplificar la UAP y centralizar las funciones de proceso de señal en el ASP, reduciendo así al mínimo la necesidad de calibración rutinaria y de ajuste en fábrica de los equipos medidores.

Esta nueva arquitectura del sistema de pruebas resulta económica incluso en las pequeñas aplicaciones (unidades de con-

Placa de procesador ASP, basada en una placa impresa estándar del Sistema 12.



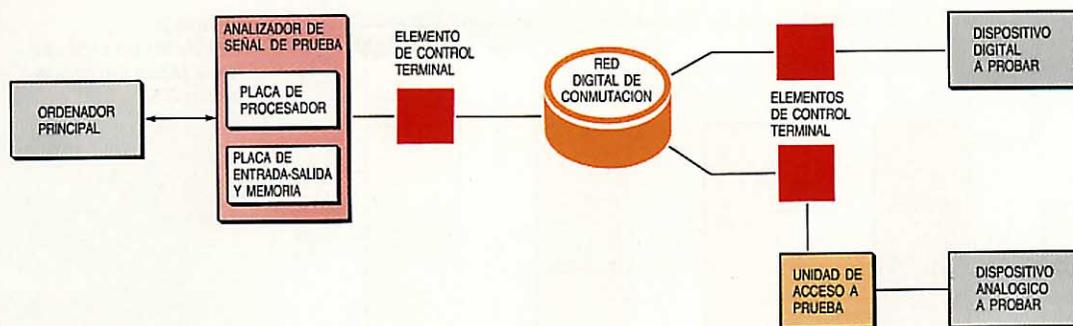


Figura 1
Arquitectura del sistema de prueba.

mutación remotas atendidas por un centro de mantenimiento, ya que sólo se requieren UAP en los conmutadores remotos). Cuando una central se digitalice enteramente, con todas sus líneas y enlaces, no se necesitará ya UAP pero el ASP permanecerá invariable.

Así, en efecto, el ASP realizará comprobaciones en el bucle digital de abonado en cuanto se introduzcan teléfonos digitales. Puede aplicarse a diversas señales, (p. ej., medida de resistencia en función de la frecuencia) y no se limita a una serie concreta de pruebas, sino que admite aplicación a pruebas adicionales. Por ser programable, el ASP podrá mantener la compatibilidad con las futuras normas de prueba para líneas digitales.

Aunque se ha concebido para utilizarse en el Sistema 12, la versatilidad del ASP le permite otras aplicaciones, sirviendo como herramienta para probar nuevos conceptos basados en proceso digital.

Consideraciones de diseño de sistema

El ASP es el corazón del sistema de mantenimiento y diagnóstico de línea, enlace y circuito de servicio para el Sistema 12.

El mantenimiento de la línea de abonado puede dividirse en tres áreas: mantenimiento del terminal de abonado (usualmente un teléfono a dos hilos), de la vía de transmisión entre dicho terminal y la central, y del circuito de terminación en la central. Esta división vale tanto para las centrales analógicas como para las digitales, y también para enlaces entre centrales y circuitos de servicio especial.

Dado que las centrales digitales y analógicas coexistirán durante bastantes años en el mundo de las redes telefónicas, sería deseable disponer de un sistema de prueba único para ambos entornos, analógico y digital. Asimismo, ya que la arquitectura del Sistema 12 permite un margen muy amplio de tamaños y la introducción de nuevos servicios y tecnologías, el sistema de prueba debería ser capaz de expansión y

de evolucionar para atender futuras necesidades telefónicas. Al comienzo del desarrollo se definieron los siguientes requisitos para las pruebas de líneas y enlaces:

- Economía y eficacia en la prueba de las facilidades de planta exterior de abonado, ya terminen en la central o en una unidad remota de abonados.
- Capacidad de probar enlaces a 2 y 4 hilos, así como circuitos de servicio (p. ej., receptores y emisores multifrecuencia).
- Flexibilidad para probar los aparatos telefónicos actuales y futuros.

El método de diseño elegido para cumplir estos requisitos se proponía tres objetivos principales. Ante todo, un bajo coste inicial del sistema, para lo cual éste debería constar de dos partes: un interfaz analógico dependiente de la aplicación (UAP), y una parte común para tratar los aspectos de medida y cálculo de las pruebas (ASP). Dado que UAP y ASP se comunican por un enlace MIC, el ASP tiene que basarse en técnicas de proceso digital de señales.

En segundo lugar, el sistema de prueba había de tener versatilidad suficiente para ser utilizado en distintas aplicaciones y cumplir una extensa gama de requisitos, que incluyen bucles digitales, concentradores y centrales supervisadas.

Por último, se tenía que expandir la capacidad de diagnóstico, lo cual implicaba la posibilidad de interpretación y proceso diferidos (pasado el periodo de medida). En consecuencia, las tareas de procesamiento muestra por muestra y la interpretación han de ser ejecutadas por procesadores diferentes.

Estas consideraciones condujeron a una configuración de prueba basada en módulos analógicos (UAP), distribuidos entre los módulos de línea y enlace, y en uno o más procesadores de señal centrales (ASP). El número de UAP conectadas a un ASP dependerá de la complejidad, frecuencia y número de las pruebas. El ASP se conecta a la central por un enlace MIC a 4 Mbit s^{-1} típico del Sistema 12.

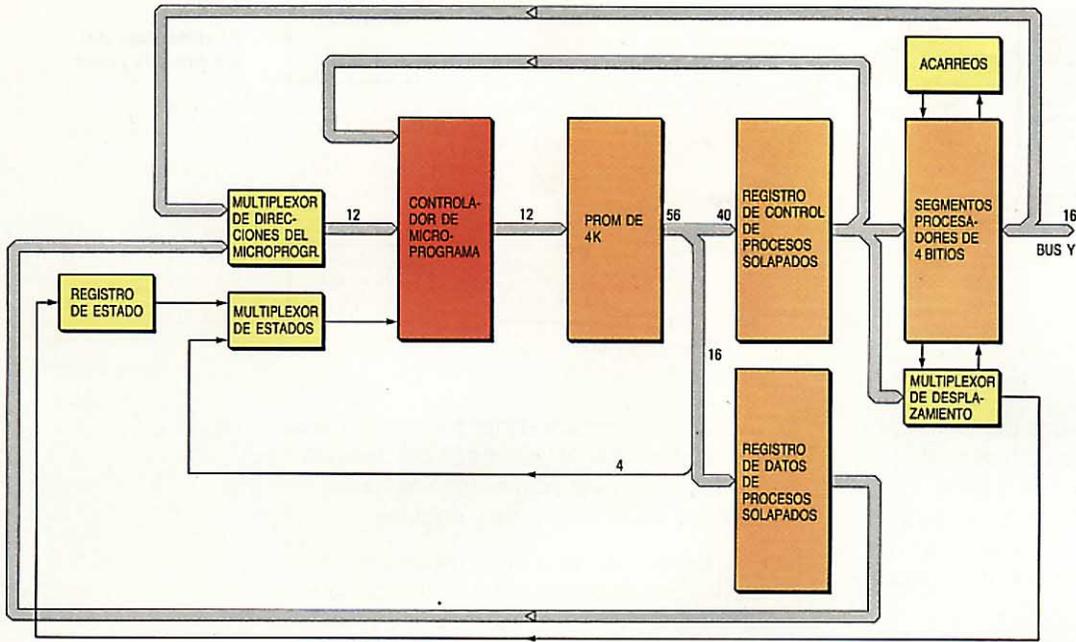


Figura 2
Diagrama de bloques de la placa del procesador ASP.

Descripción del sistema

El ASP es un procesador digital de señales, de alta velocidad, que procesa muestra por muestra, bajo control de un ordenador principal de uso general (Fig. 1). El ordenador principal recibe la petición de medida, carga el programa correspondiente en el ASP, y asigna y encamina las señales a través de la red y la UAP hacia el dispositivo a probar. El ASP ejecuta el programa de aplicación que genera una señal de prueba y analiza las muestras recibidas, devolviendo luego los resultados al citado ordenador. Las operaciones de proceso de señal más frecuentes se efectúan mediante un extenso conjunto de instrucciones. Por ejemplo, el filtrado recursivo, medida de frecuencia, generación de señal y medidas de amplitud eficaz se realizan mediante instrucciones individuales, tratando una sola muestra cada una de ellas. Esto facilita la creación de programas de aplicación para muchos tipos de medidas de líneas, enlaces y circuitos de servicio.

El ordenador principal analiza y convierte las medidas a valores físicos absolutos, y también combina varias medidas para determinar un parámetro físico, comprobando si tal parámetro se encuentra entre los límites especificados.

Soporte físico del analizador de señal de prueba

El ASP se conecta a la red mediante un enlace MIC estándar, y al procesador principal a través de un bus de control. Está realizado en dos placas de circuito impreso:

- placa de procesador, que contiene la estructura fundamental (lógica de direccionamiento del microprograma, memoria ROM para almacenamiento del mismo, y la unidad aritmética-lógica)
- placa de entrada/salida y memoria, que proporciona la estructura periférica (lógica de direccionamiento de memoria, interfaz con la red e interfaz de control con el ordenador principal).

Estas placas se interconectan mediante un bus de entrada/salida por el que circulan datos y direcciones.

La placa del procesador, un procesador de uso general de coma fija, se basa en cuatro procesadores bipolares por segmentos ("bit-slice") (Fig. 2). El microprograma que lo controla reside en una PROM bipolar de 4 k-palabras de 56 bits. Dicho microprograma incorpora funciones de autodiagnóstico y depuración de programas, un sistema operativo y un juego de instrucciones de proceso de señal orientado al análisis de averías. El tiempo de ejecución de una microinstrucción es de 200 ns; las instrucciones simples requieren 3 ciclos de microinstrucción, mientras que las más complejas pueden necesitar cientos de ciclos.

El microprograma se direcciona mediante un secuenciador bipolar. La dirección de la siguiente instrucción se determina en parte por la instrucción en curso. Cada ciclo de microinstrucción consta de dos fases: búsqueda y ejecución. Una microinstrucción se direcciona durante un periodo de reloj y se carga en un registro para procesos solapados ("pipeline") al final de este periodo, ejecutándose al siguiente ciclo de reloj.

De esta forma, mientras se ejecuta una instrucción, se está ya direccionando la siguiente. Este solape permite al ASP operar aproximadamente a velocidad doble de la de un diseño equivalente que no sea de este tipo.

La salida del mencionado registro controla el ASP. Distintos grupos de bits controlan el secuenciador del microprograma, los segmentos de procesador, el registro de instrucción, el registro de estado, los multiplexores y dispositivos periféricos. La salida del subsistema procesador forma un bus combinado de datos y direcciones, conectado a la placa de entrada/salida y memoria. Este bus se puede conectar a otras placas (p. ej., la de los multiplicadores y conversores analógico-digitales).

La placa de entrada/salida y memoria (Fig. 3) está controlada por la del procesador; esta arquitectura permite que otros tipos de placas (p. ej., las especificadas por el usuario para aplicaciones distintas a los

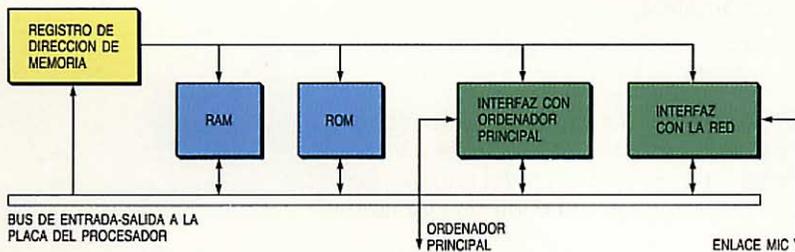


Figura 3
Diagrama de bloques de la placa de entrada/salida y memoria.

análisis de pruebas) sean conectadas a la del procesador. La placa de entrada/salida y memoria incorpora una memoria RAM de 4 k-palabras para almacenar programas de carga dinámica, variables de programa y muestras de datos; una memoria PROM bipolar de 2 k-palabras guarda las tablas de codificación y los módulos de programas de aplicación corrientemente utilizados. También existen un interfaz de control con el ordenador principal y un interfaz de datos serie con la red, datos que pueden ser transmitidos y recibidos simultáneamente.

Un bus de control y un bus bidireccional de datos y direcciones interconectan el procesador con la placa de entrada/salida y memoria, pudiendo dichos buses expandirse para atender a otros dispositivos. Los dispositivos de entrada/salida están organizados en la memoria de forma tal que el margen de direccionamiento del ASP es de 64 k-palabras. Cuando aparece una dirección en el bus, se carga ésta en un registro de direcciones de memoria presente en todas las placas excepto en la del procesador, cuyo contenido decodificado activa los dispositivos de una determinada placa.

Microprogramación

La microprogramación hace que el ASP sea un procesador de señales versátil y de fácil uso. En ella pueden distinguirse cuatro grupos: diagnósticos, herramientas de desarrollo de programas, sistema operativo y conjunto de instrucciones orientadas al tratamiento de problemas. Estas funciones necesitan actualmente 2,5 k-palabras de las 4 k-palabras disponibles para microprogramas, lo que permite posteriores expansiones.

Diagnósticos

Existen dos grupos de diagnósticos: en línea y fuera de línea. Los diagnósticos fuera de línea, de funcionamiento autónomo, se utilizan para localización de averías y no se describen aquí. Los diagnósticos en línea a su vez se dividen en internos y externos. Los internos se ejecutan enteramente en el ASP, enviándose al ordenador principal un informe sobre su éxito o fracaso. En cambio, los externos se realizan en cooperación con el ordenador principal y hacen uso de los interfaces con dicho ordenador y con la red.

Herramientas de desarrollo

Proporcionan un medio sencillo de comunicación con el ASP al objeto de probar programas de aplicación. El sistema operativo aporta funciones similares, pero la utilización de sus rutinas resulta más compleja para el desarrollo de programas. Las herramientas en cuestión incluyen:

- lectura/escritura en memoria, que transfiere información entre la memoria del ASP y el ordenador principal
- ejecución, que ejecuta el programa de prueba en la memoria del ASP
- depuración, que permite suspender la ejecución del programa en determinados casos.

Sistema operativo

El sistema operativo o monitor proporciona la vía de comunicación entre el ASP y el ordenador principal. Las operaciones del monitor incluyen la transmisión de información entre dichos procesadores, los movimientos de información dentro del ASP y la ejecución de programas de generación de señal y de medidas. Además, el monitor es responsable de la integridad de las transferencias de datos entre el ordenador principal y el ASP. Siempre que el ASP no esté ejecutando ningún programa de aplicación, se encontrará en un modo desocupado de exploración, en espera de alguna orden del monitor.

La referida comunicación entre ordenador principal y ASP se desarrolla así: el ordenador prepara un programa monitor, que es una secuencia de órdenes, y lo escribe en la memoria tampón de órdenes junto con un recuento de palabras y una suma de comprobación. Si no hay errores de transmisión, se transfiere el control al ASP y se ejecuta la secuencia de órdenes.

Repertorio de instrucciones

La facilidad de uso del ASP procede del repertorio de instrucciones orientadas a la resolución de problemas. Este repertorio se estructura a modo de un lenguaje de alto nivel, como el PL/1. Ninguna de las instrucciones se refiere a los registros internos de los segmentos procesadores, salvo para el contador de programa y el indicador de posición, por lo que la arquitectura del ASP y su soporte físico son totalmente transparentes para el usuario.

Se utiliza aritmética de complemento a dos, y los datos pueden ser almacenados en diferentes formatos: precisión simple o doble, enteros o fraccionarios.

Existen 57 instrucciones, clasificables en 8 grupos. El repertorio de instrucciones se orienta hacia la adquisición y proceso en tiempo real de una muestra cada vez, pudiendo ésta ser procesada por varias instrucciones a fin de obtener distintas medidas. Además, algunos procesos se dividen en dos fases: proceso realizado en cada muestra, y cálculos finales que sólo se efectúan al final del periodo de medida.

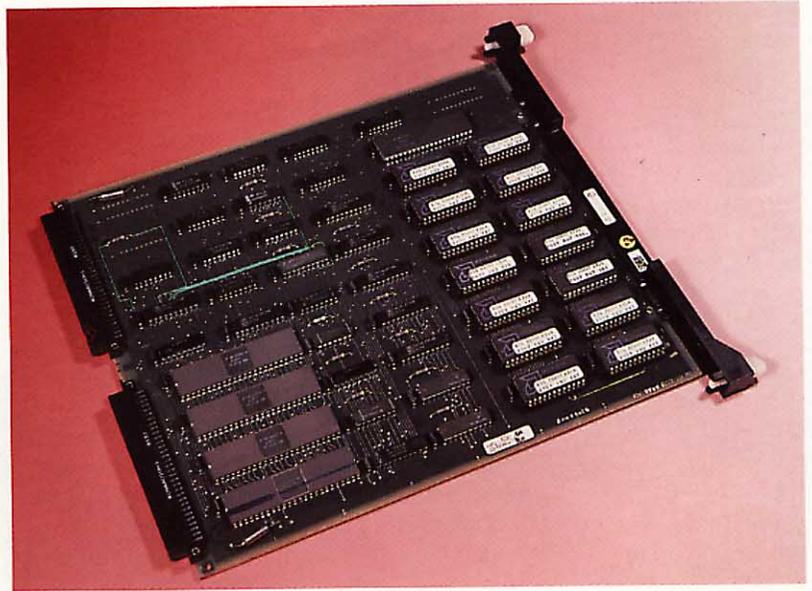
Los grupos de instrucciones son los siguientes:

Grupo misceláneo. Desempeña funciones diversas, tales como inicialización de un bloque de memoria, traslado de bloques de memoria de uno a otro lugar, borrado de indicadores, etc.

Grupo aritmético. Además de las funciones aritméticas y lógicas usuales, incorpora la función raíz cuadrada y el proceso por muestra de medias parciales y desviaciones típicas.

Grupo de medidas. Cada instrucción de este grupo se utiliza en bucle para procesar una serie de muestras individuales. Las instrucciones son:

- Medida de intervalo, que mide el tiempo transcurrido entre los cambios de estado de una señal de entrada (posiblemente multietapa). Los parámetros de entrada especifican la histéresis y el periodo de transición que definen un cambio de estado.
- Medida de frecuencia.
- Medida del espectro, aplicando el algo-



Placa de entrada/salida y memoria.

ritmo de Goertzel a la misma muestra de entrada para medir la amplitud eficaz de cada tono de un conjunto de tonos con frecuencias especificadas en la lista de parámetros.

- Medida del valor de pico de una señal de entrada.

