

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 380 962**

51 Int. Cl.:
G10L 19/08 (2006.01)
G10L 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08001922 .7**
96 Fecha de presentación: **06.10.2001**
97 Número de publicación de la solicitud: **1912207**
97 Fecha de publicación de la solicitud: **16.04.2008**

54 Título: **Procedimiento y aparato para codificación de baja tasa de transmisión de bits de habla sorda de alto rendimiento**

30 Prioridad:
17.10.2000 US 690915

45 Fecha de publicación de la mención BOPI:
21.05.2012

45 Fecha de la publicación del folleto de la patente:
21.05.2012

73 Titular/es:
**QUALCOMM INCORPORATED
5775 MOREHOUSE DRIVE
SAN DIEGO CA 92121-1714, US**

72 Inventor/es:
Huang, Pengjun

74 Agente/Representante:
Carpintero López, Mario

ES 2 380 962 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Procedimiento y aparato para codificación de baja tasa de transmisión de bits de habla sorda de alto rendimiento

Antecedentes**I. Campo de la invención**

- 5 Las realizaciones dadas a conocer se refieren al campo del procesamiento del habla. Más en particular, las realizaciones dadas a conocer se refieren a un procedimiento y un aparato novedosos y mejorados para codificación de baja tasa de transmisión de bits de segmentos sordos del habla.

II. Antecedentes

- 10 La transmisión de voz mediante técnicas digitales se ha generalizado, en particular en aplicaciones telefónicas de radio digital y larga distancia. Esto, a su vez, ha creado interés en determinar la mínima cantidad de información que puede enviarse a través de un canal mientras se mantiene la calidad percibida del habla reconstruida. Si el habla se transmite simplemente muestreando y digitalizando, se requiere una tasa de transmisión de datos del orden de sesenta y cuatro kilobits por segundo (kbps) para conseguir la calidad del habla del teléfono analógico convencional. Sin embargo, mediante el uso del análisis del habla, seguido por la codificación, transmisión y resíntesis apropiadas en el receptor, puede conseguirse una reducción significativa en la tasa de transmisión de datos.

- 15 Los dispositivos que emplean técnicas para comprimir el habla extrayendo parámetros que se refieren a un modelo de generación del habla humana se denominan codificadores del habla. Un codificador del habla divide la señal de habla entrante en bloques de tiempo, o tramas de análisis. Los codificadores del habla comprenden normalmente un codificador y un descodificador, o un códec. El codificador analiza la trama de habla entrante para extraer ciertos parámetros pertinentes, y entonces cuantifica los parámetros en una representación binaria, es decir, en un conjunto de bits o un paquete de datos binarios. Los paquetes de datos se transmiten a través del canal de comunicación a un receptor y un descodificador. El descodificador procesa los paquetes de datos, los descuantifica para producir los parámetros, y entonces resintetiza las tramas de habla utilizando los parámetros descuantificados.

- 20 La función del codificador del habla es comprimir la señal de habla digitalizada en una señal de baja tasa de transmisión de bits eliminando todas las redundancias naturales intrínsecas del habla. La compresión digital se consigue representando la trama de habla de entrada con un conjunto de parámetros y empleando cuantificación para representar los parámetros con un conjunto de bits. Si la trama de habla de entrada tiene un número de bits N_1 y el paquete de datos producido por el codificador del habla tiene un número de bits N_0 , el factor de compresión conseguido por el codificador del habla es $C_r = N_1/N_0$. El reto es mantener una alta calidad de voz del habla descodificada mientras se consigue el factor de compresión objetivo. El rendimiento de un codificador del habla depende de (1) cómo de bien se realice el modelo del habla, o la combinación del proceso de análisis y síntesis descrito anteriormente, y (2) cómo de bien se realice el proceso de cuantificación de parámetros a la tasa de transmisión de bits objetivo de N_0 bits por trama. Por tanto, el objetivo del modelo del habla es capturar la esencia de la señal de habla, o la calidad de voz objetivo, con un pequeño conjunto de parámetros para cada trama.

- 25 Los codificadores del habla pueden implementarse como codificadores en el dominio del tiempo, que intentan capturar la forma de onda del habla en el dominio del tiempo empleando procesamiento de alta resolución en el tiempo para codificar pequeños segmentos del habla (normalmente subtramas de 5 milisegundos (ms)) cada vez. Para cada subtrama, se encuentra una alta precisión representativa de un espacio de libro de código por medio de diversos algoritmos de búsqueda conocidos en la técnica. Como alternativa, los codificadores del habla pueden implementarse como codificadores en el dominio de la frecuencia, que intentan capturar el espectro del habla a corto plazo de la trama de habla de entrada con un conjunto de parámetros (análisis) y emplean un proceso de síntesis correspondiente para recrear la forma de onda del habla a partir de los parámetros espectrales. El cuantificador de parámetros conserva los parámetros representándolos con representaciones almacenadas de vectores de código según técnicas de cuantificación conocidas descritas en A. Gersho & R.M. Gray, Vector Quantization and Signal Compression (1992).