Grupo de filtrado. Este grupo contiene instrucciones para tres tipos de filtros. Desde el punto de vista del usuario son instrucciones de una línea las siguientes:

- Filtro recursivo, que se realiza mediante una cascada de N etapas de segundo orden. Para cada etapa se especifican los coeficientes progresivos y regresivos, así como los parámetros de escala.
- Filtro de promediado dinámico, que es un filtro de media dinámica de N muestras (N es una potencia de dos).
- Filtro de respuesta impulsiva finita; está formado por una función de ponderación en la que se calcula una salida para cada M muestras, de acuerdo con la estructura de respuesta impulsiva de cada uno de los N filtros independientes.

Grupo de generación de señal. Las siguientes instrucciones generan una muestra cada vez que se ejecutan:

- Generación de senoide, obteniéndose en cada punto de muestra la suma de M sinusoides en la gama de 2 a 4 kHz, y N sinusoides de frecuencia inferior a 2 kHz. La frecuencia y amplitud de cada senoide se especifican por separado, al igual que la fase inicial y su dirección.
- Generación de ruido, mediante un generador de ruido pseudoaleatorio.
- Generación exponencial, que genera una muestra de onda "rectangular" con

tiempo de subida-bajada exponencial. Los parámetros incluyen los tiempos de subida y bajada, la duración de los distintos estados, amplitudes y la componente continua de compensación.

Grupo de terminación. Estas instrucciones complementan a las del grupo de medidas y se las llama al terminar las mediciones. Las funciones son: finalizar las medidas de promediado parcial, desviación típica, espectro y frecuencia.

Grupo de control de subrutinas. El ASP, por medio de sus potentes instrucciones de control de subrutinas, permite la programación estructurada, así como el anidamiento de subrutinas. Entre las instrucciones se incluye la ejecución condicional de una subrutina.

Grupo de entrada/salida a la red. Estas instrucciones permiten al usuario conectarse con la red serie MIC. Para utilizarlas sólo se necesita un mínimo de conocimientos del interfaz serie, ya que ellas mismas controlan todos los procesos necesarios, como la sincronización de red. Las transferencias se realizan siempre entre la red y la memoria, en la que las muestras pueden ser tratadas por otras instrucciones. Las muestras se transmiten y reciben en formato lineal de 8 ó 13 bits, o bien en ley μ ó A de 8 bits. La transmisión y recepción se pueden simultanear, controlando por separado las codificaciones de entrada y salida, totalmente programables.

Los tiempos de ejecución de estas instrucciones varían entre centenares de nanosegundos y algunos microsegundos. Como el objetivo es procesar muestra por muestra, y dado que la velocidad de muestreo en la red es de 8 kHz, todas las instrucciones de manejo de muestra, incluidas las de generación y filtrado de señal, deben ser ejecutadas en menos de 125 μ s para asegurar que las listas de parámetros tengan una longitud razonable.

Aplicaciones

El ASP es un versátil procesador digital de señales, adecuado para muy diferentes aplicaciones, que juega un papel primordial en el mantenimiento y diagnóstico de líneas y enlaces del Sistema 12.

Aplicaciones en el Sistema 12

La arquitectura de control distribuido de la central digital Sistema 12 consta de módulos dotados de microprocesador, conectados con la planta exterior (líneas de abonado, enlaces, etc) y con una red digital de conmutación inteligente. El elemento de

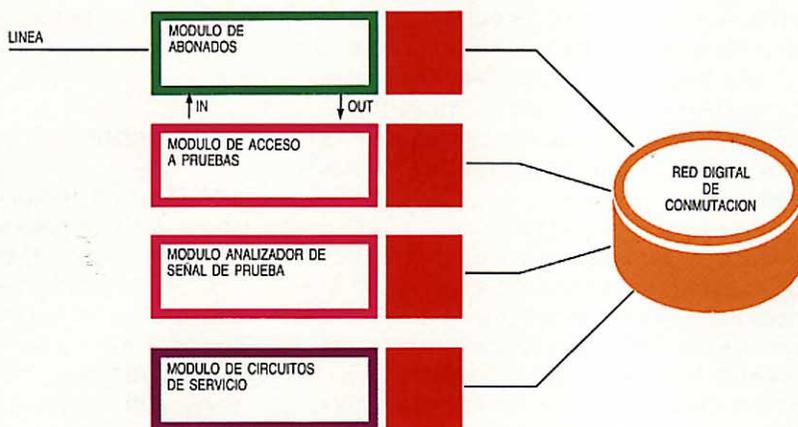


Figura 4
Configuración del sistema de prueba.

control terminal de cada módulo se comunica con la red de conmutación mediante un enlace MIC a 4 Mbit s⁻¹. Los datos se estructuran funcionalmente en tramas, conteniendo cada una 32 canales y con una frecuencia de repetición de 8 kHz, lo cual obliga a que cada muestra sea procesada por el ASP en menos de 125 μ s. Dos de los 32 canales se reservan para control y sincronización de la red, y los 30 restantes pueden llevar datos o voz. El ASP se comunica con la red a través de uno de estos 30 canales disponibles.

El ASP constituye la parte terminal de un módulo del Sistema 12, y por lo tanto puede recoger/enviar muestras de prueba de/a cualquier dispositivo conectado a otros módulos del sistema, incluidos los módulos de otras centrales, a través del enlace MIC y la red digital de conmutación.

Además de comunicarse con el ASP, el microprocesador del módulo donde éste se aloja establece caminos a través de la red para conectar el ASP al dispositivo a probar (ver figura 4). Las pruebas más simples, como la medida de ruido en la línea de abonado, necesitan un solo camino: así,

Tabla 1 — Medidas de prueba típicas que utilizan el ASP

Medidas de mantenimiento
Frecuencia y amplitud de potenciales exteriores (tensiones inducidas)
Ruido
Capacidad
Resistencia de fugas
Medidas de circuito de línea
Respuesta en frecuencia
Detección de descuelgue
Impedancia terminal
Distorsión
Medidas de circuito de servicio
Amplitud y frecuencia de tonos generados
Prestaciones del receptor MF/DTMF

pues, la línea a probar se conecta a la entrada analógica de la UAP por el bus de prueba *out*, siendo digitalizada la señal por dicha UAP y transmitidas las muestras al ASP a través de la red. Un programa en el ASP recibe las muestras y efectúa las medidas.

Las pruebas complejas, como la respuesta en frecuencia del lado de recepción de un circuito de línea, necesitan dos caminos. El programa de aplicación del ASP genera una señal de prueba que se transmite por la red a la UAP, donde se convierte en una señal analógica; ésta se envía seguidamente al dispositivo a probar a través del bus de prueba *in*. El circuito de línea digitaliza la señal y envía las muestras a través de la red al ASP, donde está el programa que recibe las muestras de respuesta y realiza las medidas al tiempo que genera y transmite la señal de prueba.

Son prestaciones fundamentales del ASP la prueba remota de dispositivos, y el que los programas de prueba no tengan que residir en el mismo módulo sino que puedan ser cargados desde cualquier generador de programas a través de la red. Se pueden realizar muchos tipos de pruebas, y sin modificaciones del soporte físico es posible añadir pruebas y mejorar las existentes. El tiempo requerido para una medida específica suele ser de 0,1 a 0,2 s, lo que permite la adquisición de datos suficientes para conseguir la precisión requerida.

En la tabla 1 se enumeran las medidas típicas.

Aplicaciones adicionales

La arquitectura de uso general del ASP, así como su flexible repertorio de instrucciones orientadas al usuario, han servido para otras aplicaciones que la prueba de enlaces y líneas. Así, el ASP se ha utilizado como potente herramienta de laboratorio para probar nuevos procedimientos basados en el proceso digital de señales; ejemplo de ello es su uso en el proceso de la voz para experiencias sobre transmisión del habla.

Esta aplicación del ASP ha hecho posible probar y verificar conceptos en las primeras etapas de desarrollo, permitiendo obtener suficientes datos experimentales y de prueba sin necesidad de soporte físico adicional. Otro ejemplo es el de la detección de señales de conversación en presencia de ruido de fondo, lo cual requiere complejos cálculos en tiempo real que no pueden ser simulados con un microprocesador de

uso general, pero que son particularmente adecuados para el ASP.

Conclusiones

El ASP es un versátil procesador digital de señales en tiempo real con un potente repertorio de instrucciones orientadas al análisis de problemas. Puede ser utilizado en muchas aplicaciones, además de aquellas para las que fue diseñado: mantenimiento de líneas, enlaces y circuitos de servicio del Sistema 12. Ha servido también como potente herramienta de laboratorio para la prueba de nuevos conceptos que hacen uso de técnicas de proceso digital de señales.

S. Das obtuvo las graduaciones BS y MS en electrónica y telecomunicaciones por la Universidad de Jadavpur, Calcuta, en 1965 y 1968, respectivamente. Antes de trasladarse a los Estados Unidos en 1969, trabajó en el proyecto del ordenador para aquella Universidad y el Instituto Indio de Estadística. Estudió ingeniería eléctrica en la Universidad de Washington, San Luis, donde se graduó DSc en 1973. El Dr. Das se incorporó entonces al centro de investigación de North Electric (hoy ITT-TNA), donde colaboró en el desarrollo de la central digital ITT 1210. Desde marzo de 1976 trabaja en el ITT Advanced Technology Center, Shelton, dirigiendo actualmente la División de Tecnología Aplicada.

Sarma Jayanthi nació en 1953, en la India. Se graduó BS en físicas e ingeniería eléctrica en 1972 y 1976, respectivamente, en la Universidad de Toledo, Ohio. Dos años más tarde obtuvo un MS en ingeniería eléctrica. El Sr. Jayanthi ingresó entonces en el ITT Advanced Technology Center, donde es miembro del grupo de diseño digital en el departamento de desarrollo tecnológico.

Francisco A. Middleton nació en México en 1950. En 1972 obtuvo la licenciatura en ingeniería eléctrica en el Instituto Politécnico Nacional de México, incorporándose a continuación a ITT. Entre 1972 y 1976 fue ingeniero de proyecto para el desarrollo de centralitas privadas. En 1976, el Sr. Middleton fue nombrado subdirector del laboratorio de desarrollo de Indetel, y dos años después ingresó en el ITT Advanced Technology Center para trabajar en el desarrollo del soporte físico (hardware) del Sistema 12. Desde enero de 1984, es director del departamento de desarrollo tecnológico de la División de Tecnología Aplicada.

Steven G. Morton obtuvo en 1972 los títulos de BSEE y MSEE en el Instituto Tecnológico de Massachussets. Se incorporó luego al laboratorio Lincoln del MIT, donde trabajó en instrumental de comunicaciones espaciales para las Fuerzas Aéreas. El Sr. Morton ingresa en ITT en Mayo de 1979, y actualmente dirige el diseño de procesadores avanzados en la División de Tecnología Aplicada del ITT Advanced Technology Center. Es responsable técnico del procesador celular matricial, un procesador paralelo, tolerante a fallos.

Proceso digital de señales en módems de transmisión de datos

Muchas de las funciones de los módems transmisores de datos exigen una considerable cantidad de cálculo. Una juiciosa elección de los algoritmos y la estructura de datos ha permitido desarrollar un módem que utiliza microprocesadores existentes en el mercado para el proceso digital de señales.

G. Bång
P. Carlsson

Standard Radio & Telefon, Estocolmo, Suecia

Introducción

Los módems de datos están diseñados para transformar las señales digitales de datos de tal forma que pueda transmitirse la información digital por la red telefónica, recuperándola en el lado receptor con la mayor seguridad y exactitud posible. Las administraciones de telecomunicación imponen requisitos estrictos en cuanto a la amplitud de la señal que se transmite, y el ancho de banda normalmente se limita a 3.000 Hz. Para confinar en este margen la potencia de señal de línea, la señal digital de datos en el transmisor debe ser convertida a un formato analógico. Así, el transmisor y el receptor del módem pueden considerarse como conversores digital-analógico y analógico-digital, respectivamente (Fig. 1).

En los primeros módems, el interfaz digital-analógico estaba cerca del lado digital. Se procesaban digitalmente palabras de un solo bitio, mientras que la modulación, demodulación, recuperación de portadora y otras funciones del módem eran realizadas por circuitos analógicos. Sin embargo, en cuanto el soporte físico para el proceso digital de señales ha llegado a ser económico frente a los circuitos analógicos, tal modo de proceso se ha aplicado a más funciones de complejidad numérica, con la consiguiente mejora en el funcionamiento.

Figura 1
Módems de datos considerados como conversores analógico-digital y digital-analógico.



En el nuevo módem 2284 de SRT (Standard Radio & Telefon), un módem para transmisión en dúplex total por dos hilos a 2400 bit s^{-1} (CCITT V.22 bis), se han aprovechado las ventajas del proceso digital de señales.

Líneas telefónicas

Las líneas telefónicas degradan las señales de diversas formas, las más importantes de las cuales son la distorsión lineal, el ruido, los desplazamientos de frecuencias, la fluctuación de fase y la distorsión armónica. La distorsión lineal puede dividirse en distorsión de amplitud (atenuación) y de retardo de grupo. El control automático de ganancia puede compensar sobradamente la primera, mientras que un igualador automático adaptativo puede hacer frente a la distorsión de retardo de grupo y a la distorsión residual de amplitud.

Nunca desaparece el ruido en un sistema, pero sus efectos adversos pueden contrarrestarse mediante adecuados filtros de transmisión y recepción. Puede haber desplazamientos de frecuencia y fluctuaciones de fase en circuitos que incluyan enlaces por portadora. Ambas degradaciones se pueden detectar y minimizar utilizando un bucle de enganche de fase.

La distorsión armónica es más difícil de compensar, pero normalmente su nivel es muy bajo y no afecta de modo apreciable a la calidad de la transmisión.

Funciones del módem

El módem 2284 de SRT es capaz de operar en dúplex total a 2400 ó 1200 bit s^{-1} por líneas a dos hilos, conmutadas o alquiladas. Utiliza transmisión binaria en serie, síncrona o asíncrona, con 8, 9, 10 u 11 bitios por

Las velocidades de muestreo, de bits y baudios, se extraen pasado el CAG. Las señales procedentes de los dos filtros paso banda, centrados en torno a 900 y 1500 Hz respectivamente, se mezclan después. La frecuencia de 600 Hz, obtenida por diferencia, se filtra y se aplica a un bucle digital de enganche de fase para reducir la fluctuación de fase de las señales de reloj. Esto permite obtener un sistema estable y progresivo de recuperación de reloj.

Las muestras entregadas por el conversor analógico-digital, a 3600 Hz, se llevan al igualador, que es un filtro transversal con dos juegos de coeficientes variables. Los valores iniciales de tales coeficientes corresponden a filtros paso banda centrados en 1200 Hz, uno de los cuales es la transformada Hilbert del otro. Las dos señales de salida del igualador se muestrean a 600 Hz.

Por ser enteras las relaciones entre la frecuencia de muestreo (3600 Hz), la frecuencia portadora (1200 Hz) y la velocidad de baudios (600 Hz), no se necesita ningún procesamiento adicional para obtener la señal en banda de base. Tomando regiones de decisión apropiadas en el cuantificador, puede decidirse cuál es el punto de la señal que más probablemente se ha enviado. La diferencia entre el punto de señal realmente recibido y el punto fijado como objetivo forma una señal de error, cuyas componentes en fase y en cuadratura se utilizan para actualizar los dos juegos de coeficientes del igualador. La actualización está regida por las ecuaciones siguientes:

$$(C_k)_{nuevo} = (C_k)_{anterior} - \alpha * (ERR) * (S_k) \times$$

$$(D_k)_{nuevo} = (D_k)_{anterior} - \alpha * (ERR) * (S_k) \text{ y}$$

donde

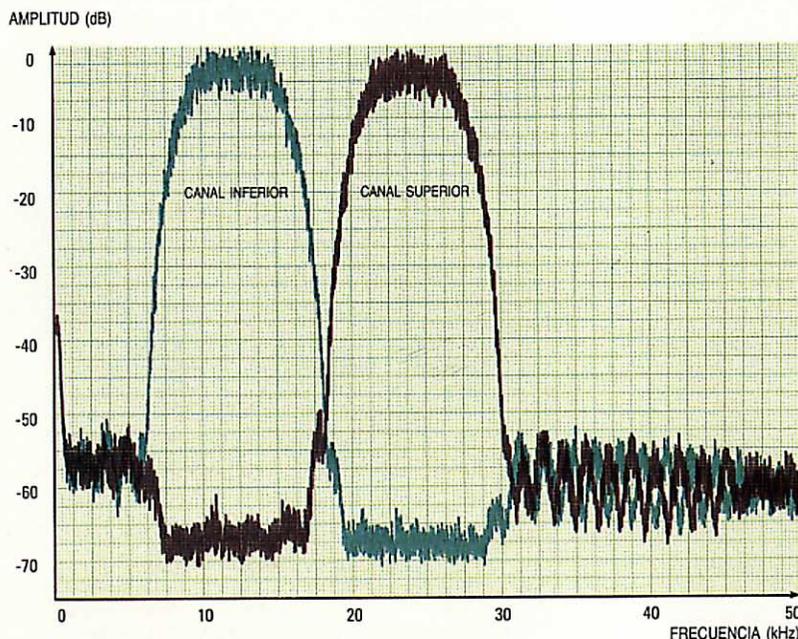
- (C), (D) - dos conjuntos de coeficientes
 ERR - señal de error del igualador
 S - valor de entrada
 α - paso de la iteración
 k - k-ésimo coeficiente o n° de orden de entrada
 x, y - componentes en fase y cuadratura, respectivamente.

Esta configuración del igualador tiene la ventaja de ser insensible a los errores de fase del reloj. Su sencilla estructura, que no necesita interacoplamiento entre los canales en fase y en cuadratura, facilita su realización. Cuando la señal de línea recibida está degradada por un desplazamiento de frecuencia, es necesario compensar este desplazamiento antes de cuantificar, lo cual se consigue utilizando un bucle de enganche de fase de segundo orden controlado por el error de fase obtenido en el cuantificador. Ya que la fase se ajusta inmediatamente antes del cuantificador, no se produce

ningún retardo adicional en el bucle y, por tanto, se mejora la capacidad de seguimiento de fase.

La señal cuantificada se decodifica y desaleatoriza utilizando técnicas conven-

Figura 3
Espectros transmitidos en los canales superior e inferior.



cionales, y los datos se entregan al usuario. Cuando se recibe el canal superior, el decodificador debe tener en cuenta la inversión de fase introducida por el modulador del receptor.