- 30 Un codificador del habla en el dominio del tiempo ampliamente conocido es el codificador Predictivo Lineal Excitado por Código (CELP) descrito en L.B. Rabiner & R.W. Schafer, Digital Processing of Speech Signals 396-453 (1978), que está incorporado en su totalidad en el presente documento por referencia. En un codificador CELP, las correlaciones a corto plazo, o redundancias, en la señal de habla se eliminan mediante un análisis de predicción lineal (LP), que encuentra los coeficientes de un filtro formante a corto plazo. Aplicar el filtro de predicción a corto plazo a la trama de habla entrante genera una señal de residuo LP, que se modela y cuantifica adicionalmente con parámetros de filtro de predicción a largo plazo y un libro de código estocástico posterior. Por tanto, la codificación CELP divide la tarea de codificar la forma de onda del habla en el dominio del tiempo en las tareas separadas de codificación de los coeficientes de filtro a corto plazo LP y codificar el residuo LP. La codificación en el dominio del tiempo puede realizarse a una tasa fija (es decir, utilizando el mismo número de bits, N_0 , para cada trama) o a una tasa variable (en la que se utilizan diferentes tasas de transmisión de bits para diferentes tipos de contenidos de trama). Los codificadores de tasa variable intentan utilizar sólo la cantidad de bits necesaria para codificar los

parámetros de códec a un nivel adecuado para obtener una calidad objetivo. Un codificador CELP de tasa variable ejemplar se describe en la patente estadounidense n.º 5.414.796, que está transferida al cesionario de las realizaciones dadas a conocer actualmente e incorporada en su totalidad en el presente documento por referencia.

5 Los codificadores en el dominio del tiempo tales como el codificador CELP se basan normalmente en un alto número de bits, N_0 , por trama para conservar la precisión de la forma de onda del habla en el dominio del tiempo. Normalmente tales codificadores proporcionan una calidad de voz excelente dado el número de bits, N_0 , por trama relativamente grande (por ejemplo, 8 kbps o superior). Sin embargo, a bajas tasas de transmisión de bits (4 kbps e inferiores), los codificadores en el dominio del tiempo no conservan la alta calidad y el rendimiento robusto debido al número limitado de bits disponibles. A bajas tasas de transmisión de bits, el espacio de libro de código limitado
10 recorta la capacidad de ajuste de forma de onda de codificadores en el dominio del tiempo convencionales, que se utilizan tan satisfactoriamente en aplicaciones comerciales de tasa superior.

Normalmente, los esquemas CELP emplean un filtro de predicción a corto plazo (STP) y un filtro de predicción a largo plazo (LTP). Se emplea un enfoque de Análisis por Síntesis (AbS) en un codificador para encontrar los retardos y ganancias LTP, así como los mejores índices y ganancias de libro de código estocástico. Los
15 codificadores CELP del estado de la técnica actual tales como el Codificador de Tasa Variable Mejorada (EVRC) pueden conseguir habla sintetizada de buena calidad a una tasa de transmisión de datos de aproximadamente 8 kilobits por segundo.

También se conoce que el habla sorda no muestra periodicidad. El ancho de banda consumido que codifica el filtro LTP en los esquemas CELP convencionales no se utiliza tan eficazmente para habla sorda como para habla sonora,
20 en la que la periodicidad del habla es fuerte y el filtrado LTP es significativo. Por lo tanto, es deseable un esquema de codificación más eficaz (es decir tasa de transmisión de bits inferior) para habla sorda.

Para codificar a tasas de transmisión de bits inferiores, se han desarrollado diversos procedimientos de codificación espectral del habla, o en el dominio de la frecuencia, en los que la señal de habla se analiza como una evolución variable en el tiempo de espectros. Véase, por ejemplo, R.J. McAulay & T.F. Quatieri, Sinusoidal, Coding, in Speech
25 Coding and Synthesis ch. 4 (W.B. Kleijn & K.K. Paliwal eds., 1995). En codificadores espectrales, el objetivo es modelar, o predecir, el espectro del habla a corto plazo de cada trama de entrada de habla con un conjunto de parámetros espectrales, en lugar de imitar con precisión la forma de onda del habla variable en el tiempo. Entonces se codifican los parámetros espectrales y se crea una trama de habla de salida con los parámetros descodificados. El habla resultante sintetizada no coincide con la forma de onda del habla de entrada original, aunque ofrece una
30 calidad percibida similar. Ejemplos de codificadores en el dominio de la frecuencia que son bien conocidos en la técnica incluyen codificadores de excitación multibanda (MBE), codificadores de transformación sinusoidal (STC) y codificadores de armónicos (HC). Tales codificadores en el dominio de la frecuencia ofrecen un modelo paramétrico de alta calidad que presenta un conjunto de parámetros compacto que puede cuantificarse con precisión con el bajo número de bits disponibles a bajas tasas de transmisión de bits.

35 Sin embargo, la codificación a baja tasa de transmisión de bits impone la limitación crítica de una resolución de codificación limitada, o un espacio de libro de código limitado, que limita la efectividad de un único mecanismo de codificación, haciendo que el codificador no pueda representar diversos tipos de segmentos del habla bajo diversas condiciones de fondo con igual precisión. Por ejemplo, los codificadores en el dominio de la frecuencia, de baja tasa de transmisión de bits, convencionales no transmiten información de fase para tramas del habla. En su lugar, la
40 información de fase se reconstruye utilizando un valor de fase inicial, aleatorio, generado artificialmente y técnicas de interpolación lineal. Véase, por ejemplo, H. Yang et al., Quadratic Phase Interpolation for Voiced Speech Synthesis in the MBE Model, in 29 Electronic Letters 856-57 (mayo de 1993). Debido a que la información de fase se genera artificialmente, incluso aunque las amplitudes de las senoides se conserven perfectamente mediante el proceso de cuantificación-descuantificación, el habla de salida producida por el codificador en el dominio de la
45 frecuencia no estará alineada con el habla de entrada original (es decir, los pulsos principales no estarán en sincronización). Por lo tanto ha resultado difícil adoptar cualquier medida de rendimiento en bucle cerrado, tal como, por ejemplo, relación señal-ruido (SNR) o SNR perceptiva, en codificadores en el dominio de la frecuencia.