Unidad de control

Un cierto número de funciones de control se realizan digitalmente, incluyendo:

- la marcación automática, que acelera el establecimiento de la conexión, ahorra el tiempo consumido por el operador en repetidas marcaciones e incrementa la seguridad del sistema
- la conversión arrítmico/síncrono, admitiendo todo tipo de protocolos, tales como el control de alto nivel del enlace de datos, el control de enlace de datos síncrono y el X.25 del CCITT
- detección automática de la velocidad binaria, que determina si el módem está trabajando con otro módem V.22 bis a 2400 bit s⁻¹ o con un módem V.22 a 1200 bit s⁻¹
- extenso repertorio de facilidades de diagnóstico.

Realización

En relación con el soporte físico había que tomar dos decisiones importantes. La primera era dónde colocar el interfaz analó-

gico-digital con el fin de sacar el máximo provecho de las diferentes ventajas de ambas técnicas. La segunda decisión era si el proceso digital de señal debía centralizarse en un potente procesador adecuado a ese fin, o bien descentralizarse en un sistema distribuido, con menores demandas de potencia de procesamiento para cada parte.

Interfaz analógico-digital

La tecnología digital aventaja a la analógica en las siguientes áreas:

- Flexibilidad, al poder realizarse diferentes filtros y funciones de modulación, con el mismo soporte físico.
- Inteligencia, ya que se facilita mucho la toma de decisiones complejas como se requiere en la parte detectora del receptor y del igualador adaptativo.
- Precisión, al poderse reproducir exactamente las mismas funciones sin variaciones debidas a dispersión de los valores de los componentes, minimizando por tanto el ajuste y las pruebas.

No obstante, existen algunas desventajas importantes asociadas a la técnica digital:

- Errores de cuantificación introducidos por la limitada resolución de los convertidores analógico-digital y digital-analógico y por la longitud de palabra utilizada en los cálculos.
- Posible dificultad de alcanzar la velocidad de cálculo requerida, al tener que utilizar una frecuencia de muestreo que sea al menos el doble de la máxima componente frecuencial presente en la señal a procesar.

Considerando estas ventajas y desventajas se eligió la solución descrita a continuación.

En el transmisor, la conformación y modulación de la amplitud en cuadratura se realizan de forma enteramente digital, con el fin de conseguir la flexibilidad necesaria para cambiar los canales de transmisión y añadir los tonos de guarda. También permite realizar el filtro conformador como filtro de respuesta impulsiva finita con la propiedad de fase lineal. Se utiliza un conversor digital-analógico de 12 bits para mantener el ruido de cuantificación a un nivel suficientemente bajo, tal que no interfiera con el canal recibido.

Se eligió la elevada velocidad de muestreo de 9600 Hz con el fin de desplazar hacia arriba las frecuencias del espectro que son múltiplos de la frecuencia de muestreo, lo cual permite aligerar las exigencias impuestas a los filtros paso bajo analógicos que van después.

El problema en el lado receptor es más complejo, a consecuencia del amplio margen dinámico de la señal recibida. Según la recomendación del CCITT, el módem debe ser capaz de recibir señales con niveles de sólo -48 dBm. Si el transmisor genera una señal de 0 dBm y hay desadaptación entre la híbrida y la línea, podría ocurrir que las fugas desde el transmisor hacia el receptor no se atenuaran en más de 6 dB. En el caso más desfavorable, esto podría dar lugar a que el nivel de la señal recibida fuera 42 dB inferior al nivel de las fugas.

El procesador digital de señal requiere al menos 7 bits de resolución en la señal recibida, lo cual significa que, si no hay ninguna supresión de fugas en el lado analógico, la situación más crítica requeriría un conversor analógico-digital con resolución no inferior a 14 bits. La inclusión de un filtro analógico de canal seguido de un CAG, reduciría este requisito a 8 bits.

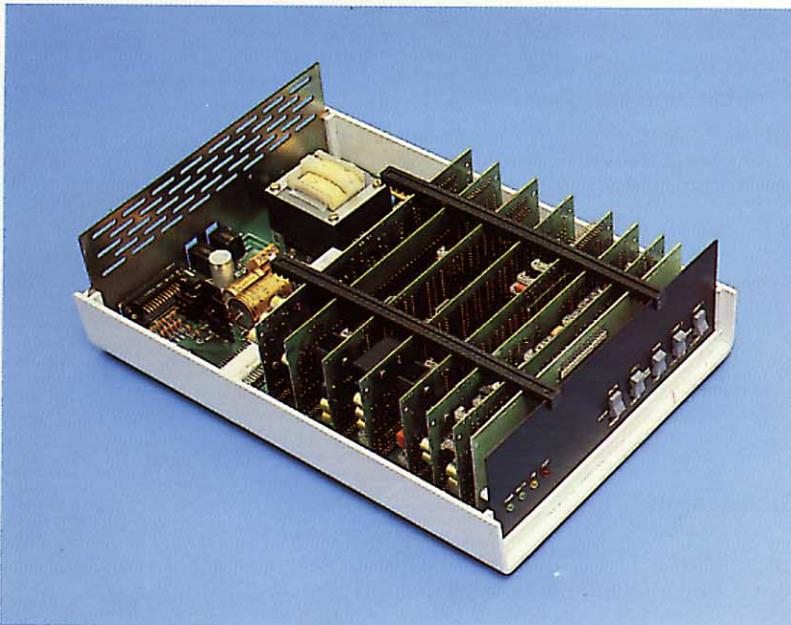
La traslación del canal superior al inferior permite reducir la velocidad de muestreo hasta 3600 Hz y no tener solapamientos.

Partición digital

Había que elegir entre un sistema centralizado con un potente procesador digital de señal que compartiera su tiempo entre diferentes funciones del módem, y un sistema distribuido en el que el trabajo se repartiera entre un cierto número de procesadores, menos potentes, de aplicación general. Se escogió la segunda alternativa considerando el coste y la modularidad del sistema distribuido, que da lugar a un diseño más estructurado.

El hecho de que el transmisor y el receptor trabajen entre sí de modo asíncrono justificó su realización en dos sistemas diferentes. El sistema receptor está dividido

Módem de datos 2284 de SRT, que utiliza en gran medida el proceso digital de señal.





Módem 2284 utilizado con un ordenador personal ITT XTRA*.

en dos subsistemas, uno para la igualación adaptativa y el otro para realizar las funciones de demodulación restantes.

Se incluye un sistema adicional de control en el interfaz de usuario, el cual administra también el control general del módem: elección del canal y de velocidad de baudios, conversión arrítmico/síncrono y establecimiento de bucles de prueba.

Sistema transmisor

El transmisor se encarga de la aleatorización, la codificación diferencial, la modulación de amplitud en cuadratura, la conformación, la generación del tono de guarda y la sincronización externa.

Se han podido eliminar todas las multiplicaciones en tiempo real mediante el uso de una tabla de consulta, que contiene los valores pre-multiplicados almacenados en ROM. De esta forma, las funciones del transmisor se pueden realizar mediante un único microprocesador de 8 bits de uso general.

La modulación de amplitud en cuadratura y la conformación por filtro transversal de respuesta impulsiva finita se realizan codificando los datos entrantes como indicadores de la tabla, y acumulando los valores pre-multiplicados seleccionados de la tabla a lo

largo de la longitud de símbolo del filtro. Los tonos de guarda de 550 y 1800 Hz se generan añadiendo los valores de una tabla de senos a la suma acumulada. El resultado se introduce en un convertidor digital-analógico de 12 bits a una velocidad de muestreo de 9600 Hz, generándose así el espectro transmitido (Fig. 3).

La sincronización con un reloj exterior se realiza con dos temporizadores en el circuito integrado, que trabajan como un bucle de enganche de fase para impedir que las fluctuaciones del reloj distorsionen la señal transmitida.

Sistema receptor

El receptor está dividido en dos subsistemas; uno realiza la igualación adaptativa y el otro se ocupa del CAG, de la recuperación de portadora, demodulación, decodificación y desaleatorización.

La capacidad del igualador para compensar la distorsión lineal depende de la longitud de la respuesta impulsiva del igualador (número de coeficientes). Otro factor que afecta a la distorsión residual es la longitud de la palabra utilizada en los cálculos.

Al igualador se le autoriza un tiempo de convergencia limitado a 500 ms por la secuencia de protocolo V.22 bis. Después de este periodo, el igualador se actualiza continuamente con el fin de seguir las variaciones de la distorsión lineal durante la transmisión de datos. Para cumplir los requisitos de tiempo de convergencia y minimizar la distorsión residual, es necesario poder realizar multiplicaciones con bastante rapidez y elevada resolución, y disponer de una zona de RAM suficientemente grande para contener las muestras de la señal recibida y los coeficientes variables. La solución elegida fue un microprocesador con 256 octetos de RAM interna, capaz de multiplicar dos palabras de 16 bits en $8 \mu\text{s}$, y con una arquitectura interna de 16 bits. Se utiliza el mismo tipo de procesador para realizar las restantes funciones del receptor.

Sistema de control

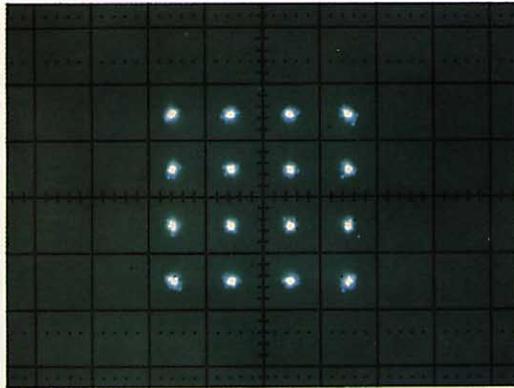
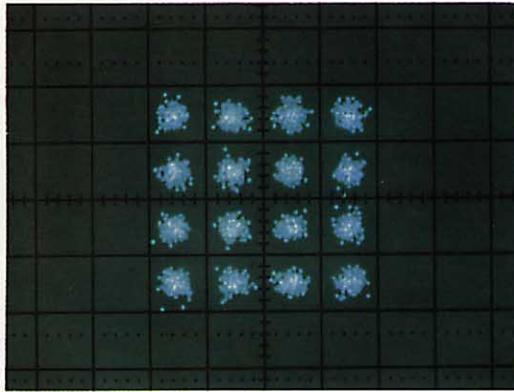
El procesador de control se relaciona con el usuario y gobierna los sistemas transmisor y receptor. Las funciones de la unidad de control requieren capacidad de temporización, facilidades de entrada/salida y memoria RAM. Para el sistema de control se ha seleccionado el mismo tipo de procesador en un solo circuito integrado que en el transmisor.

Conclusiones

Las limitaciones de anchura de banda impuestas a las señales transmitidas por los

* Marca registrada del Sistema ITT

Figura 4
Puntos objetivos de la señal modulada en amplitud y en cuadratura, una vez demodulada: durante la convergencia del igualador (izquierda), y después de la convergencia del igualador (derecha).



canales telefónicos de voz, indican que la velocidad de cálculo necesaria para su proceso digital es ya asequible con la tecnología actual.

El proceso digital de señales ha demostrado ser un poderoso medio para tratar los problemas de distorsión que aparecen al transmitir datos por la red telefónica. El uso de un igualador adaptativo ha hecho posible que el módem 2284 logre una tasa de error de bitio menor de 10^{-5} para una relación señal-ruido de 15 dB (ancho de banda de ruido de 3 kHz), aun en líneas con fuerte distorsión.

La figura 4 muestra los puntos objetivo en la señal demodulada (modulación de amplitud en cuadratura) durante y después de la

convergencia del igualador. La señal se ha transmitido por una línea que consta de siete secciones de portador, 140 km de cable cargado (132 mH/1600 m), y 17 km de cable local de 0,7 mm.

La flexibilidad del proceso digital de señales es una ventaja esencial, ya que permite utilizar el mismo soporte físico para diferentes funciones del módem, como es la reversión del módem 2284 al modo de 1200 bit s^{-1} . Además, minimiza la necesidad de ajustar los componentes durante su producción y simplifica las comprobaciones.

Bibliografía

- 1 D. D. Falconer: Jointly Adaptive Equalization and Carrier Recovery in Two-Dimensional Digital Communication Systems: *Bell System Technical Journal*, marzo 1976, volumen 55, nº 3, págs. 317–334.
- 2 D. D. Falconer: Analysis of a Gradient Algorithm for Simultaneous Passband Equalization and Carrier Phase Recovery: *Bell System Technical Journal*, abril 1976, volumen 55, nº 4, págs. 409–428.
- 3 K. H. Mueller y J. J. Werner: An Efficient Passband Equalizer Structure for Data Transmission: *Institute of Electrical and Electronics Engineers Transactions on Communications*, marzo 1982, volumen Com-30, nº 3, págs. 538–541.
- 4 Recomendación V.22 bis del CCITT.

Göran Bång nació en 1943 en Falun, Suecia. Se graduó MSc en ingeniería eléctrica, en 1967, por el Real Instituto de Tecnología de Estocolmo. En 1968 se incorporó a la división de transmisión de SRT, donde participó en varios proyectos de módems de datos. A continuación, pasó algún tiempo trabajando en el grupo de módems de STL en Harlow, en el desarrollo de igualadores adaptativos automáticos. Actualmente el Sr. Bång trabaja como ingeniero superior en SRT, en proyectos que utilizan proceso digital de señales.

Peter Carlsson nació en 1957 en Estocolmo, Suecia. Se le otorgó el MSc en ingeniería eléctrica en 1981 por el Real Instituto de Tecnología de aquella ciudad. Ese mismo año se incorporó a la división de transmisión de SRT, donde ha estado trabajando en la aplicación de proceso digital de señales en módems de datos.

Compensador de eco para módems de datos a dos hilos

Es posible transmitir datos en dúplex a dos hilos merced a las técnicas de compensación de eco. Se aplica proceso digital de señales al diseño de un compensador de eco para módems de datos a dos hilos, del cual se ha desarrollado un prototipo para la red pública de datos nórdica.

T. Alvestad
T. J.-C. Eriksen

Centro de Investigación de
Standard Telefon og Kabelfabrik A/S, Oslo,
Noruega

Introducción

Las redes de cables de pares de cobre trenzados, de gran utilización en Noruega y muchos otros países para comunicaciones locales, requieren un par de hilos por cada abonado. El número de pares instalados se basa en predicciones sobre las líneas analógicas de teléfonos que necesitará una determinada área.

Sin embargo, estas redes fueron instaladas para transmisión analógica, mientras que hoy ya se prevé que en los próximos años la transmisión digital predomine en las redes de telecomunicaciones. Como ilustración del crecimiento de las comunicaciones digitales, STK (Standard Telefon og Kabelfabrik) ha fabricado unos 40.000 módems de banda base (módems para líneas alquiladas) desde su introducción en 1972.

Cuando se creó la red pública de datos nórdica (RPDN), se esperaba que en algunas áreas hubiera falta de líneas. Esto sucedía en parte por tener que proporcionar las líneas para datos adicionalmente a las necesarias para cubrir la demanda de líneas telefónicas analógicas, y en parte porque la transmisión digital requería cuatro hilos (dos pares) por cada conexión. Ahora, sin embargo, las nuevas instalaciones pueden utilizar módems a dos hilos provistos de compensador de eco. El coste de añadir uno de tales módems queda normalmente contrapesado por el ahorro de utilizar dos hilos en vez de cuatro.

En el mercado noruego, STK suministra placas de módems de datos para insertar en equipos terminales de circuito de datos¹ fabricados por Standard Radio & Telefon, asociada sueca a ITT. En 1980, el Centro de Investigación de STK comenzó el desarrollo

de un módem a dos hilos para reemplazar los existentes a cuatro hilos, siendo el objetivo diseñar el lado de abonado del par del módem que conecta el equipo terminal de circuito de datos a la red pública de datos nórdica. Las especificaciones técnicas fundamentales se dan en la tabla 1. Se realizaron varias simulaciones a fin de conseguir un diseño que cumpliera todos los requisitos.

Al mismo tiempo que se desarrollaba el módem a dos hilos basado en compensación de eco, se comenzó un diseño basado en la técnica "ping-pong" o de separación en el tiempo, para comparar las dos alternativas. Quedó probado que el método de compensación de eco es el más eficaz en las distancias de transmisión experimentadas.

Tabla 1 — Especificaciones principales para el módem de datos a dos hilos de STK

Longitud máxima de la línea	15 km (usando cable de 0,6 mm de diámetro y 45 nF km ⁻¹)
Velocidades de transmisión	12 kbit s ⁻¹ , 6 kbit s ⁻¹ , 3 kbit s ⁻¹ y 750 bit s ⁻¹ (la velocidad de señalización en línea es siempre de 12 kbit s ⁻¹)
Código de línea bifase en espacio	
Tiempo desde el encendido hasta el comienzo de la transferencia de datos para la máxima distancia	2 a 3 segundos
Control automático de ganancia e igualador automático	
Interfaz con 5V TTL para la lógica del equipo terminal de circuito de datos	
La correlación entre las señales transmitidas y recibidas no afecta a los coeficientes del filtro	

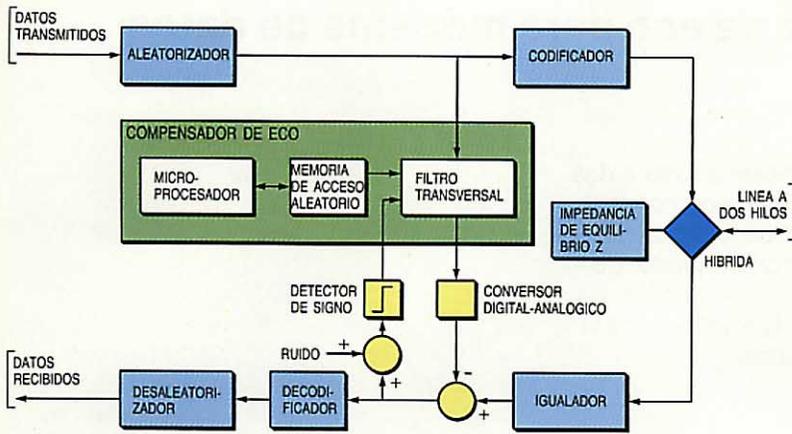


Figura 1
Diagrama de bloques del módem de datos a dos hilos.

Descripción funcional

La figura 1 presenta un diagrama de bloques del módem de datos a dos hilos, que cumple todos los requisitos para transmisión dúplex de datos en banda base sobre líneas a dos hilos a una velocidad de 12 kbit s⁻¹.

Antes de ser enviados a la línea, los bits de datos se aleatorizan y codifican adecuadamente. La aleatorización es necesaria para reducir la probabilidad de correlación entre las señales transmitida y recibida, ya que esta correlación podría interferir el correcto funcionamiento del compensador de eco. No obstante, el compensador desarrollado por STK incorpora una opción para insensibilizarse a tal correlación.