Una técnica efectiva para codificar habla eficazmente a baja tasa de transmisión de bits es la codificación multimodo. Se han empleado técnicas de codificación multimodo para realizar codificación del habla a baja tasa conjuntamente
50 con un proceso de decisión de modo de bucle abierto. Una técnica de codificación multimodo de este tipo se describe en Arnitava Das et al., Multimode and Variable-Rate Coding of Speech, in Speech Coding and Synthesis ch. 7 (W.B. Kleijn & K.K. Paliwal eds., 1995). Los codificadores multimodo convencionales aplican diferentes modos, o algoritmos de codificación-descodificación, a diferentes tipos de tramas de habla de entrada. Cada modo, o proceso de codificación-descodificación, se adapta para representar un cierto tipo de segmento del habla, tal como, por ejemplo, habla sonora, habla sorda, o ruido de fondo (no habla) de la manera más eficaz. Un mecanismo de
55 decisión de modo de bucle abierto examina la trama de habla de entrada y toma una decisión con respecto a qué modo aplicar a la trama. La decisión de modo de bucle abierto normalmente se realiza extrayendo un número de parámetros de la trama de entrada, evaluando los parámetros para ciertas características temporales y espectrales, y basando una decisión de modo en la evaluación. La decisión de modo se realiza por tanto sin conocer de antemano la condición exacta del habla de salida, es decir, cómo de cerca estará el habla de salida del habla de
60 entrada en cuanto a calidad de voz u otras mediciones de rendimiento. Un modo de decisión de bucle abierto

ejemplar para un códec del habla se describe en la patente estadounidense n.º 5.414.796, que fue transferida al cesionario de las realizaciones dadas a conocer actualmente e incorporada en su totalidad en el presente documento por referencia.

5 La codificación multimodo puede ser a tasa fija, utilizando el mismo número de bits N_0 para cada trama, o a tasa variable, en la que se utilizan diferentes tasas de transmisión de bits para diferentes modos. El objetivo en la codificación a tasa variable es utilizar sólo la cantidad de bits necesaria para codificar los parámetros de códec a un nivel adecuado para obtener la calidad objetivo. Como resultado, puede obtenerse la misma calidad de voz objetivo que la de un codificador de tasa superior, a tasa fija, a una tasa promedio inferior significativa utilizando técnicas de tasa de transmisión de bits variable (VBR). Un codificador del habla de tasa variable ejemplar se describe en la
10 patente estadounidense n.º 5.414.796, transferida al cesionario de las realizaciones dadas a conocer actualmente y previamente incorporada en su totalidad en el presente documento por referencia.

Actualmente hay un aumento del interés en la investigación y necesidades comerciales fuertes para desarrollar un codificador del habla de alta calidad que opere a medias a bajas tasas de transmisión de bits (es decir, en el intervalo de 2,4 a 4 kbps y por debajo). Las áreas de aplicación incluyen telefonía inalámbrica, comunicaciones por
15 satélite, telefonía por Internet, diversas aplicaciones multimedia y de flujo continuo (*streaming*) de voz, correo de voz, y otros sistemas de almacenamiento de voz. Las fuerzas conductoras son la necesidad de alta capacidad y la demanda de rendimiento robusto en situaciones de pérdida de paquetes. Los diversos esfuerzos recientes de estandarización de la codificación del habla son otra fuerza conductora directa que impulsa la investigación y el desarrollo de algoritmos de codificación del habla a baja tasa. Un codificador del habla de baja tasa crea más
20 canales, o usuarios, por ancho de banda de aplicación permisible, y un codificador del habla de baja tasa acoplado con una capa adicional de codificación de canal adecuada puede adecuarse al presupuesto de bits global de especificaciones de codificadores y proporcionar un rendimiento robusto en condiciones de errores de canales.

Por lo tanto, la codificación del habla VBR multimodo es un mecanismo efectivo para codificar habla a baja tasa de transmisión de bits. Los esquemas multimodo convencionales requieren el diseño de esquemas, o modos, de
25 codificación eficaces para diversos segmentos del habla (por ejemplo, sordo, sonoro, transición) así como un modo para ruido de fondo, o silencio. El rendimiento global del codificador del habla depende de cómo se comporte cada modo, y la tasa promedio del codificador depende de las tasas de transmisión de bits de los diferentes modos para segmentos del habla sordos, sonoros y otros. Con el fin de conseguir la calidad objetivo a una tasa promedio baja, es necesario diseñar modos de alto rendimiento, eficaces, algunos de los cuales deben trabajar a bajas tasas de
30 transmisión de bits. Normalmente, los segmentos del habla sonoros y sordos se capturan a altas tasas de transmisión de bits, y los segmentos de ruido de fondo y de silencio se representan con modos que trabajan a una tasa significativamente inferior. Por tanto, existe una necesidad de una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento que capture con precisión un alto porcentaje de segmentos del habla sordos mientras que utiliza un número de bits por trama mínimo.

35 Adicionalmente se llama la atención sobre el documento WO 00/30074, que da a conocer una técnica de codificación a baja tasa de transmisión de bits para segmentos del habla sordos, incluyendo las etapas de extraer coeficientes de energía de alta resolución en el tiempo de una trama de habla, cuantificar los coeficientes de energía, generar una envolvente de energía de alta resolución en el tiempo a partir de los coeficientes de energía cuantificados, y reconstruir una señal residual conformando un vector de ruido generado aleatoriamente con valores
40 cuantificados de la envolvente de energía. La envolvente de energía puede generarse con una técnica de interpolación lineal. Puede obtenerse una medición de postprocesamiento y compararse con un umbral predefinido para determinar si el algoritmo de codificación se está comportado adecuadamente.

Sumario

45 Según la presente invención, se proporciona un procedimiento para descodificar segmentos del habla sordos, como se exponen en la reivindicación 1, y un descodificador, como se expone en la reivindicación 9. En las reivindicaciones dependientes se describen realizaciones de la invención.

Las realizaciones dadas a conocer se dirigen a una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento que captura con precisión segmentos del habla sordos mientras que utiliza un número de bits por trama
50 mínimo. Por consiguiente, en un aspecto de la invención, un procedimiento de descodificación de segmentos del habla sordos incluye recuperar un grupo de ganancias cuantificadas utilizando índices recibidos para una pluralidad de subtramas; generar una señal de ruido aleatorio que comprende números aleatorios para cada una de la pluralidad de subtramas; seleccionar un porcentaje predeterminado de los números aleatorios de mayor amplitud de la señal de ruido aleatorio para cada una de la pluralidad de subtramas; ajustar a escala los números aleatorios de mayor amplitud seleccionados mediante las ganancias recuperadas para cada subtrama para producir una señal de
55 ruido aleatorio ajustada a escala; filtrar paso banda y conformar la señal de ruido aleatorio ajustada a escala; y seleccionar un segundo filtro basándose en un indicador de selección de filtro recibido y conformar adicionalmente la señal de ruido aleatorio ajustada a escala con el filtro seleccionado.

Breve descripción de los dibujos

Las características, objetos, y ventajas de las realizaciones dadas a conocer resultarán más evidentes a partir de la descripción detallada expuesta a continuación tomada conjuntamente con los dibujos, en los que caracteres de referencia similares identifican lo mismo en todo el documento y en los que:

- 5 La figura 1 es un diagrama de bloques de un canal de comunicación que termina en cada extremo con codificadores del habla;
- La figura 2A es un diagrama de bloques de un codificador que puede utilizarse en un codificador del habla de baja tasa de transmisión de bits de alto rendimiento;
- 10 La figura 2B es un diagrama de bloques de un descodificador que puede utilizarse en un codificador del habla de baja tasa de transmisión de bits de alto rendimiento;
- La figura 3 ilustra un codificador del habla sorda de baja tasa de transmisión de bits de alto rendimiento que podría utilizarse en el codificador de la figura 2A;
- La figura 4 ilustra un descodificador del habla sorda de baja tasa de transmisión de bits de alto rendimiento que podría utilizarse en el descodificador de la figura 2B;
- 15 La figura 5 es un diagrama de flujo que ilustra etapas de codificación de una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento para habla sorda;
- La figura 6 es un diagrama de flujo que ilustra etapas de descodificación de una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento para habla sorda;
- 20 La figura 7A es un gráfico de una respuesta en frecuencia de un filtrado paso bajo para su uso en análisis de energía de banda;
- La figura 7B es un gráfico de una respuesta en frecuencia de un filtrado paso alto para su uso en análisis de energía de banda;
- La figura 8A es un gráfico de una respuesta en frecuencia de un filtro paso banda para su uso en un filtrado perceptivo;
- 25 La figura 8B es un gráfico de una respuesta en frecuencia de un filtro de conformación preliminar para su uso en un filtrado perceptivo;
- La figura 8C es un gráfico de una respuesta en frecuencia de un filtro de conformación que puede utilizarse en un filtrado perceptivo final; y
- 30 La figura 8D es un gráfico de una respuesta en frecuencia de otro filtro de conformación que puede utilizarse en un filtrado perceptivo final.

Descripción detallada de las realizaciones preferidas

Las realizaciones dadas a conocer proporcionan un procedimiento y un aparato para codificación de habla sorda a baja tasa de transmisión de bits de alto rendimiento. Las señales de habla sorda se digitalizan y convierten en tramas de muestras. Cada trama de habla sorda se filtra mediante un filtro de predicción a corto plazo para producir bloques de señales a corto plazo. Cada trama se divide en múltiples subtramas. Entonces se calcula una ganancia para cada subtrama. Estas ganancias posteriormente se cuantifican y transmiten. Entonces, se genera y se filtra un bloque de ruido aleatorio mediante los procedimientos descritos en detalle posteriormente. Este ruido aleatorio filtrado se ajusta a escala mediante las ganancias de subtrama cuantificadas para formar una señal cuantificada que representa la señal a corto plazo. En un descodificador se genera y se filtra una trama de ruido aleatorio de la misma manera que el ruido aleatorio en el codificador. El ruido aleatorio filtrado en el descodificador entonces se ajusta a escala mediante las ganancias de subtrama recibidas, y se pasa a través de un filtro de predicción a corto plazo para formar una trama de habla sintetizada que representa las muestras originales.

Las realizaciones dadas a conocer presentan una técnica de codificación novedosa para una variedad de habla sorda. A 2 kilobits por segundo, el habla sorda sintetizada es equivalente perceptivamente a la producida por los esquemas CELP convencionales que requieren tasas de transmisión de datos mucho más altas. Un alto porcentaje (aproximadamente el veinte por ciento) de segmentos de habla sorda pueden codificarse según las realizaciones dadas a conocer.