El código de línea es el bifase en espacio (Fig. 2), un código de dos niveles caracterizado por unas buenas propiedades de sincronización y componente continua. La señal de datos codificada se transmite a línea a través de una bobina híbrida convencional. Se ha optimizado la impedancia de equilibrio para líneas largas (15 km de cable de 0,6 mm de diámetro), con una previsión de atenuación de 10 a 20 dB. A pesar de la óptima adaptación, una parte sustancial de la señal a transmitir se introduce en el

receptor e interfiere a la señal recibida de línea.

Esta interferencia, o señal de eco, puede ser mucho mayor que la señal del extremo lejano. De hecho, la relación entre esta señal y la señal de eco es del orden de -40 dB cuando el módem opera a la máxima distancia (15 km), en virtud de lo cual hay que atenuar la señal de eco en más de 60 dB para lograr una relación señal-ruido de 20 dB. Esta es la misión del compensador de eco, cuyo bloque funcional se incluye en la figura 1.

El compensador de eco se basa en técnicas de proceso digital de señal. Consiste en un filtro transversal adaptativo que simula el trayecto de transmisión del eco desde el transmisor, y pasando por la híbrida conectada a la línea de abonado, hasta la entrada del receptor. Si dicho camino es lineal, podrá representarlo un filtro transversal con la precisión que se desee, simulando este filtro la no linealidad de la codificación bifase en espacio. Los cambios en la longitud de la línea, en sus parámetros, y las reflexiones causadas por desadaptaciones, son automáticamente compensadas por ser el filtro adaptativo.

Las muestras de compensación (eco simulado por el compensador) se generan, mediante técnicas de proceso digital de señal, a partir de los datos transmitidos y los coeficientes almacenados, pasando luego por un convertor digital-analógico cuya salida se resta de la señal recibida. Esto se realiza cuatro veces en cada intervalo de bitio, obteniéndose cuatro muestras de la señal del extremo lejano (Fig. 2), de las cuales se extraen los datos y la temporización. Después de la compensación, la señal contiene un error que no está correlado con la señal lejana y puede, por tanto, utilizarse para ajustar los coeficientes del filtro.

Filtro adaptativo

Soporte físico

El compensador de eco puede dividirse en tres partes principales, como se muestra en la figura 1:

- filtro transversal, que opera en tiempo real
- RAM para almacenar los coeficientes del filtro
- microprocesador (Intel 8085), para actualización de dichos coeficientes.

En la figura 3 se indica la estructura del filtro y el bucle de ajuste. Nótese que el ajuste del filtro, efectuado por el microprocesador, no es un proceso en tiempo real. El filtro hace uso del conjunto de coeficientes

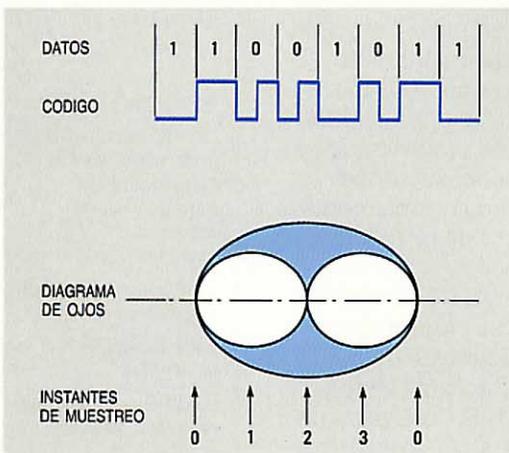
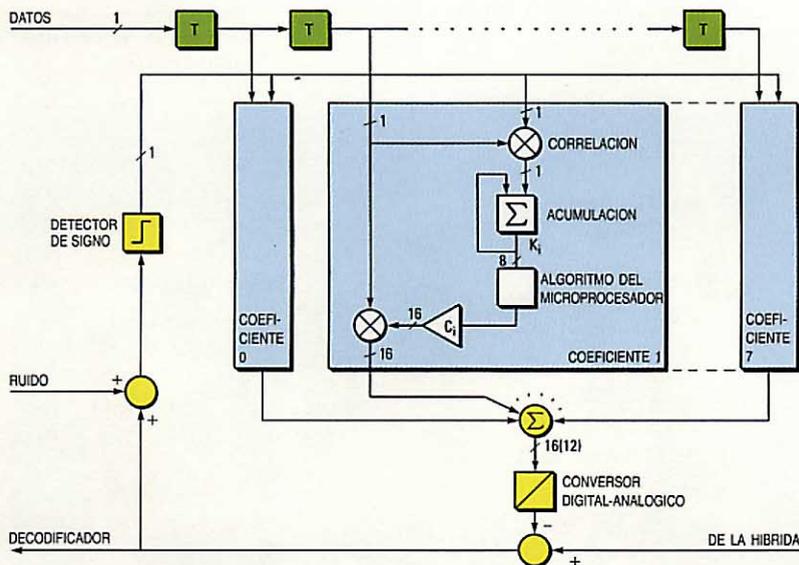


Figura 2
Código de línea bifase en espacio.

anteriores mientras se calculan unos nuevos.

Los circuitos aritméticos realizan dos operaciones principales: compensación y correlación. El filtro consta de 8 etapas, pero como el filtrado se realiza 4 veces por intervalo de bitio, existe un total de 32 coeficientes. Para cada coeficiente (C_i) hay un coeficiente de correlación (K_i); unos y otros

Figura 3
Filtro transversal adaptativo.
 T - intervalo de bitio
 C_i - coeficiente de compensación
 K_i - coeficiente de correlación.



se almacenan en secciones distintas de la RAM.

Durante la compensación, el dato (1 bitio) se multiplica por el correspondiente coeficiente (16 bits), leído de la RAM. Dicha multiplicación es una operación de O-exclusivo. Para todas las etapas se utiliza el mismo circuito (cálculos en serie), a fin de minimizar el equipo necesario. En un acumulador se suman los resultados de las multiplicaciones y se genera la muestra final de compensación con una resolución de 16 bits.

En el módem se emplea un convertor digital-analógico de 12 bits, que desecha los cuatro bits menos significativos. No

obstante, el compensador en sí trabaja con resolución de 16 bits.

La correlación consiste en un O-exclusivo entre el signo del error y los datos transmitidos. El resultado de esta correlación (-1 ó +1) se suma al valor del coeficiente de correlación anterior. Se ha demostrado que debe añadirse ruido incoherente y controlado a la señal recibida para tener una buena estimación del error de compensación cuando la resolución es sólo de 1 bitio (signo)². El ruido añadido debe tener una distribución de probabilidad rectangular. También se calcula la correlación entre los datos transmitidos y recibidos, y si resulta demasiado elevada, una facilidad incorporada en el sistema inhibe la actualización de los coeficientes. No obstante, esta correlación se minimiza aleatorizando los datos transmitidos y recibidos mediante distintos polinomios.

Cuando se detecta un desbordamiento (± 127) en alguno de los coeficientes de correlación, se interrumpe al microprocesador y se calcula un nuevo conjunto de coeficientes que se escriben en la sección de correlación de la RAM. Con ello no se perturba el funcionamiento normal del filtro. Las dos secciones se intercambian luego y el filtro opera con el nuevo conjunto de coeficientes.

Soporte lógico

El soporte lógico (programación) se divide en unas 25 subrutinas modulares, que utilizan lenguaje ensamblador a fin de aprovechar mejor el limitado espacio de memoria.

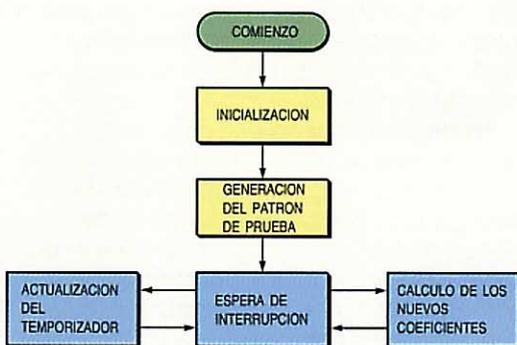
Pasada la inicialización, el filtro en tiempo real se carga con un conjunto de coeficientes de prueba con objeto de verificar su correcto funcionamiento, reponiéndose después. La figura 4 muestra cómo se genera la programación.

El siguiente módulo de programa a ejecutar es el que selecciona uno de los dos igualadores y ajusta la ganancia. Esto forma parte de la inicialización del soporte físico, y no debe corregirse una vez comenzada la transferencia de datos. Al revés de lo que sucede en la transmisión a cuatro hilos, una pequeña variación en la fase o en la ganancia afecta mucho a las prestaciones del módem. Una variación lenta no presenta ningún problema, por cuanto el compensador de eco es adaptativo.

La igualación y la ganancia se escogen sólo una vez, por lo que hay que cuidar de que los espúreos de la línea no ocasionen un ajuste incorrecto.

Inicializados ya los circuitos y la programación, el procesador espera una de las dos interrupciones posibles. La primera proviene de un temporizador de 10 ms, que

Figura 4
Flujo de la programación para el filtro adaptativo.



regula el tiempo que tarda en saturarse un registro de coeficientes de correlación. La segunda es provocada por saturación de uno de tales registros, que hace detenerse al temporizador e indicar lo que puede llamarse el tiempo de correlación. Este tiempo es muy corto cuando los coeficientes difieren mucho de su valor correcto, como sucede en el arranque, y bastante más largo cuando los coeficientes se aproximan a sus valores óptimos.

El algoritmo de un bitio podría ser una solución de compromiso entre rapidez de adaptación y precisión (con la máxima longitud de la línea). Para obviar este problema en algunos diseños se efectúa un ajuste rápido tras el encendido, reduciendo posteriormente, y de forma gradual el paso de adaptación. El inconveniente de esta solución es su lenta adaptación tras la fase inicial.

Debido al uso de un procesador para calcular los nuevos coeficientes en tiempo diferido, se ha elegido un algoritmo más elaborado, que no sacrifica la rapidez de adaptación a la precisión. Sus variables de entrada son las siguientes:

- coeficientes anteriores
- coeficientes de correlación nuevos, uno de los cuales está saturado (± 127)
- tiempo de correlación
- valor de pico del ruido añadido actual.

Como base para el paso de adaptación, se calcula un valor que es proporcional al coeficiente de correlación y al valor de pico de ruido, e inversamente proporcional al tiempo de correlación. Si este valor fuera bajo, se ajustaría conforme a su cuadrado el incremento del coeficiente correspondiente, mientras que si fuese grande, se escogería una relación lineal. De ese modo, un coeficiente se va ajustando más lentamente al acercarse a su valor óptimo. Por otra parte, se obtiene una respuesta rápida cuando se necesita.

Tras el cálculo de los incrementos, éstos se suman a los 32 coeficientes que les corresponden. Antes de almacenar en la RAM el nuevo conjunto de coeficientes se comprueba si existe algún desbordamiento. Los coeficientes de correlación y el temporizador se vuelven al estado inicial y se ajusta el ruido, antes de completarse una iteración.

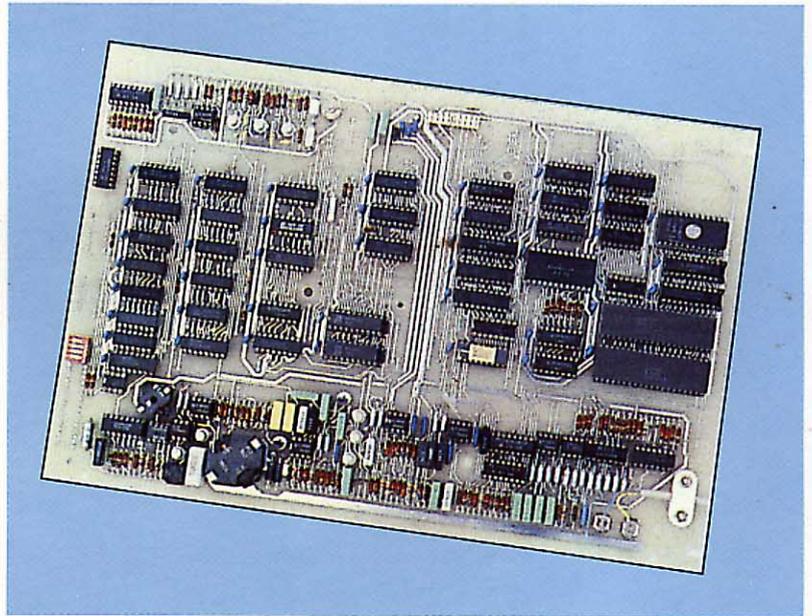
Prestaciones

Varias placas prototipo del módem a dos hilos con compensador de eco se han sometido a pruebas de laboratorio, en un entorno de red real. El programa de pruebas, establecido por la Administración

noruega de Telecomunicaciones, incluía medidas de alcance de transmisión, tasa de errores de bitio en función de diferentes fuentes de ruido, e interrupciones de línea simuladas. Estas pruebas demostraron que las prestaciones del módem cumplen todas las especificaciones:

- 15 km de alcance de transmisión, con cable de 0,6 mm de diámetro (50 dB de atenuación) y tasa de error de 10^{-7}
- 2 a 3 s de tiempo de convergencia del compensador de eco, a partir del encendido

Placa prototipo del módem de datos a dos hilos de STK.



- sincronización de bitio mantenida para interrupciones menores de 200 ms.

El comportamiento del algoritmo de correlación del signo utilizado para ajustar el filtro transversal adaptativo ha sido satisfactorio, con resultados que concuerdan con las simulaciones por ordenador.

Además, el módem ha sido probado y comparado con otros módems banda base a dos hilos dentro del programa de pruebas europeo de La Haya, organizado por la CEPT (Conferencia Europea de Correos y Telecomunicaciones). La finalidad de estas pruebas fue obtener información útil para la futura normalización de los interfaces de módems de datos a dos hilos en redes públicas de datos.

El compensador de eco que hemos descrito se está utilizando también en una nueva versión a dos hilos del módem banda base DCB 19200 de STK. La parte de tiempo real del filtro adaptativo — los circuitos aritméticos — se ha montado con componentes estándar TTL Schottky de bajo consumo. Esta parte es muy adecuada para

realizarse mediante un circuito LSI de diseño a medida, reduciendo así el soporte físico necesario y consiguiendo un compensador de eco más compacto.

Conclusiones

Se ha desarrollado un nuevo compensador de eco que, mediante técnicas de proceso digital de señales, permite la transmisión de datos dúplex a dos hilos. Este compensador, basado en un filtro transversal adaptativo, se está utilizando en módems a dos hilos desarrollados por STK. Las pruebas realizadas demuestran que el módem satisface todas las especificaciones.

Referencias

- 1 T. Hedberg: Equipo terminal de circuito de datos flexible para redes de datos por conmutación de

circuitos: *Comunicaciones Eléctricas* (1983), volumen 58, n° 2, págs. 176—180.

- 2 E. Jørgensen y K. O. Kjølås: Echo Canceling System Based on the Sign Correlation Algorithm: National Telecommunications Conference 1981.

Torgeir Alvestad nació en Bø i Telemark, Noruega, en 1955. Estudió ingeniería electrónica en la Universidad Técnica de Trondheim, graduándose en 1980. Seguidamente se incorporó al Centro de Investigación de STK, trabajando en transmisión digital y proceso digital de señales en relación con módems de datos. El Sr. Alvestad se dedica actualmente a terminales y líneas de abonado para la Red Digital de Servicios Integrados (RDSI).

Tom Jens-Christian Erksen nació en Oslo, Noruega, en 1950. Estudió ingeniería electrónica en la Universidad Técnica de Trondheim, graduándose en 1973. Cumplió su servicio nacional en el Centro de Investigación de la Defensa Noruega antes de incorporarse al Elektrisk Bureau en 1975. El Sr. Erksen ingresó en el Centro de Investigación de STK en 1980, como jefe de proyecto para transmisión digital. Actualmente tiene a su cargo el grupo de terminales de RDSI.

Compensación de eco adaptativa para transmisión de datos en banda base

El desarrollo de nuevos algoritmos y circuitos más refinados abre la posibilidad de aplicar el proceso digital de señales al campo de las comunicaciones. En la actualidad se dispone de varios algoritmos con diferentes características en lo que respecta a velocidad de convergencia, error en régimen permanente, complejidad de cálculo y de realización física.

M. Gilsanz

F. Pedrón

J. A. Siles

Centro de Investigación de Standard Eléctrica S.A., Madrid, España

Introducción

En transmisión a dos hilos, el transmisor y el receptor están acoplados a la línea por medio de una bobina híbrida, construida de forma que el paso de señal del lado transmisor al receptor sea teóricamente cero, suponiendo una perfecta adaptación entre la impedancia de la línea y la impedancia de equilibrio de la híbrida (Fig. 1). Sin embargo, en la práctica, la impedancia de la línea es diferente en cada conexión, y puede variar incluso durante una comunicación establecida provocando un paso de señal indeseado del lado transmisor al receptor, denominado eco próximo.

Otro tipo de eco, el eco lejano, se produce por cambios de diámetro en la línea, empalmes, e imperfecta adaptación en la híbrida del extremo remoto (Fig. 1).

En comunicaciones de un solo sentido, o bidireccionales semidúplex, los problemas de ecos se pueden solventar abriendo el camino de recepción, como suele hacerse en los supresores de eco. Sin embargo, esto no es posible cuando la conexión es

dúplex total y hay que recurrir entonces a compensadores de eco, basados en simular localmente el trayecto del eco. La señal de eco estimada se resta de la señal entrante, quedando la señal transmitida por el extremo distante limpia de esta señal perturbadora (Fig. 2).

Desgraciadamente la realización práctica del compensador de eco no es tan fácil, por ser el trayecto del eco desconocido y variable de una conexión a otra, o incluso durante una comunicación. La elección de un filtro simulador con características fijas basado en el promedio estadístico de canales no es del todo satisfactoria, dejando como única solución posible el uso de filtros con coeficientes adaptativos^{1, 2}.