En la figura 1 un primer codificador 10 recibe muestras del habla digitalizadas $s(n)$ y codifica las muestras $s(n)$ para su transmisión sobre un medio 12 de transmisión, o canal 12 de comunicación, a un primer descodificador 14. El descodificador 14 descodifica las muestras del habla codificadas y sintetiza una señal de habla de salida $S_{SYNTH}(n)$. Para la transmisión en sentido opuesto, un segundo codificador 16 codifica muestras del habla digitalizadas $s(n)$,

que se transmiten sobre un canal 18 de comunicación. Un segundo descodificador **20** recibe y descodifica las muestras del habla codificadas, generando una señal de habla de salida sintetizada $S_{SYNTH}(n)$.

Las muestras del habla, $s(n)$, representan señales de habla que se han digitalizado y cuantificado según cualquiera de diversos procedimientos conocidos en la técnica entre los que se incluyen, por ejemplo, modulación de código de pulso (PCM), μ -law o A-law comprimido-expandido. Como se conoce en la técnica, las muestras del habla, $s(n)$, se organizan en tramas de datos de entrada en las que cada trama comprende un número predeterminado de muestras del habla digitalizadas $s(n)$. En una realización ejemplar, se emplea una tasa de muestreo de 8 kHz, con cada trama de 20 ms comprendiendo 160 muestras. En las realizaciones descritas posteriormente, la tasa de transmisión de datos puede variar de trama a trama desde 8 kbps (tasa completa) a 4 kbps (media tasa) a 2 kbps (cuarto de tasa) a 1 kbps (octavo de tasa). Como alternativa, pueden utilizarse otras tasas de transmisión de datos. Como se utiliza en el presente documento, los términos "tasa completa" o "alta tasa" se refieren en general a tasas de transmisión de datos que son mayores o iguales a 8 kbps, y los términos "media tasa" o "baja tasa" se refieren en general a tasas de transmisión de datos que son menores o iguales a 4 kbps. Variar la tasa de transmisión de datos es beneficioso porque pueden emplearse selectivamente tasas de transmisión de bits inferiores para tramas que contienen relativamente menos información de habla. Como entienden los expertos en la técnica, pueden utilizarse otras tasas de muestreo, tamaños de trama, y tasas de transmisión de datos.

El primer codificador 10 y el segundo descodificador 20 comprenden juntos un primer codificador del habla, o códec del habla. De manera similar, el segundo codificador 16 y el primer descodificador **14** comprenden juntos un segundo codificador del habla. Los expertos en la técnica entienden que los codificadores del habla pueden implementarse con un procesador de señal digital (DSP), un circuito integrado de aplicación específica (ASIC), lógica de puertas discreta, *firmware*, o cualquier módulo de software programable convencional y un microprocesador. El módulo de software podría residir en memoria RAM, memoria flash, registros, o cualquier otra forma de medio de almacenamiento de escritura conocido en la técnica. Como alternativa, cualquier procesador, controlador, o máquina de estado convencionales podría sustituirse por el microprocesador. ASIC ejemplares diseñados específicamente para codificación del habla se describen en la patente estadounidense n.º 5.727.123, transferida al cesionario de las realizaciones dadas a conocer actualmente e incorporada en su totalidad en el presente documento por referencia, y la patente estadounidense n.º 5.784.532, titulada APPLICATION SPECIFIC INTEGRATED CIRCUIT (ASIC) FOR PERFORMING RAPID SPEECH COMPRESSION IN A MOBILE TELEPHONE SYSTEM, transferida al cesionario de las realizaciones dadas a conocer actualmente e incorporada en su totalidad en el presente documento por referencia.

La figura 2A es un diagrama de bloques de un codificador, ilustrado en la figura (**10, 16**), que puede emplear las realizaciones dadas a conocer actualmente. Una señal de habla, $s(n)$, se filtra mediante un filtro **200** de predicción a corto plazo. El habla en sí misma, $s(n)$ y/o la señal residual de predicción lineal $r(n)$ a la salida del filtro 200 de predicción a corto plazo proporcionan la entrada a un clasificador **202** del habla.

La salida del clasificador **202** del habla proporciona la entrada a un conmutador 203 permitiendo el conmutador 203 seleccionar un codificador (**204,206**) de modo correspondiente basándose en un modo clasificado del habla. Un experto en la técnica entendería que el clasificador **202** del habla no se limita a clasificación de habla sonora y sorda y que también puede clasificar transición, ruido de fondo (silencio), u otros tipos de habla.

El codificador **204** de habla sonora codifica el habla sonora mediante cualquier procedimiento convencional tal como por ejemplo, CELP o interpolación de forma de onda prototipo (PWI).

El codificador **205** de habla sorda codifica el habla sorda a una baja tasa de transmisión de bits según las realizaciones descritas posteriormente. El codificador **206** de habla sorda se describe en detalle con referencia a la figura 3 según una realización.

Después de la codificación mediante o bien el codificador **204** o bien el codificador **206**, un multiplexor 208 forma un flujo de bits por paquetes que comprende paquetes de datos, modo del habla, y otros parámetros codificados para la transmisión.

La figura 2B es un diagrama de bloques de un descodificador, ilustrado in la figura 1 (**14, 20**), que puede emplear las realizaciones dadas a conocer actualmente.

El demultiplexor **210** recibe un flujo de bits por paquetes, demultiplexa datos del flujo de bits, y recupera paquetes de datos, el modo del habla, y otros parámetros codificados.

La salida del demultiplexor **210** proporciona la entrada a un conmutador **211** permitiendo el conmutador **211** seleccionar un descodificador (212, 214) de modo correspondiente basándose en un modo clasificado del habla. Un experto en la técnica entendería que el conmutador **211** no se limita a modos de habla sonora y sorda y también puede reconocer transición, ruido de fondo (silencio) u otros tipos de habla.

El descodificador **212** de habla sonora descodifica el habla sonora realizando las operaciones inversas del codificador **204** sonoro.

En una realización, el descodificador **214** de habla sorda descodifica el habla sorda transmitida a una baja tasa de transmisión de bits como se describe posteriormente en detalle con referencia a la figura 4.

Después de la descodificación mediante o bien el descodificador **212** o bien el descodificador **214**, se filtra una señal residual de predicción lineal sintetizada mediante un filtro **216** de predicción a corto plazo. El habla sintetizada en la salida del filtro **216** de predicción a corto plazo se pasa a un procesador **218** postfiltro para generar el habla de salida final.

La figura 3 es un diagrama de bloques detallado del codificador **206** del habla sorda de baja tasa de transmisión de bits de alto rendimiento ilustrado en la figura 2A. La figura 3 detalla el aparato y la secuencia de operaciones de una realización del codificador sordo.

Las muestras del habla digitalizadas, $s(n)$, se introducen al analizador **302** de codificación predictiva lineal (LPC) y al filtro **304** LPC. El analizador **302** LPC produce coeficientes predicativos lineales (LP) de las muestras del habla digitalizadas. El filtro **304** LPC produce una señal residual de habla, $r(n)$, que se introduce al componente 306 de cálculo de ganancia y al analizador **314** de energía de banda no ajustada a escala.

El componente **306** de cálculo de ganancia divide cada trama de muestras del habla digitalizadas en subtramas, calcula un conjunto de ganancias de libro de código, a las que se hace referencia en lo sucesivo en el presente documento como ganancias o índices, para cada subtrama, divide las ganancias en subgrupos, y normaliza las ganancias de cada subgrupo. La señal residual de habla $r(n)$, $n=0, \dots, N-1$, se segmenta en K subtramas, donde N es el número de muestras residuales en una trama. En una realización, $K=10$ y $N=160$. Se calcula una ganancia, $G(i)$, $i=0, \dots, K-1$, para cada subtrama como sigue:

20

$$G(i) = \sum_{k=0}^{N/K-1} r(i * N / K + k)^2, \quad i=0, \dots, K-1,$$

y

$$G(i) = \sqrt{\frac{G(i)}{N/K}}.$$

El cuantificador **308** de ganancia cuantifica las K ganancias, y posteriormente se transmite el índice de libro de código de ganancia para las ganancias. La cuantificación puede realizarse utilizando esquemas de cuantificación de vectores o lineal convencionales, o cualquier variante. Un esquema realizado es la cuantificación de vectores de múltiples etapas.

La salida de señal residual del filtro **304** LPC, $r(n)$, se pasa a través de un filtro paso bajo y un filtro paso alto en el analizador **314** de energía de banda no ajustada a escala. Los valores de energía de $r(n)$, E_1 , E_{lp1} , y E_{hp1} , se calculan para la señal residual, $r(n)$. E_1 es la energía en la señal residual, $r(n)$. E_{lp1} es la energía de banda baja en la señal residual, $r(n)$. E_{hp1} es la energía de banda alta en la señal residual, $r(n)$. Las respuesta en frecuencia de los filtros paso bajo y paso alto del analizador **314** de energía de banda no ajustada a escala, en una realización, se muestran en la figura 7A y en la figura **7B**, respectivamente. Los valores de energía E_1 , E_{lp1} , y E_{hp1} se calculan como sigue:

35

$$E_1 = \sum_{i=0}^{N-1} r^2(n),$$

$$r_{lp}(n) = \sum_{i=1}^{M_{lp}-1} r_{lp}(n-i) * a_{lp}(i) + \sum_{j=0}^{N_{lp}-1} r(n-j) * b_{lp}(j), \quad n=0, \dots, N-1,$$

$$r_{hp}(n) = \sum_{i=1}^{M_{hp}-1} r_{hp}(n-i) * a_{hp}(i) + \sum_{j=0}^{N_{hp}-1} r(n-j) * b_{hp}(j), \quad n=0, \dots, N-1,$$

$$E_{lp1} = \sum_{i=0}^{N-1} r_{lp}^2(i),$$

$$E_{hp1} = \sum_{i=0}^{N-1} r_{hp}^2(i).$$

5 Los valores de energía E_1 , E_{lp1} , y E_{hp1} se utilizan posteriormente para seleccionar filtros de conformación en el filtro **316** de conformación final para procesar una señal de ruido aleatorio de modo que la señal de ruido aleatorio se parezca lo más posible a la señal residual original.

10 El generador **310** de números aleatorios genera números aleatorios uniformemente distribuidos, con varianza unidad, entre -1 y 1 para cada una de las **K** subtramas emitidas por el analizador **302** LPC. El selector **312** de números aleatorios selecciona entre una mayoría de los números aleatorios de baja amplitud en cada subtrama. Se retiene una fracción de los números aleatorios de mayor amplitud para cada subtrama. En una realización, la fracción de números aleatorios retenida es del 25%.

15 La salida de números aleatorios para cada subtrama del selector **312** de números aleatorios se multiplica entonces por las respectivas ganancias cuantificadas de la subtrama, emitidas desde el cuantificador **308** de ganancia, mediante el multiplicador **307**. La salida de señal aleatoria ajustada a escala del multiplicador **307**, $\hat{r}_1(n)$, se procesa entonces mediante filtrado perceptivo.