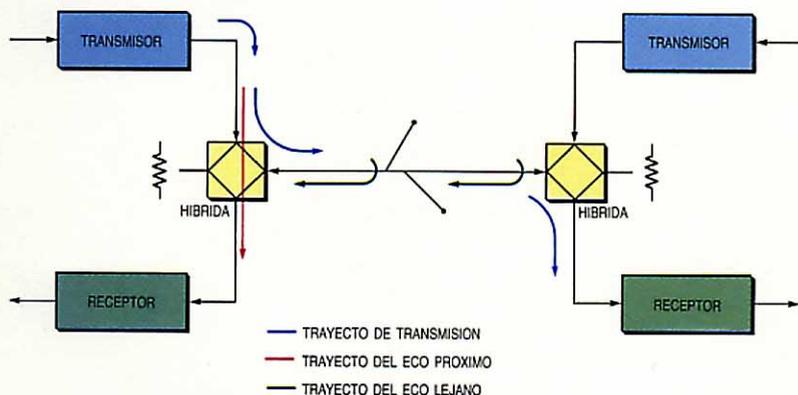
Requisitos del compensador de eco

El eco próximo tiene un retardo muy corto y un nivel de señal elevado. La atenuación típica de la señal en la híbrida es de unos 10 dB, mientras que la atenuación en la línea puede llegar a 40 ó 50 dB. Por consiguiente, el nivel del eco puede superar en 30 ó 40 dB el de la señal entrante. Como se requiere una relación señal-ruido de al menos 20 dB, el compensador deberá reducir el nivel del eco en unos 50–60 dB.

El eco lejano es una señal con un nivel muy bajo, pero con un retardo apreciable (del orden de milisegundos) que depende de la longitud de la línea y de su respuesta impulsiva. En este caso el requisito principal del compensador es una respuesta larga o retardada, en vez de una reducción importante en el nivel de la señal.

En transmisión en banda base, el trayecto del eco se puede considerar lineal, con excepción de ciertas faltas de linealidad³:

Figura 1
Diagrama de bloques de un sistema de transmisión de datos dúplex a dos hilos.



- no linealidad de la conversión analógico/digital
- asimetría entre los impulsos positivos y negativos transmitidos
- saturación en la bobina de acoplamiento.

Tales irregularidades deben corregirse con el compensador de eco o con circuitos adicionales, ya sea aumentando el nivel de compensación, o teniendo en cuenta las faltas de linealidad al diseñar el algoritmo de compensación.

La compensación del eco, que en la realidad será sólo parcial, produce una señal residual que se puede considerar como un ruido a sumar a la señal proveniente del extremo lejano. El efecto es reducir la apertura del ojo, que podría llegar a ser insuficiente para una buena recepción.

Debe también tenerse en cuenta la aritmética de precisión finita en la realización del compensador de eco. Además de las entradas y salidas de los conversores analógico-digital y digital-analógico, hay que examinar en detalle el algoritmo a fin de determinar los pasos más críticos en lo que atañe a la precisión, sobre todo cuando se utilizan multiplicadores.

Se deben establecer compromisos entre complejidad de cálculo, velocidad de convergencia y ruido alrededor del punto de convergencia. El primer factor es cada vez menos crítico, debido a los avances tecnológicos. En cuanto a la velocidad de convergencia, su aumento provoca en muchos algoritmos un aumento del error en régimen permanente; habrá que evaluar, pues, la importancia relativa de cada factor.

Algoritmos adaptativos

Los algoritmos adaptativos siguen un criterio de minimización, con una señal de error relacionada con el grado de adaptación. Uno de los criterios más usuales es el de error cuadrático medio, ya que las ecuaciones que de él se obtienen son de una complejidad limitada y resiste suficientemente a pequeños cambios en torno al punto óptimo sin degradar seriamente el comportamiento del sistema.

En este artículo se examinan los siguientes algoritmos: gradiente, signo, mínimos cuadrados (Kalman), Kalman rápido y celosía. La señal de error que va a utilizarse siempre es (Fig. 2):

$$e(j) = f(j) + c(j) + n(j) - \hat{c}(j) \quad (1)$$

donde $c(j)$ es la señal de eco, $\hat{c}(j)$ la señal de eco estimada, $n(j)$ el ruido aditivo, y $f(j)$ es la señal transmitida por el extremo lejano.

Si no hubiera señal del extremo lejano, la potencia del error se aproximaría asintótica-

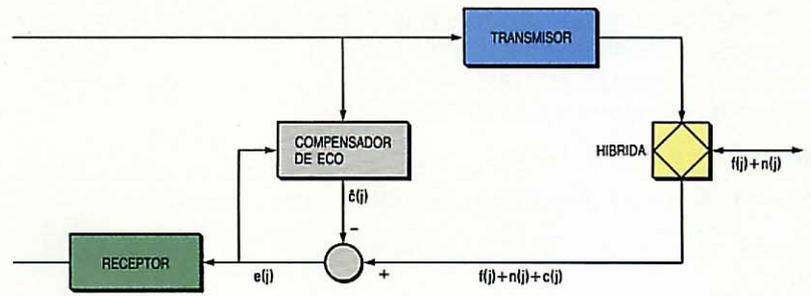


Figura 2
Configuración básica de un compensador de eco para un sistema de transmisión en dúplex total.

mente a cero. No ocurrirá esto si en la entrada del compensador están presentes la señal distante y/o el ruido aditivo.

Algoritmo de gradiente

Este algoritmo es el más utilizado, basado en el mínimo de la señal de error cuadrático medio. Los coeficientes del filtro se actualizan mediante la ecuación:

$$\mathbf{a}(n+1) = \mathbf{a}(n) - \mu \nabla \{E[e^2(n)]\} \quad (2)$$

donde $\mathbf{a}(n)$ es el vector de coeficientes en el instante n , μ es la constante de adaptación y ∇ el gradiente de la esperanza del error cuadrático.

En lugar del valor promedio del error, se hace una primera aproximación tomando el valor instantáneo del error. En este caso, y suponiendo un filtro de respuesta impulsiva finita (RIF), la ecuación (2) se convierte en:

$$\mathbf{a}(n+1) = \mathbf{a}(n) + 2\mu e(n) \mathbf{x}(n). \quad (3)$$

Dado que la superficie de error $E[e^2(n)]$ es parabólica y tiene sólo un mínimo, la expresión (3) convergerá siempre a ese mínimo, y el error en régimen permanente dependerá del valor de μ utilizado. Dicho valor está comprendido entre cero y el inverso del mayor autovalor de la matriz de autocorrelación de la señal. Se sabe también que, cuanto mayor es la dispersión de los autovalores (la relación entre el mayor y el más pequeño de los mismos), más lenta es la convergencia de los coeficientes.

Normalmente μ es constante, aunque a veces tiene un valor mayor al principio y uno más pequeño en las proximidades del mínimo. Esto aumenta la velocidad de convergencia y reduce el error en régimen permanente.

Si se utiliza un filtro RII (respuesta impulsiva infinita) en lugar de un filtro RIF, la ecuación (2) ya no es válida. En este caso el error es función de la señal entrante y de los valores anteriores de la señal de salida. Por lo tanto, la diferenciación de la señal de error con respecto a los coeficientes del numerador y del denominador da como resultado funciones cuadráticas que, por supuesto, no son lineales.

Se han desarrollado algunos algoritmos para resolver este problema, tales como el de Feintuch⁴ y el de Horvath⁵. Para este último las ecuaciones de actualización son:

$$\mathbf{a}(n+1) = \mathbf{a}(n) + \mu_1 [-2e(n) \nabla e_a(n)] \quad (4)$$

$$\mathbf{b}(n+1) = \mathbf{b}(n) + \mu_2 [-2e(n) \nabla e_b(n)] \quad (5)$$

donde, para un filtro RII de M ceros y N polos:

$$\nabla e_a^T(n) = \left[\frac{\partial e(n)}{\partial a_0} \quad \frac{\partial e(n)}{\partial a_1} \quad \dots \quad \frac{\partial e(n)}{\partial a_M} \right] \quad (6)$$

$$\nabla e_b^T(n) = \left[\frac{\partial e(n)}{\partial b_1} \quad \frac{\partial e(n)}{\partial b_2} \quad \dots \quad \frac{\partial e(n)}{\partial b_N} \right] \quad (7)$$

(T quiere decir traspuesto).

La obtención física de estas derivadas es directa, como se puede ver en la figura 3.

La superficie de error no tiene, en este caso, un mínimo único. Se alcanzará uno de los posibles mínimos, pero no se asegura que sea el menor de todos.

La principal ventaja de un filtro RII es que su respuesta impulsiva es más larga que la de un filtro RIF para igual número de coeficientes. Ambos tipos de filtro se están utilizando conjuntamente en compensadores de eco para aprovechar las ventajas de cada uno. El filtro RIF sirve para compensar el eco próximo, más fuerte y sin retardo, mientras que el RII se utiliza para compensar el eco lejano. En la figura 4 se presenta la realización práctica.

Algoritmo del signo

Este algoritmo es una versión simplificada del algoritmo de gradiente, ya que usa el signo del error instantáneo en vez de su valor. La ecuación (3) se transforma en

$$\mathbf{a}(n+1) = \mathbf{a}(n) + 2\mu [\text{sgn}/e(n)] \mathbf{x}(n). \quad (8)$$

La introducción de la operación signo en (8) dificulta el análisis teórico del comportamiento del algoritmo, y reduce las prestaciones, ya que la convergencia es más lenta y la potencia del error residual no disminuye aunque la señal distante se haga más pequeña.

Como estos algoritmos son muy atractivos desde el punto de vista de cálculo, se han estudiado recientemente modificaciones para mejorar su comportamiento general.

Algoritmo de mínimos cuadrados (Kalman)

En vez de hacer mínimo el error cuadrático promedio, es posible minimizar la suma de los errores instantáneos al cuadrado^{6,7}. En este caso se usa el error cuadrático ponderado:

$$E(n) = \sum_{k=-\infty}^n w(n-k) e^2(k) \quad (9)$$

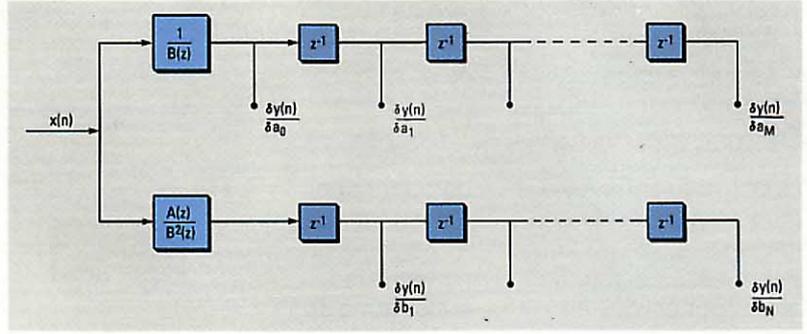


Figura 3 Derivadas parciales para el algoritmo de Horvath.

donde $e(k)$ es el error instantáneo y $w(n)$ la ventana de ponderación. La ventana más utilizada es la exponencial α^n , donde $\alpha < 1$ y $n \geq 0$.

La ecuación de actualización de coeficientes se transforma en:

$$\mathbf{a}(n+1) = \mathbf{a}(n) + \mathbf{k}(n+1) \mathbf{e}(n+1) \quad (10)$$

donde $\mathbf{k}(n+1)$, el vector de ganancia Kalman, se puede obtener de:

$$\mathbf{k}(n+1) = P^{-1}(n+1) \mathbf{x}(n+1). \quad (11)$$

La inversión de la matriz P requiere del orden de $N^3/3$ multiplicaciones y/o divisiones, siendo N el número de etapas del filtro.

Sin embargo, la matriz inversa se puede obtener de la ecuación recursiva:

$$P^{-1}(n+1) = \frac{1}{\alpha} \left[P^{-1}(n) - \frac{P^{-1}(n) \mathbf{x}(n+1) \mathbf{x}^T(n+1) P^{-1}(n)}{\alpha + \mathbf{x}^T(n+1) P^{-1}(n) \mathbf{x}(n+1)} \right] \quad (12)$$

la cual requiere alrededor de $2N^2$ multiplicaciones.

La complejidad de estas operaciones se reduce transmitiendo señales binarias o ternarias. Si pudieran imponerse ciertas condiciones a estas señales, sería posible simplificar los cálculos; por desgracia, no es posible variar a voluntad del diseñador las propiedades de las señales transmitidas, y tales condiciones no se cumplen. No obstante, al principio de una comunicación no se transmiten datos reales y podrían utilizarse secuencias especiales con las propiedades requeridas⁸, demostrando que para señales pseudoaleatorias el sistema converge en N iteraciones.

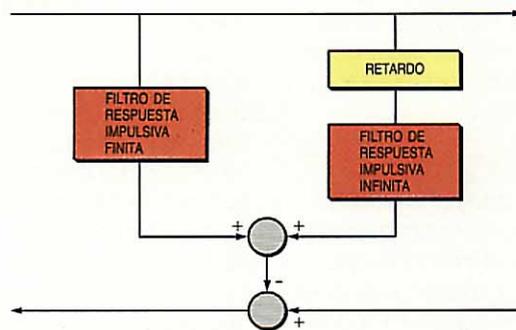


Figura 4 Compensador de eco con filtros RIF y RII.

Optimización del algoritmo de Kalman (Kalman rápido)

El algoritmo de Kalman converge muy aprisa, pero requiere un cálculo considerable. Se optimiza este algoritmo aprovechando que el vector de señal de entrada en cualquier momento es el del instante de muestreo anterior, retardado en una o más muestras y con una o más muestras añadidas. Tratando matemáticamente esta propiedad, el algoritmo de Kalman se realiza con sólo unas $10N$ multiplicaciones y/o divisiones en las mismas condiciones de convergencia.

Filtro en celosía adaptativa

El valor del error a la salida de un filtro RIF viene dado por:

$$e(k) = d(k) - \sum_{i=0}^N a_i(n) b_i(k) \quad (13)$$

donde

- $d(k)$ – señal a estimar
- \sum – respuesta de un filtro RIF con coeficientes a_i y señal de entrada b_i .

Sustituyendo $e(k)$ de (13) en la (9), diferenciando con respecto a a_i e igualando a cero, se tiene una expresión en cuyo miembro izquierdo aparece el vector de correlación cruzada de las señales $d(k)$ y $b(k)$, y en el derecho la suma de los productos de los coeficientes por las funciones de correlación b_i .

Si b_i fueran señales ortogonales, la matriz de autocorrelación sería diagonal y los cálculos se simplificarían mucho. Estas señales ortogonales se pueden obtener de un filtro en celosía, ya que los coeficientes de error hacia atrás cumplen esta condición una vez reducida al mínimo la energía residual del filtro. La ecuación de minimización se convierte en:

$$\mathbf{c}(n) = \mathbf{a}(n) R_{XX} \quad (14)$$

siendo el vector de correlación cruzada

$$\mathbf{c}(n) = \sum_{k=-\infty}^n \alpha^{n-k} d(k) b(k)$$

R_{XX} – matriz de autocorrelación de $b(k)$.

El vector $\mathbf{c}(n)$ se puede obtener recursivamente de la última ecuación. Asimismo el valor de R_{XX} resulta de minimizar de forma recursiva el error residual en el filtro de celosía. La figura 4 es un diagrama de bloques de un filtro en celosía adaptativo.

Si en vez de utilizar la ecuación (9) se minimiza el gradiente del error, se obtiene un conjunto de ecuaciones en el que también están presentes las funciones de autocorrelación; con las señales ortogona-

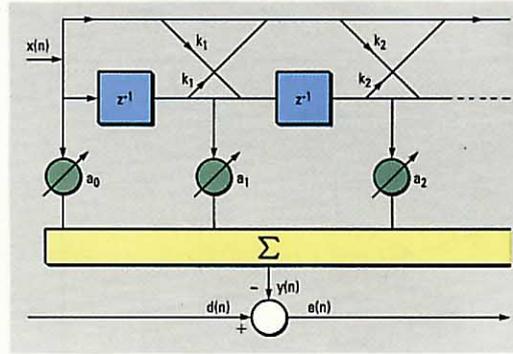


Figura 5
Diagrama de bloques de un filtro adaptativo con estructura en celosía.

les se llega a simplificaciones similares, aunque la convergencia es más lenta.

En ambos casos, los coeficientes de correlación parcial de la celosía y del filtro RIF se adaptan por separado, haciéndolo primero los de la celosía.

Hay todavía otro algoritmo de celosía, la celosía de mínimos cuadrados, en el que los errores se minimizan actualizando a la vez los coeficientes del filtro y de la celosía, lo cual acelera al máximo la convergencia. La figura 6 y la tabla 1 se refieren a los dos últimos casos^{7,9}.

Criterios de diseño para compensadores de eco

El entorno de trabajo del compensador de eco impone una serie de requisitos, tales como la velocidad de transmisión y las particularidades del trayecto del eco. El diseñador puede además seleccionar algunas características, como las siguientes:

- algoritmo de compensación del eco
- arquitectura general, incluyendo la posición de los conversores analógico/digital
- compensación a la velocidad de símbolo o a una fracción de ella
- precisión finita en operaciones digitales y resultados intermedios.

Selección de algoritmos

Para seleccionar un algoritmo deben considerarse primordialmente la velocidad de

Tabla 1 – Comparación de la complejidad de cálculo

Algoritmo	Número de multiplicaciones y/o divisiones por iteración (del orden de)
De signo	N
De gradiente	$2N$
Kalman rápido	$10N$
Celosía de gradiente	$10N$
Celosía de mínimos cuadrados	$20N$
Kalman de mínimos cuadrados	N^2

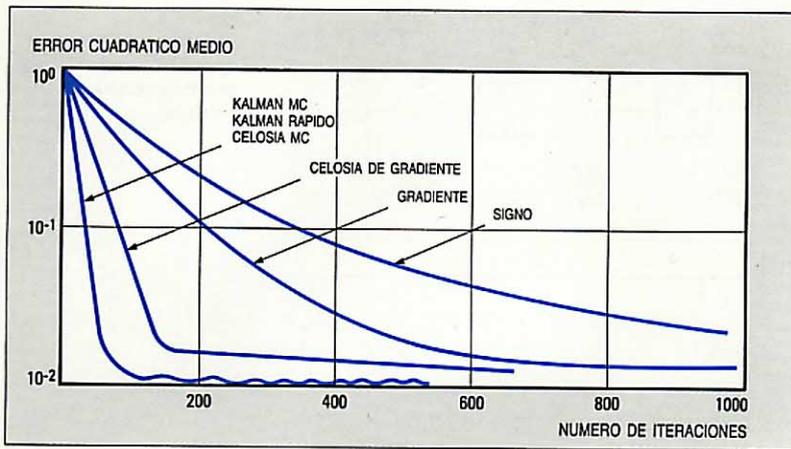


Figura 6
Comparación de las
frecuencias de conver-
gencia para algoritmos
adaptativos.
MC - mínimos
cuadrados.

convergencia, el error en régimen permanente, la complejidad de cálculo y la complejidad del circuito. Aunque estos factores no sean independientes, se estudian por separado para mayor sencillez.

Poco se adelanta utilizando filtros en celosía en un compensador de eco. En efecto, el filtro en celosía se utiliza para descomponer la señal transmitida en componentes incorrelados antes de entrar al filtro adaptativo. En transmisión de datos, éstos se transmiten ya incorrelados, especialmente si hay aleatorizadores. Si la señal transmitida es de voz, las muestras sucesivas suelen estar muy correladas, aunque las características de la correlación varían rápidamente con el tiempo. En el primer caso la celosía es innecesaria, mientras que en el segundo los coeficientes de la celosía necesitan continua adaptación.

El algoritmo de Kalman logra la máxima velocidad de convergencia y el mínimo error en régimen permanente. Los de Kalman y Kalman rápido dan el mismo resultado, salvo en el diferente número de operaciones por iteración. La convergencia puede alcanzarse en unos pocos múltiplos de N muestras, siendo N el número de etapas del filtro. Usando señales pseudoaleatorias durante el periodo de adiestramiento, podría llegarse a la convergencia en N muestras.

Con el algoritmo de gradiente, la velocidad de convergencia se relaciona con las propiedades estadísticas de la señal y la constante de adaptación μ . Si μ se acerca a su valor máximo, la velocidad de convergencia aumenta pero el ruido en régimen permanente es también más alto. Si μ es pequeña, disminuirá el error en régimen permanente pero la convergencia será mucho más lenta. Por tanto hay que adoptar un valor de compromiso para μ , que puede ser más elevado en el periodo de adiestramiento o en un periodo inicial de comunicación semidúplex, pasando a un valor más

pequeño cuando esté presente la señal distante.

El algoritmo de signo es el más lento y con mayor ruido en régimen permanente de todos los algoritmos sujetos a las mismas condiciones. Su atractivo reside en la facilidad de cálculo.

La figura 6 muestra las diferentes velocidades de adaptación en función del número de iteraciones. La tabla 1 resume el número de multiplicaciones y divisiones por iteración que necesita cada algoritmo. En ambas figura y tabla los resultados se relacionan con N , el número de etapas del filtro.

La complejidad del circuito compensador de eco depende de varios factores. Así, pues, la longitud del filtro está relacionada con la longitud del eco, la facilidad de cálculo depende del algoritmo seleccionado, y la velocidad de cálculo es función del número de operaciones por iteración y de la frecuencia de muestreo.

Otra factor es la memoria necesaria, aunque no sea muy importante si se emplean memorias comerciales o un circuito VLSI desarrollado para una aplicación específica. Si se utiliza un procesador de señal comercial, hay que revisar los criterios de diseño, pues algunos tienen limitada memoria interna.

Distintas configuraciones del compensador de eco

Existe tres configuraciones alternativas¹⁰:

- sustracción de señales analógicas
- sustracción de señales muestreadas
- sustracción de señales digitales.

El filtro adaptativo que simula el trayecto del eco es un filtro digital, mientras que la señal entrante es analógica. Las diferentes configuraciones se centran en el interfaz entre la señal analógica y el filtro digital, o en otras palabras, dónde debe situarse la conversión analógico/digital.

En el caso de sustracción de señal analógica, se equipan un conversor digital-analógico y un filtro paso bajo a la salida del filtro adaptativo para producir la señal analógica réplica del eco, la cual se resta de la señal entrante. A la salida del sustractor analógico, un muestreador y un conversor analógico-digital producen la señal de error que necesita el algoritmo de compensación. En este tipo de compensador, a menos que el ancho de banda de la señal esté bien limitado, puede haber insuficiente compensación en ambos extremos de la banda debido al solapamiento. Por esta razón, el filtro tiene requisitos más severos.

En la sustracción de señal muestreada, se necesita anteponer un muestreador al sustractor. La salida digital del filtro adapta-

tivo se debe convertir de digital a analógico, aunque no se necesite filtro paso bajo. A la salida del sustractor, un conversor analógico-digital produce la señal de error digital a partir de la señal de error muestreada. Si el receptor requiere una señal digital, debería bastar con la salida del conversor analógico-digital anterior. Sin embargo, en el caso de receptor analógico, se necesita un filtro paso bajo aunque con exigencias menos severas que los de sustracción analógica. Son adecuados conversores sencillos analógico-digital y digital-analógico, ya que sólo su precisión finita afecta a la compensación de ecos.

En el caso de sustracción de señal digital, hay que muestrear y cuantificar la señal entrante antes de pasar al sustractor, al cual se conecta directamente la salida del filtro adaptativo. Si el receptor es digital, el compensador de eco no requiere ninguna conversión adicional, pero si es analógico necesitará un conversor digital-analógico y un filtro paso bajo. Los requisitos de este filtro no son demasiado exigentes, pero sí es crítico el conversor analógico-digital para la señal entrante, debido a que el proceso de compensación y la señal que llegue al receptor se podrían degradar si el número de bitios y el tamaño del paso cuantificador no fueran correctos.

Selección de la frecuencia de muestreo

Otra alternativa a evaluar es la frecuencia de muestreo y de compensación^{10, 11}. Los datos entran al compensador a la velocidad de símbolo, y el compensador debe entregar señal al sustractor como mínimo a esa velocidad. El muestreador y/o el conversor analógico-digital que preceden al sustractor deben trabajar a la misma velocidad. Esta operación de muestreo debe guardar sincronismo con los símbolos procedentes del extremo lejano, así como el compensador con los símbolos transmitidos, por lo cual los transmisores próximo y lejano requieren relojes sincronizados.

El muestreo se puede hacer a una frecuencia más alta que la de símbolo. Es ventajoso elegir una frecuencia de muestreo igual o mayor que la frecuencia de Nyquist: en efecto, con sustracción analógica, se pueden evitar los problemas de solapamiento de las bandas, y si la sustracción es digital o muestreada, el instante de muestreo puede independizarse de la fase de la señal entrante, con lo que no se necesitará sincronizar los relojes de ambos transmisores.

Esta mayor frecuencia de muestreo implica que el filtro adaptativo (simulador del trayecto del eco) trabaje a frecuencia más alta, como un filtro interpolador, suavizando

los requisitos del filtro paso bajo, en caso de utilizarse.

Para evitar tener diferentes frecuencias de señal a la entrada y a la salida, el filtro interpolador se puede realizar con cierto número de compensadores de eco en paralelo, todos con la misma entrada y con salidas multiplexadas en el sustractor, produciéndose para cada compensador una señal de error diferente.

Los compensadores tratan de reproducir la respuesta impulsiva del trayecto del eco utilizando las muestras obtenidas de la señal entrante. Variando los instantes de muestreo, al par que se mantiene la frecuencia de muestreo, se obtienen distintas respuestas impulsivas. Cada compensador convergerá, pues, de forma independiente, pese a recibir la misma secuencia de datos de entrada. Debe advertirse que el multiplexar los compensadores de eco no afecta a la velocidad de adaptación.

Selección del número de bitios

Otro criterio de diseño importante es la longitud de palabra, o lo que es lo mismo, la exactitud con que se representa cada variable en el compensador¹⁰. Uno de los principales parámetros que influyen en dicha exactitud es el paso de cuantificación del conversor analógico-digital, que debe ser capaz de discriminar el eco residual especificado para el compensador. Una vez definidos este parámetro y el margen dinámico de la señal de eco, se puede elegir el número de bitios del conversor analógico-digital. Para la compensación de más de 50 dB se utilizan conversores de 8 a 12 bitios.

Otros aspectos del compensador de eco, tales como el acumulador, los coeficientes y sus ajustes, y el conversor digital-analógico se pueden especificar con respecto al número de bitios que utilice el conversor analógico-digital^{10, 12}.

También debe tenerse en cuenta la precisión de las operaciones, especialmente las multiplicaciones. En el compensador de eco para datos, no se altera la precisión en las multiplicaciones, ya que al ser los datos +1, -1 ó 0 sólo hay cambios de signo o multiplicación por cero. El problema es más serio en compensadores de eco para voz, donde las multiplicaciones pueden degradar la precisión. En algunos casos se pueden ahorrar multiplicaciones escogiendo valores adecuados para las constantes. Por ejemplo, en el algoritmo de gradiente puede elegirse una potencia de 2 para la constante μ , reemplazando la multiplicación por un desplazamiento, lo que es más breve y evita la pérdida de precisión en el multiplicador.

Conclusiones

A la hora de desarrollar un compensador de eco, el diseñador puede elegir entre muchas opciones, que hay que estudiar con atención para lograr un diseño equilibrado y eficiente. Sin embargo, las continuas mejoras en circuitos integrados y los algoritmos disponibles facilitan cada día más la obtención de compensadores de eco con el comportamiento requerido.

Las realizaciones físicas de los diseños actuales no han pasado más allá de usar los algoritmos de signo o de gradiente, por su relativa sencillez. Sin embargo, no son ya adecuados para alcanzar las prestaciones deseadas en un compensador de eco. Algunos de los nuevos algoritmos reseñados en este artículo prometen un comportamiento mucho mejor para un futuro no muy lejano.

Referencias

- 1 T. A. Claasen y W. F. Mecklenbräuker: Overview of Adaptive Techniques in Signal Processing: Signal Processing II: Theories and Applications: H. W. Schusler (editor), 1983, págs. 747-754.
- 2 C. W. K. Gritton y D. W. Lin: Echo Cancellation Algorithms: *Institute of Electrical and Electronics Engineers Acoustical Speech and Signal Processing Magazine*, abril 1984, págs. 30-37.
- 3 O. Agazzi, D. G. Messerschmidt y D. A. Hodges: Nonlinear Echo Cancellation of Data Signals: *Institute of Electrical and Electronics Engineers Transactions on Communications*, noviembre 1982, volumen COM-30, págs. 2421-2433.
- 4 P. L. Feintuch: An Adaptive Recursive LMS Filter: *Proceeding of the Institute of Electrical and Electronics Engineers*, noviembre 1976, págs. 1622-1624.
- 5 S. Horvath: A New Adaptive Recursive LMS Filter. También S. Cappellini y G. A. Constantinides: *Digital Signal Processing*, Academic Press.
- 6 M. S. Mueller: Least-Squares Algorithms for Adaptive Equalizers: *Bell System Technical Journal*, octubre 1981, volumen 60, págs. 1905-1925.
- 7 J. G. Proakis: *Digital Communications*: McGraw Hill, págs. 412-445.
- 8 J. Salz: On the Start-up Problem in Digital Echo Cancellers: *Bell System Technical Journal*, julio-agosto 1983, volumen 62, págs. 1353-1364.

- 9 B. Friedlander: Lattice Filters for Adaptive Processing: *Proceedings of the Institute of Electrical and Electronics Engineers*, volumen 70, n° 8, págs. 829-867.
- 10 N. A. Verhoeckx, H. v. d. Elzen, F. A. M. Sniijders y P. J. V. Gerwen: Digital Echo Cancellation for Baseband Data Transmission: *Institute of Electrical and Electronics Engineers Transactions on Acoustical Speech and Signal Processing*, diciembre 1979, volumen ASSP-27, n° 6, págs. 768-781.
- 11 D. G. Messerschmitt: Echo Cancellation in Speech and Data Transmission: *Institute of Electrical and Electronics Engineers Journal on Selected Areas in Communication*, marzo 1984, volumen SAC-2, págs. 283-297.
- 12 N. A. Verhoeckx y T. A. Claasen: Some Considerations on the Design of Adaptive Digital Filters Equipped with the Sign Algorithm: *Institute of Electrical and Electronics Engineers Transactions on Communications*, marzo 1984, volumen COM-32, n° 3, págs. 258-266.

M. Gilsanz nació en Segovia, España, en 1948. En 1972 obtuvo el título de Ingeniero de Telecomunicación en la Universidad Politécnica de Madrid, y el año siguiente ingresó en el Centro de Investigación de SESA, trabajando en diferentes proyectos de sistemas de datos. Desde 1980 el Sr. Gilsanz trabaja en el área de proceso digital de señales, siendo nombrado jefe del grupo en 1982.

F. Pedrón nació en Valencia, España, en 1953. En 1976 se graduó Ingeniero de Telecomunicación en la Universidad Politécnica de Madrid. Trabajó después en SITRE en módems para transmisión de datos y control remoto, siendo responsable del grupo. En 1984 el Sr. Pedrón ingresó en el Centro de Investigación de SESA para trabajar en la utilización del proceso digital de señales para la transmisión de datos. También está interesado en la aplicación de estas técnicas a los sistemas de navegación y posicionamiento.

J. A. Siles nació en 1950 en Valladolid, España. Acabó los estudios en la Escuela Técnica Superior de Ingenieros Industriales de Sevilla, especialidad eléctrica, en 1972. En 1973 ingresó en SESA, en el departamento de desarrollo de circuitos electrónicos para las centrales PENTACONTA*, donde llegó a ser responsable de los equipos de señalización multifrecuencia. El Sr. Siles fue transferido al Centro de Investigación en 1982 para trabajar en el grupo de proceso digital de señales, en el cual actualmente se dedica al desarrollo de técnicas de proceso de voz.

Aplicación del proceso digital de señal a las pruebas e interfaz de enlace del Sistema 12

Las técnicas de proceso digital de señales ofrecen ventajosas prestaciones en muchas áreas anteriormente dominadas por la tecnología analógica. Por ello, se han aplicado a la realización del interfaz y la prueba de enlaces en la central internacional Sistema 12 de Acilia, aun siendo su entorno esencialmente analógico.

M. Fucito

G. Mocerino

FACE Sud, Salerno, Italia

Introducción

Con la transmisión se inició una migración desde la tecnología analógica a la digital, firmemente arraigada hoy en el campo de la conmutación. La tecnología digital permite mejorar las prestaciones del sistema, incluyendo facilidades potenciadas, un mantenimiento más completo, supervisión y control más sencillos, equipos de conmutación y transmisión integrados, mayor facilidad y tamaños más reducidos.

Ejemplos en que la tecnología digital ha desplazado a la analógica son los subsistemas de prueba de línea y de enlace, los dispositivos de control de eco, las facilidades de conferencia y los interfaces de señalización dentro de banda asociada al canal. Sin embargo, aun realizándose de modo digital estas funciones, deben poder operar correctamente en un entorno mixto analógico-digital, cumpliendo toda especificación aplicable al correspondiente equipo analógico. El proceso digital de señal (PDS) ofrece, pues, ventajas considerables en una central telefónica, aunque esté conectada a medios de transmisión analógicos.

Una aplicación importante es la central internacional Sistema 12 de Acilia, principal centro operativo de Italcable en las cercanías de Roma¹. Italcable explota todos los servicios intercontinentales de telecomunicación en Italia, y por ello el entorno de la central de Acilia es en su mayoría analógico, garantizando un interfaz eficiente con el mismo las características de la aplicación interurbana internacional del Sistema 12. Los principales requisitos de esta central incluyen:

- Adecuado interfaz con enlaces analógicos y digitales.
- Capacidad para diferentes sistemas de señalización: R2 analógico y digital,

CCITT n° 5 analógico y digital, y CCITT n° 6 y 7 para señalización por canal común.

- Supresión de eco en enlaces internacionales.
- Mayor capacidad de prueba de enlaces, tanto en transmisión como en señalización, sean analógicos o digitales, para garantizar una alta fiabilidad en el servicio.
- Mayores posibilidades de conexión en conferencia múltiple.

Para atender a estas necesidades se ha desarrollado una serie de equipos con PDS, entre los que se incluyen: supresores digitales de eco (SDE), interfaz digital de señalización de línea CCITT n° 5 (ID5), módulo digital de conferencia, módulo de prueba de enlaces y detector digital de voz. Todos estos equipos, enteramente digitales, se instalan como parte de la central, y su desarrollo se ha orientado a potenciar las prestaciones globales, utilizando mejor los medios de transmisión y aumentando la calidad del servicio.

Supresor digital de eco

El SDE (Fig. 1) es un semisupresor de eco enteramente digital (Recomendación G.164 del CCITT), que sobre todo se utiliza en largas conexiones internacionales e intercontinentales para reducir los efectos (eco) de las reflexiones causadas por las bobinas híbridas o conexiones a dos hilos².

Es necesario procesar digitalmente las señales de entrada a transmisión y recepción, a fin de obtener una indicación fiable de la presencia de voz en ambos sentidos de transmisión. El proceso incluye un filtro paso alto que elimina ruidos de baja fre-

cuencia, seguido de un filtro de acentuación que aumenta la sensibilidad a las altas frecuencias y compensa los sonidos fricativos. Las señales así filtradas pasan a través de detectores de envolvente, obteniéndose salidas que corresponden al nivel de habla real, junto con una alta inmunidad al ruido y una rápida detección.

Para poder inhibir el tono remoto, las señales transmitidas y recibidas se combinan e introducen en un detector de tono basado en filtros de banda eliminada. Las rigurosas exigencias en cuanto a velocidad de operación, inmunidad al ruido y compatibilidad con los supresores analógicos de eco reclaman un cuidadoso diseño de la estructura del filtro y de la técnica de estimación de potencia. Todos los filtros incorporan células RII (respuesta impulsiva infinita) de 2º orden para optimizar el compromiso entre prestaciones y costes.

A fin de mejorar la inmunidad al ruido, los detectores de envolvente tienen constantes de tiempo de activación y desactivación controladas independientemente.

El equipo SDE para 30 canales consta de cuatro placas de circuito impreso, y puede asignarse a cualquier módulo de enlaces analógicos o digitales o bien utilizarse como recurso en reserva. La figura 2 muestra una configuración típica.