Para mejorar la calidad perceptiva y mantener la naturalidad del habla sorda cuantificada, se realiza un proceso de filtrado perceptivo de dos etapas sobre la señal aleatoria ajustada a escala, $\hat{r}_1(n)$.

20 En la primera etapa del proceso de filtrado perceptivo, la señal aleatoria ajustada a escala $\hat{r}_1(n)$ se pasa a través de dos filtros fijos en el filtro **318** perceptivo. El primer filtro fijo del filtro **318** perceptivo es un filtro **320** paso banda que elimina las frecuencias del extremo bajo y el extremo alto de $\hat{r}_1(n)$ para producir la señal $\hat{r}_2(n)$. La respuesta en frecuencia del filtro **320** paso banda, en una realización, se ilustra en la figura 8A. El segundo filtro fijo del filtro **318** perceptivo es el filtro **322** de conformación preliminar. La señal, $\hat{r}_2(n)$, calculada por el elemento **320**, se pasa a través del filtro **322** de conformación preliminar para producir la señal $\hat{r}_3(n)$. La respuesta en frecuencia del filtro **322** de conformación preliminar, en una realización, se ilustra en la figura 8B.

25 Las señales $\hat{r}_2(n)$, calculada por el elemento **320**, y $\hat{r}_3(n)$, calculada por el elemento **322**, se calculan como sigue:

$$\hat{r}_2(n) = \sum_{i=1}^{M_{bp}-1} \hat{r}_1(n-i) * a_{bp}(i) + \sum_{j=0}^{N_{bp}-1} \hat{r}_1(n-j) * b_{bp}(j), n=0, \dots, N-1,$$

y

$$\hat{r}_3(n) = \sum_{i=1}^{M_{sp1}-1} \hat{r}_2(n-i) * a_{sp1}(i) + \sum_{j=0}^{N_{sp1}-1} \hat{r}_2(n-j) * b_{sp1}(j), n=0, \dots, N-1.$$

La energía de las señales $\hat{r}_2(n)$ y $\hat{r}_3(n)$ se calcula como E_2 y E_3 respectivamente. E_2 y E_3 se calculan como sigue:

$$E_2 = \sum_{i=0}^{N-1} \hat{r}_2^2(n),$$

30

y

$$E_3 = \sum_{i=0}^{N-1} \hat{r}_3^2(n).$$

En la segunda etapa del proceso de filtrado perceptivo, la señal $\hat{f}_3(n)$, emitida desde el filtro **322** de conformación preliminar, se ajusta a escala para tener la misma energía que la señal residual original $r(n)$, emitida desde el filtro **304** LPC, basándose en E_1 y E_3 .

- 5 En el analizador **324** de energía de banda ajustada a escala, la señal aleatoria ajustada a escala y filtrada, $\hat{f}_3(n)$, calculada por el elemento (**322**), se somete al mismo análisis de energía de banda realizado previamente sobre la señal residual original, $r(n)$, mediante el analizador **314** de energía de banda no ajustada a escala.

La señal, $\hat{f}_3(n)$, calculada por el elemento **322**, se calcula como sigue:

$$\hat{r}_3(n) = \sqrt{\frac{E_1}{E_3}} \hat{f}_3(n), n=0, \dots, N-1.$$

- 10 La energía de banda paso bajo de $\hat{f}_3(n)$, se indica como E_{lp2} , y la energía de banda paso alto de $\hat{f}_3(n)$ se indica como E_{hp2} . La energía de banda alta y banda baja de $\hat{f}_3(n)$ se comparan con las energías de banda alta y banda baja de $r(n)$ para determinar el siguiente filtro de conformación a utilizar en el filtro **316** de conformación final. Basándose en la comparación de $r(n)$ y $\hat{f}_3(n)$, se elige o bien ningún filtrado adicional, o bien uno de dos filtros de conformación fijos para producir la coincidencia más próxima entre $r(n)$ y $\hat{f}_3(n)$. La forma de filtro final (o ningún filtrado adicional) se determina comparando la energía de banda en la señal original con la energía de banda en la señal aleatoria.

- 15 La relación, R_l , de la energía de banda baja de la señal original con la energía de banda baja de señal aleatoria prefiltrada ajustada a escala se calcula como sigue:

$$R_l = 10 * \log_{10} (E_{lp1} / E_{lp2}).$$

La relación, R_h , de la energía de banda alta de la señal original con la energía de banda alta de señal aleatoria prefiltrada ajustada a escala se calcula como sigue:

20
$$R_h = 10 * \log_{10} (E_{hp1} / E_{hp2})$$

Si la relación R_l es inferior a -3, se utiliza un filtro de conformación final paso alto (filtro 2) para procesar adicionalmente $\hat{f}_3(n)$ para producir \hat{f}_n .

Si la relación R_h es inferior a -3, se utiliza un filtro de conformación final paso bajo (filtro 3) para procesar adicionalmente $\hat{f}_3(n)$ para producir $\hat{f}(n)$.

- 25 En caso contrario, no se realiza ningún procesamiento adicional de $\hat{f}_3(n)$, de modo que $\hat{f}(n) = \hat{f}_3(n)$.

La salida del filtro **316** de conformación final es la señal residual aleatoria cuantificada $\hat{f}(n)$. La señal $\hat{f}(n)$ se ajusta a escala para tener la misma energía que $\hat{f}_2(n)$.

La respuesta en frecuencia del filtro de conformación final paso alto (filtro 2) se muestra en la figura 8C. La respuesta en frecuencia del filtro de conformación final paso bajo (filtro 3) se muestra en la figura 8D.