Se ha prestado especial atención a la fiabilidad, realizando una vigilancia continua

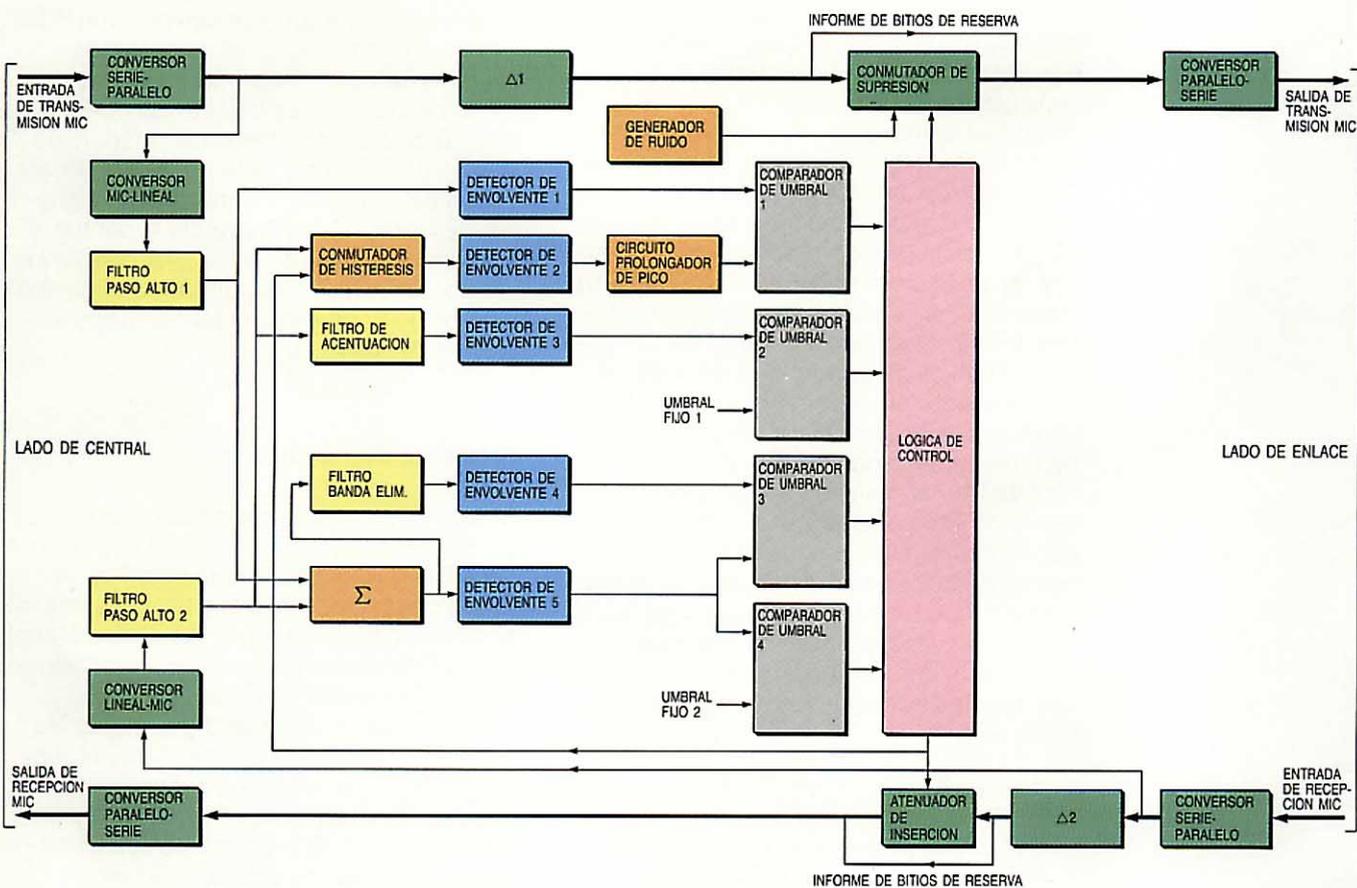
y en línea del equipo físico². Por medio de dos canales de reserva se hace una especie de análisis de firma que detecta cualquier fallo con la mayor celeridad, reduciendo así el tiempo de detección de avería y eliminando la necesidad de rutinas de prueba. La posibilidad de bucle en dúplex a la salida del SDE permite a la central realizar pruebas funcionales totalmente autónomas del SDE, conforme a las recomendaciones del CCITT³.

Interfaz de señalización digital de línea CCITT nº 5

El ID5 consiste en un transmisor y un receptor totalmente digitales para el sistema de señalización nº 5 del CCITT, con dos frecuencias dentro de banda (2400 y/o 2600 Hz) para señalización de línea¹. Vigila continuamente la conversación normal y detecta los tonos de señalización válidos en cualquier enlace CCITT nº 5. Los problemas mayores atañen al margen dinámico, imitación de la señalización por el habla, e inmunidad al ruido. El transmisor de señalización requiere que el nivel y frecuencia de la señal sean estables y precisos, así como la pureza de su espectro.

La figura 3 es un diagrama de bloques del ID5. El lado de transmisión contiene un generador de tono de 2400/2600 Hz con

Figura 1 Diagrama de bloques del supresor digital de eco.



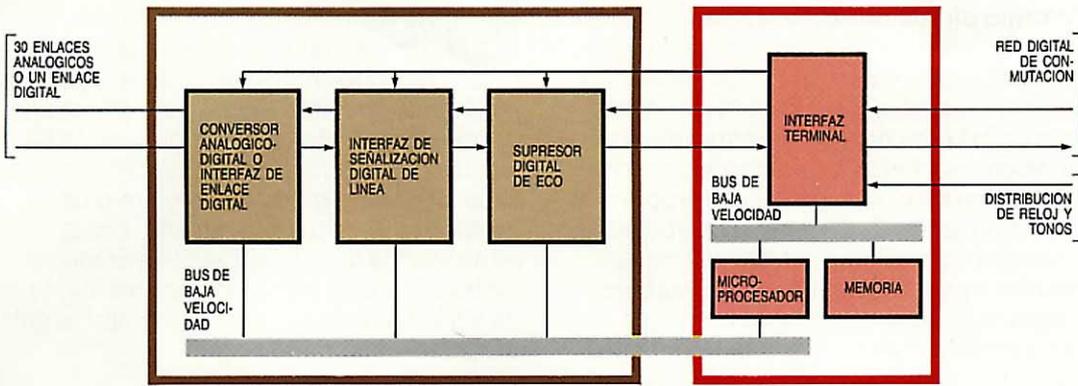


Figura 2
Conexión del supresor digital de eco especializado dentro de los enlaces analógicos y digitales del Sistema 12, con señalización CCITT nº 5.

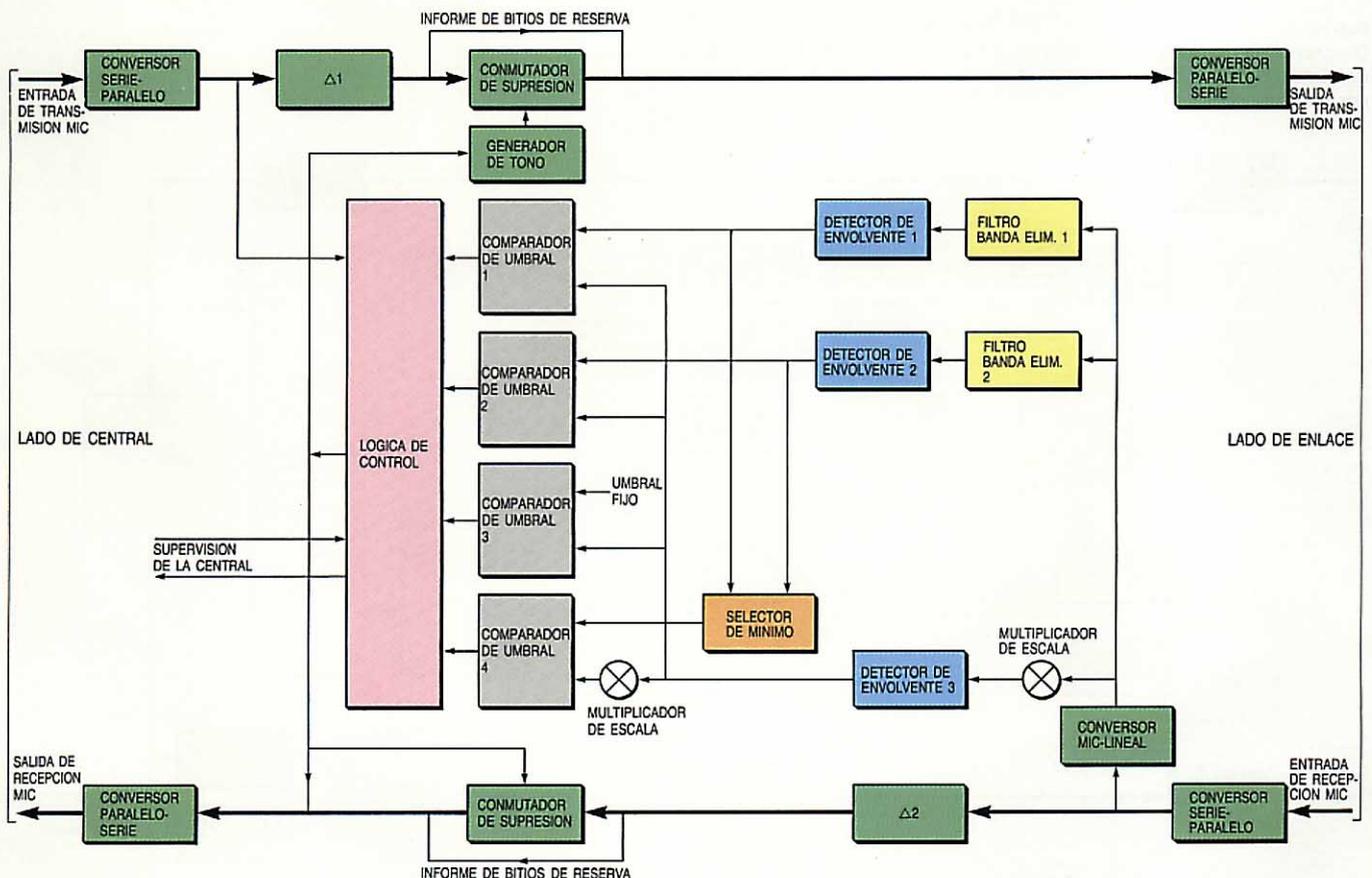
nivel fijo de salida, realizado como una PROM con tabla de consulta de longitud adecuada. El filtrado en la parte de recepción se basa en filtros de banda eliminada, seguidos de detectores de envolvente con umbral adaptativo para detectar los tonos de señalización. Si sólo hay un tono, se hace una comprobación extra para mejorar la inmunidad a la imitación de señal por los forjantes del habla⁴.

Para garantizar una operación rápida, hay un par de secciones RII de 2º orden en cada filtro, seguidas de un detector de envolvente. Las constantes de tiempo de activación y desactivación, y la ganancia del umbral adaptativo, se han optimizado de

acuerdo con la máxima distorsión en tiempo de las señales que permite el CCITT. El equipo de 30 canales ID5 podrá insertarse en los módulos de enlaces analógicos o digitales del Sistema 12 cuando la señalización sea CCITT nº 5 (Fig. 2).

Debido a sus similares exigencias de proceso y señalización, el ID5 y el SDE utilizan mucho soporte físico común, obteniéndose las funciones mediante diferentes microprogramaciones. Esta comunidad de equipo se refleja en el mantenimiento, incluyendo la continua observación en línea y las facilidades de bucle en dúplex, de acuerdo con las recomendaciones del CCITT.

Figura 3
Diagrama de bloques del interfaz digital para señalización CCITT nº 5.



Módulo digital de conferencia

El MDC realiza todas las funciones necesarias para establecer conexiones múltiples con un máximo de 30 abonados, ubicándose en una central local o de tránsito como recurso en reserva conectado a la red digital de conmutación Sistema 12. Los abonados analógicos accederán al MDC utilizando medios normales de transmisión y conmutación, que generalmente consisten en portadores a 2 y 4 hilos y bobinas híbridas. Las funciones principales del MDC son la mezcla de señales de abonado y la protección contra la inestabilidad causada por las reflexiones (ecos) en las conexiones locales (nacionales) de abonado².

El MDC (Fig. 4) consta de dos grandes partes: el mezclador de señales de abonado, y los mecanismos de lógica de entrada-salida y de conmutación de pérdidas por voz para proteger contra la inestabilidad. Se utiliza un algoritmo de mezcla no recursivo al que no afecta ninguna inestabilidad intrínseca o ciclo límite.

El circuito de conmutación por voz se asemeja a un semisupresor de eco. Vigila cada conexión de abonado para detectar cualquier señal de voz (en una o ambas direcciones), de modo casi idéntico al del supresor de eco. Según se esté hablando o no, la línea de abonado se abre en el caso de una sola dirección, o se atenúa en un factor fijo cuando operan ambas direcciones.

Para asegurar la estabilidad con este método de control de eco, hubo que modifi-

car algo la Recomendación G.164 del CCITT a fin de optimizar las prestaciones para conexiones múltiples. El MDC usa mucho soporte físico común al SDE, aparte del mezclador. Mediante programación se le puede configurar para varias conferencias simultáneas independientes entre grupos de abonados de distinto tamaño. Hasta 30 abonados pueden conectarse al mismo módulo físico. El MDC incluye un sistema de vigilancia continua en línea similar al del SDE.

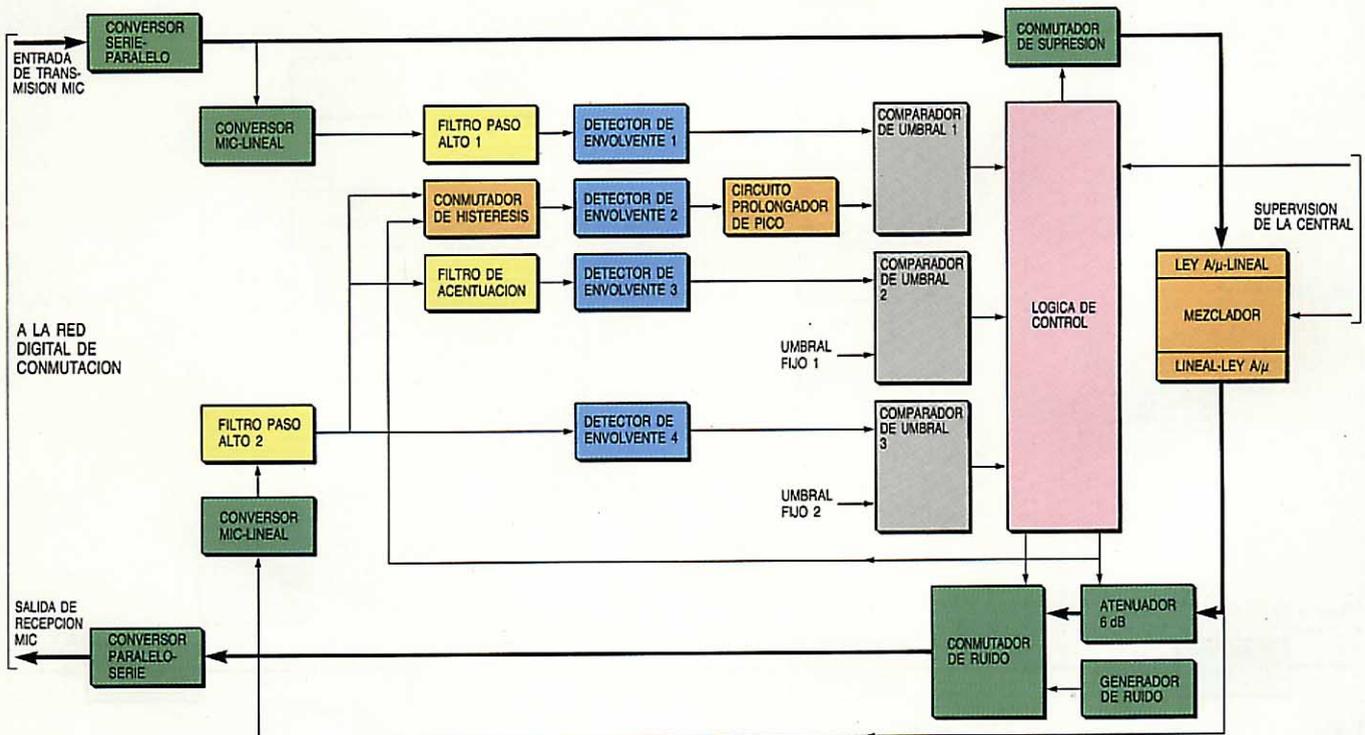
Módulo de prueba de enlaces

El módulo de prueba de enlaces (Fig. 5) permite probar la transmisión y señalización en enlaces analógicos y digitales. Entre los requisitos básicos se incluye la capacidad de realizar hasta 30 pruebas simultáneas de enlace por módulo⁴. El concepto básico de diseño fue repartir la funcionalidad global (ATME n° 2, MCA/DRA* italianas, tasa de errores de bitio, tasa de error en línea, etc.) en varias funciones básicas, cada una hecha realidad en circuitos optimizados.

Tras un detallado análisis de los procedimientos de prueba de enlace actuales y previstos en el mercado europeo, se prepararon las siguientes facilidades:

- generador digital de tono
- instrumentos digitales de medida

Figura 4
Diagrama de bloques del módulo digital de conferencia.



* Siglas en italiano del generador automático de llamada y del contestador automático, respectivamente.

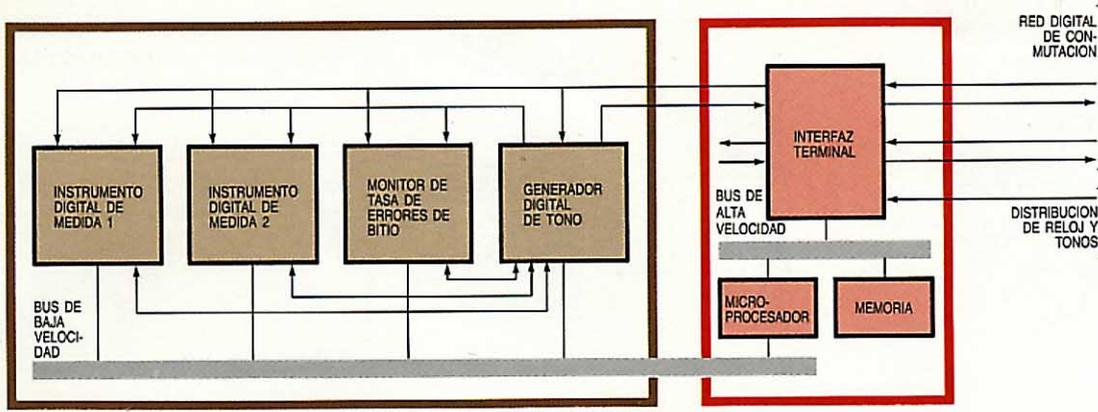


Figura 5
Diagrama de bloques
del módulo de prueba
de enlaces.

- monitor de tasa de errores de bito
- contador digital físico
- interfaz con mesa de pruebas.

Cada función utiliza un soporte físico especializado, capaz de realizar sus tareas con varios canales mediante multiplexación en el tiempo: tal soporte consiste en una placa enchufable, con lo que el módulo podrá configurarse con las mínimas placas para cada aplicación concreta, atendiendo a la funcionalidad y modularidad (número de canales). Esta solución de diseño resulta económica cuando hay más de cuatro o cinco canales en el mismo dispositivo físico (en el caso actual, hay 30 canales).

Generador digital de tono

El generador de tono totalmente digital puede producir una mezcla de hasta tres tonos sinusoidales diferentes, con niveles y frecuencias controlados por separado. El margen de frecuencia real es de 4 a 4000 Hz, en pasos de 8 Hz y con gran estabilidad; el de niveles es de +3 a -60 dBm0, en pasos de 0,5 dB. La frecuencia y el nivel se pueden seleccionar dinámicamente por programa. Para lograr tal flexibilidad se ha elegido un algoritmo especial que se apoya en una sola tabla de consulta, la cual contiene la senoide básica de 4 Hz al nivel máximo. Los tonos reales se obtienen accediendo a la referida tabla por medio de un algoritmo, el cual permite generar todos los tonos utilizados en los sistemas normales de señalización multifrecuencia con desviación de frecuencia nula (R2, n° 5, ATME n° 2, etc.). Además, atendiendo a la cuantificación definida en las especificaciones MIC (ley A ó μ), la pureza espectral y el ancho de banda de ruido son muy elevados. El soporte físico consiste en una placa con 30 canales.

Instrumento digital de medida

Este receptor, microprogramable y totalmente digital, puede realizar 16 funciones distintas e independientes en 16 canales. Las funciones se determinan mediante

programación permanente que especifica las características del filtro, el método de estimación de potencia y la lógica de decisión que correla y analiza los resultados. Se emplea con abundancia el PDS tanto en la placa del procesador del filtro como en la placa de estimación de potencia.

La primera de dichas placas es un procesador microprogramable, y en ella pueden realizarse filtros RII hasta de orden 12° utilizando cualquier célula estándar o filtros múltiples RIF (respuesta impulsiva finita). Estos eficientes filtros en tiempo real se constituyen con un máximo de 32 bancos de filtros diezmadores RIF, cada uno compuesto de hasta 32 filtros de un máximo de 128 etapas por filtro, dependiendo de la magnitud de diezmado permitida. La estabilidad del filtrado RII, así como la gran flexibilidad de los coeficientes permitidos, quedan aseguradas por el soporte físico, que incluye un multiplicador paralelo 16×16 bits, una unidad aritmética-lógica y una memoria de muestras de señal (ambas de 24 bits), más la aritmética de saturación*.

La placa de estimación de potencia es un procesador de señal microprogramable, que realiza dicha estimación mediante suma de cuadrados y promedio móvil. Incluye un multiplicador paralelo de 16×16 bits, una unidad aritmética-lógica y memoria de muestras de 32 bits, y un conversor rápido lineal-logarítmico con margen dinámico de 72 dB y tiempo máximo de conversión de 2 μ s.

Combinando adecuados microprogramas en las placas de procesador de filtrado y estimación de potencia se puede cubrir una amplia gama de aplicaciones, como:

- detección de tonos multifrecuencia para R2, CCITT n° 5, ATME, CCITT n° 2, etc.
- medida selectiva de nivel de alta precisión

* Cada número fuera del margen $-(2^{u-1}-1)$ a $+(2^{u-1}-1)$ (siendo u el número de bits) es sustituido por el valor $-(2^{u-1}-1)$ ó $+(2^{u-1}-1)$, respectivamente.

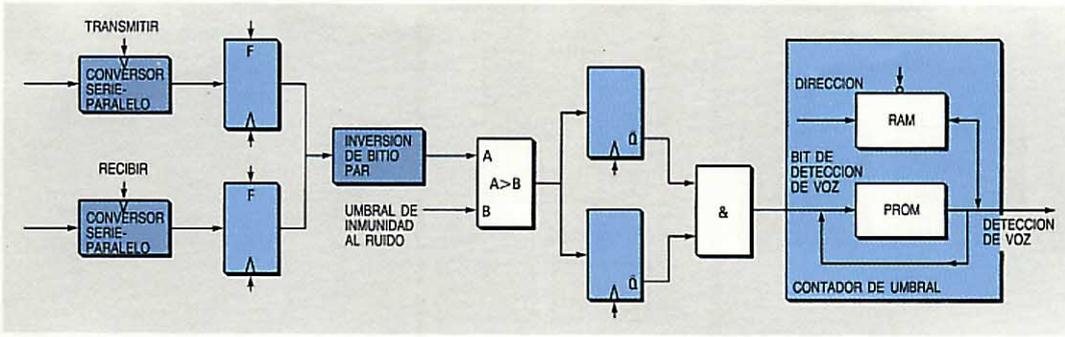


Figura 6
Detector digital de voz.

- medida de ruido con ponderación sofo-métrica y con rechazo de tono por interpolación de voz en el tiempo asignado (TASI)
- discriminación de frecuencia.

Detector digital de voz

El detector de habla totalmente digital (Fig. 6) se equipa en los módulos de enlaces analógicos manuales. Permite al operador informarse del estado de la conversación, mejorando así el control de las conexiones manuales. Las dos entradas serie MIC se derivan de los enlaces de transmisión-recepción y se aplican a un comparador de umbral para rechazar el ruido y detectar, fiable e instantáneamente, la presencia de voz en cualquiera de esos trayectos.

Un contador de umbral integra las muestras instantáneas de habla a fin de evitar una detección errónea de la voz. Los parámetros del contador de umbral y el tiempo de persistencia se han determinado mediante un programa de PDS que simula su comportamiento en presencia de señal oral sintetizada.

Conclusiones

La central interurbana internacional Sistema 12 de Acilia ha ofrecido la oportunidad de desarrollar equipo digital auxiliar para el soporte de conexiones internacionales. Este equipo nuevo, con sus correspondientes programas y herramientas de apoyo, permite a la central Sistema 12 cumplir los requisitos y procedimientos de las distintas administraciones, sólo con un reducido

número de bloques de equipo diferentes y con los programas de aplicación adecuados.

La reciente evolución tecnológica favorece la comunidad de soporte físico entre dichos bloques, y la mejora de sus prestaciones. La primera aplicación consistirá en el desarrollo de equipo de prueba de red, que ampliará las posibilidades del módulo de red Sistema 12.

Referencias

- 1 M. Della Bruna y F. Minuti: Central digital Sistema 12: Central interurbana internacional de Acilia: *Comunicaciones Eléctricas*, 1985, volumen 59, nº 1/2, págs. 145-153.
- 2 P. Haerle y M. della Bruna: Central digital ITT 1240: Aplicación a la red interurbana: *Comunicaciones Eléctricas*, 1981, volumen 56, nº 2/3, págs. 235-247.
- 3 Libro Amarillo del CCITT, volumen III, Fa III-1, Recomendación G.164.
- 4 Libro Amarillo del CCITT, volumen VI, Fa VI-2, Recomendaciones G.140, G.141, G.142, G.143, G.144 y G.145.

Michele Fucito nació en Meta, Nápoles, en 1950. Se graduó Dr. Ing. en ingeniería electrónica por la Universidad de Nápoles en 1976, entrando luego en los laboratorios de I+D de Industrie Pirelli SpA en Milán, donde trabajó en desarrollo de cables de telecomunicación. Desde 1979, el Dr. Fucito trabaja en FACE Sud Selettronica, Salerno, en desarrollo de equipos y proceso digital de señal, siendo hoy responsable del diseño del soporte físico. Es miembro de la Asociación Electrotécnica Italiana.

Giuseppe Mocerino nació en Brusciano, Nápoles, en 1956. En 1982 se graduó Dr. Ing. en ingeniería electrónica (telecomunicación y proceso digital de señal) en la Universidad Técnica de Nápoles, entrando en el mismo año en el centro de diseño de FACE Sud Selettronica, donde ahora se dedica al desarrollo de equipos y proceso digital de señales.

Dunn, J. G.

Proceso de señal: tecnología y perspectivas

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 252–259

El proceso de señales es una tecnología importante en un amplio abanico de aplicaciones que incluyen las telecomunicaciones, teoría de la información, radar, sonar y proceso de voz y de imagen. El autor describe ante todo el papel del proceso de señal en el campo de las comunicaciones, como base para los demás artículos de este número de *Comunicaciones Eléctricas* sobre dicha tecnología. Examina algunas de las técnicas más importantes, como la aproximación de mínimos cuadrados y la estimación y decisión estadística, contemplando después la influencia de la tecnología digital. El artículo concluye con una mirada al futuro del proceso de señal por extrapolación de las tendencias actuales.

Carmody, J.; Rothweiler, J.

Codificación del habla a 800 y 400 bit s⁻¹

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 260–265

Los autores describen una de las principales técnicas para conseguir transmitir el habla de modo inteligible a velocidades menores de 2400 bit s⁻¹. Primero revisan la popular técnica de codificación por predicción lineal, y luego consideran otras tres técnicas: cuantificación vectorial, codificación directa/diferencial y codificación entrelazada.

Christensen, R.; Johnsen, Ø.; Patovan, B.

Proceso de la voz en centrales telefónicas públicas

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 266–272

El acceso a muchos servicios de telecomunicación es a veces complejo, teniendo los usuarios que memorizar numerosos códigos de acceso a distintos servicios. Además, tienen que marcar números largos, difíciles de recordar. Las técnicas modernas de codificación del habla están abriendo paso a sistemas de registro y reconocimiento del habla que facilitarán a los abonados el acceso a los servicios existentes y futuros de comunicación por el camino más natural: la palabra hablada. Los autores analizan algunos métodos de codificación de la voz a velocidades comprendidas entre 100 bit s⁻¹ y 64 kbit s⁻¹ y consideran las velocidades binarias más apropiadas para diversas aplicaciones, como el reconocimiento del habla, teleconferencia y los dispositivos de grabación de voz.

Mulla, H.; Vaughan, J. F.

Aplicación del reconocimiento y la síntesis de voz a los servicios de las PABX

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 273–280

Los diseñadores de PABX se enfrentan a un reto cada vez más apremiante: proporcionar nuevos servicios avanzados manteniendo la sencillez de la operación. El reconocimiento y la síntesis de voz añadirá una nueva dimensión al hacer los futuros sistemas PABX agradables al usuario. Para explotar plenamente la tecnología, el usuario debe poder acceder por diálogo oral a todo el conjunto de facilidades de la PABX. Los autores examinan las principales aplicaciones del reconocimiento y síntesis de voz en este área y exponen la experiencia adquirida con el sistema experimental de ITT. Esto ha demostrado que el concepto de integrar el proceso de voz en una PABX es viable y será aplicado a futuros productos, haciendo conocer asimismo a los diseñadores los problemas que tienen que resolver.

Immendorfer, M.

Marcador telefónico activado por la voz

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 281–285

El autor describe el desarrollo de un marcador telefónico oral utilizable con un teléfono de abonado. El usuario sólo tiene que pronunciar frente al teléfono el nombre del abonado a quien quiere llamar, y el marcador oral reconoce ese nombre, marcando el número correspondiente de una forma enteramente automática. Cada usuario puede definir su vocabulario particular, compuesto de nombres de personas, compañías y ciudades específicas. El prototipo lleva incorporado un procesador de señales en una pastilla para el reconocimiento de voz, y un microprocesador de control que incluye una agenda telefónica electrónica.

Bostelmann, G.; Pirsch, P.

Codificación de señales de vídeo

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 286–294

Las señales de vídeo no son sino uno de los muchos tipos de señales que pueden transmitirse por una futura RDSI de banda ancha. No obstante, para que dicha transmisión sea posible, habrá que codificar previamente tales señales y reducir su velocidad hasta 140 Mbit s⁻¹ o menos. Los autores exponen las distintas técnicas existentes para codificar señales de vídeo, incluyendo la modulación por impulsos codificados (MIC), la MIC diferencial y la codificación por transformada. Utilizando la técnica más ventajosa, es posible reducir la velocidad binaria hasta 69 Mbit s⁻¹ con equipos relativamente sencillos. Se pueden lograr velocidades binarias de 34 Mbit s⁻¹ o incluso menores, pero los codecs necesarios son considerablemente más complejos y peores en rendimiento económico.

Ali, A.

Filtros integrados para señales de vídeo

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 295–299

Aunque las cámaras de vídeo y los aparatos de televisión sean analógicos, las señales pueden procesarse y transmitirse con más eficacia utilizando técnicas digitales. Se necesitan, pues, filtros digitales de buen rendimiento económico que puedan mejorar la calidad de la imagen. El autor describe una estrategia de diseño utilizable para el desarrollo de eficientes filtros con estructura repetitiva, lo cual les hace particularmente adecuados para realizarse en VLSI, y en especial en tecnología CMOS.

Hargrave, P. J.; Ward, C. R.

Nuevo enfoque de la conformación de haz adaptativa

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 300–305

La rápida adaptación será requisito esencial de muchas aplicaciones militares futuras, a fin de obtener una mejor dinámica en plataformas de antenas, contrarrestar artificiosas interferencias y variar con agilidad la frecuencia de las ondas. Los autores describen una nueva técnica para conformación de haz adaptativa en antena que satisface tal finalidad. Esta técnica evita el cálculo explícito de la matriz de covarianza estimada, reduciendo así la sensibilidad a la falta de precisión aritmética. El algoritmo propuesto se puede realizar en una eficiente arquitectura solapada por medio de un ordenamiento sistólico triangular.

Cotton, J. M.; Masterson, G. E.

Cálculo de la transformada discreta de Fourier en el procesador celular

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 306–311

El procesador celular es un ordenador del tipo de una instrucción para datos múltiples, basado en un circuito CMOS de 16 × 1 bitio. La arquitectura se presta a aplicaciones que van desde gestión de bases de datos e inteligencia artificial a gráficos por ordenador y proceso digital de señales. Este artículo analiza la aplicación del procesador celular a la ejecución de un algoritmo para obtener la transformada discreta de Fourier. Los autores obtienen los tiempos de ejecución para transformadas de Fourier de varios tamaños.

Kneib, K. N.

Tendencias en procesadores digitales de señales programables de circuito integrado único

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 312–319

La autora revisa el desarrollo de arquitecturas para procesadores digitales de señales (PDS) programables en un solo circuito integrado. En la actualidad, los diseñadores pueden escoger entre una amplia gama de PDS de este tipo. Los factores que deben considerarse al elegir un dispositivo incluyen la capacidad aritmética, la de memoria dentro y fuera del circuito, el conjunto de instrucciones, la capacidad de entrada/salida, el soporte físico externo, y la disponibilidad y tipo de herramientas de desarrollo. Se hace resaltar la relevancia de estas características para las funciones del proceso de señal, mientras que otras facilidades deseables, no conseguidas hasta ahora, se podrían realizar con dispositivos de nuevas generaciones que aprovechen los avances en diseño y tecnología.

Das, S.; Jayanthi, S.; Middleton, F.; Morton, S.

Procesador digital de señales para aplicaciones de telecomunicación

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 320—326

El analizador de señales de prueba es un procesador digital de señales programable y orientado a la búsqueda de averías, que ha sido desarrollado para la central digital Sistema 12 de ITT. Los autores describen esta versátil herramienta de medida y generación de señal, utilizada en numerosas aplicaciones dentro de la central. El analizador realiza el procesamiento muestra por muestra, a elevada velocidad. El extenso conjunto de instrucciones lleva a cabo funciones de proceso de señal para muchos tipos de medidas en subsistemas tales como los módulos de abonados, enlaces y circuitos de servicio del Sistema 12.

Bång, G.; Carlsson, P.

Proceso digital de señales en módems de transmisión de datos

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 327—332

Los módems de datos, tanto en el lado transmisor como en el receptor de una conexión, actúan como interfaz entre las señales analógicas y digitales. En el pasado, los módems estaban contruidos principalmente con circuitos analógicos, pero a medida que los microprocesadores se vuelven más potentes y de mejor rendimiento económico, éstos van asumiendo muchas de las funciones de los módems. Incluso han podido introducirse nuevas características que han elevado sus prestaciones, entre ellas el uso de igualadores digitales adaptativos para compensar la diferente degradación de las líneas. Otra ventaja de los módems basados en microprocesadores es la facilidad con que pueden variarse sus parámetros. Los autores describen un módem nuevo que aprovecha al máximo las ventajas del proceso digital de señales.

Alvestad, T.; Eriksen, T. J.-C.

Compensador de eco para módems de datos a dos hilos

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 333—337

Utilizando técnicas de compensación de eco pueden transmitirse datos en dúplex por líneas a dos hilos. Se ha demostrado que la compensación de eco da mejores resultados en distancias largas que el método de separación en el tiempo. Se aplica el proceso digital de señal en la

realización de un compensador de eco para un módem de datos a dos hilos que opera a 12 kbit s⁻¹ y con una atenuación máxima de línea de 50 dB, permitiendo un alcance de 15 km en cable de 0,6 mm de diámetro. El compensador de eco consta de dos partes principales: un filtro transversal que opera en tiempo real, y un microprocesador para ajustar los coeficientes del filtro según el algoritmo de correlación del signo. El tiempo de ajuste del filtro es de 2 a 3 segundos contados a partir del encendido.

Gilsanz, M.; Pedrón, F.; Siles, J. A.

Compensación de eco adaptativa para transmisión de datos en banda base

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 338—344

El desarrollo de algoritmos adaptativos potentes y refinados, unido a los avanzados circuitos electrónicos, posibilita la aplicación de técnicas de proceso digital de señales al campo de las comunicaciones. Entre las primeras aplicaciones de tales técnicas figuran los compensadores de eco adaptativos, que están reemplazando a los supresores de eco convencionales y permiten transmisión bidireccional dúplex total a velocidades de datos más elevadas. Los autores describen los algoritmos más importantes y los principales criterios de diseño para compensadores de eco.

Fucito, M.; Mocerino, G.

Aplicación del proceso digital de señal a las pruebas e interfaz de enlace del Sistema 12

Comunicaciones Eléctricas (1985), volumen 59, n° 3, págs. 345—350

Se afirma la tendencia hacia la digitalización de las telecomunicaciones en todas las áreas de la red. Sin embargo, la enorme inversión en la planta analógica obliga a que ésta siga en servicio durante muchos años. Para la central Sistema 12 de Acilia se ha desarrollado una gama de equipos totalmente digitales, con procesadores controlados por microprogramación, que sustituirán a los equivalentes equipos analógicos. Los nuevos diseños que incluyen un supresor de eco, interfaz de señalización de línea CCITT n° 5, módulo de conferencia, módulo de prueba de enlaces y detector de voz, servirán para potenciar el equipo de interfaz con enlaces analógicos y digitales.

Oficinas Editoriales

La correspondencia relacionada con las diferentes versiones de Electrical Communication debe dirigirse al editor correspondiente:

Michael Deason
Electrical Communication
Great Eastern House
Edinburgh Way
Harlow, Essex
England

Otto Grewe
Elektrisches Nachrichtenwesen
Lorenzstrasse 10
7000 Stuttgart 40
Deutsche Bundesrepublik

Antonio Soto
Comunicaciones Eléctricas
Ramírez de Prado, 5
28045 Madrid
España

Jean-Pierre Dartois
Revue des Télécommunications
30 Avenue George V
75008 Paris
France