- 30 Se genera un indicador de selección de filtro para indicar qué filtro (filtro 2, filtro 3, o ningún filtro) se seleccionó para el filtrado final. El indicador de selección de filtro se transmite posteriormente de modo que un descodificador pueda reproducir el filtrado final. En una realización, el indicador de selección de filtro consiste en dos bits.

- 35 La figura 4 es un diagrama de bloques detallado del descodificador **214** del habla sorda de baja tasa de transmisión de bits de alto rendimiento ilustrado en la figura 2. La figura 4 detalla el aparato y la secuencia de operaciones de una realización del descodificador de habla sorda. El descodificador de habla sorda recibe paquetes de datos sordos y sintetiza habla sorda a partir de los paquetes de datos realizando las operaciones inversas del codificador **206** de habla sorda ilustrado en la figura 2.

- 40 Los paquetes de datos sordos se introducen al descuantificador **406** de ganancia. El descuantificador **406** de ganancia realiza la operación inversa del cuantificador **308** de ganancia en el codificador sordo ilustrado en la figura 3. La salida del descuantificador **406** de ganancia es K ganancias sordas cuantificadas.

El generador **402** de número aleatorios y el selector **404** de números aleatorios realizan exactamente las mismas operaciones que el generador **310** de números aleatorios y el selector **312** de números aleatorios en el codificador sordo de la figura 3.

La salida de números aleatorios para cada subtrama del selector **404** de números aleatorios se multiplica entonces por la respectiva ganancia cuantificada de la subtrama, emitida desde el descuantificador **406** de ganancia, mediante el multiplicador **405**. La salida de señal aleatoria ajustada a escala del multiplicador **405**, $\hat{r}_1(n)$, se procesa entonces mediante un filtrado perceptivo.

5 Se realiza un proceso de filtrado perceptivo de dos etapas idéntico al proceso de filtrado perceptivo del codificador sordo de la figura 3. El filtro **408** perceptivo realiza exactamente las mismas operaciones que el filtro 318 perceptivo en el codificador sordo de la figura 3. La señal aleatoria $\hat{r}_1(n)$ se pasa a través de dos filtros fijos en el filtro **408** perceptivo. El filtro **407** paso banda y el filtro **409** de conformación preliminar son exactamente iguales que el filtro **320** paso banda y el filtro **322** de conformación preliminar utilizados en el filtro **318** perceptivo en el codificador sordo de la figura 3. Las salidas después del filtro **407** paso banda y el filtro **409** de conformación preliminar se indican como $\hat{r}_2(n)$ y $\hat{r}_3(n)$, respectivamente. Las señales $\hat{r}_2(n)$ y $\hat{r}_3(n)$ se calculan igual que en el codificador sordo de la figura 3.

15 La señal $\hat{r}_3(n)$ se filtra en el filtro **410** de conformación final. El filtro **410** de conformación final es idéntico al filtro 316 de conformación final en el codificador sordo de la figura 3. El filtro **410** de conformación final realiza o bien conformación final paso alto, conformación final paso bajo, o bien no se realiza ningún filtrado final adicional, según determine el indicador de selección de filtro generado en el codificador sordo de la figura 3 y recibido en el paquete de bits de datos en el descodificador **214**. La señal residual cuantificada de salida, $\hat{r}(n)$, del filtro **410** de conformación final se ajusta a escala para que tenga la misma energía que $\hat{r}_2(n)$.

20 La señal aleatoria cuantificada, $\hat{r}(n)$, se filtra mediante el filtro **412** de síntesis LPC para generar la señal de habla sintetizada, $\hat{s}(n)$.

Un postfiltro **414** posterior podría aplicarse a la señal de habla sintetizada, $\hat{s}(n)$, para generar el habla de salida final.

La figura 5 es un diagrama de flujo que ilustra las etapas de codificación de una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento para habla sorda.

25 En la etapa **502**, a un codificador de habla sorda (no mostrado) se proporciona una trama de datos muestras del habla digitalizadas sordas. Se proporciona una nueva trama cada 20 milisegundos. En una realización, en la que el habla sorda se muestrea a una tasa de 8 kilobits por segundo, una trama contiene 160 muestras. El flujo de control avanza hasta la etapa **504**.

En la etapa **504**, la trama de datos se filtra mediante un filtro LPC, produciendo una trama de señal residual. El flujo de control avanza hasta la etapa **506**.

30 Las etapas **506 a 516** describen etapas de procedimiento para el cálculo y la cuantificación de ganancia de una trama de señal residual.

La trama de señal residual se divide en subtramas en la etapa **506**. En una realización, cada trama se divide en diez subtramas de dieciséis muestras cada una. El flujo de control avanza hasta la etapa **508**.

35 En la etapa **508**, se calcula una ganancia para cada subtrama. En una realización se calculan ganancias de diez subtramas. El flujo de control avanza hasta la etapa **510**.

En la etapa **510**, las ganancias de subtrama se dividen en subgrupos. En una realización, 10 ganancias de subtrama se dividen en dos subgrupos de cinco ganancias de subtrama cada uno. El flujo de control avanza hasta la etapa **512**.

40 En la etapa **512**, las ganancias de cada subgrupo se normalizan, para producir un factor de normalización para cada subgrupo. En una realización, se producen dos factores de normalización para dos subgrupos de cinco ganancias cada uno. El flujo de control avanza hasta la etapa **514**.

45 En la etapa **514**, los factores de normalización producidos en la etapa **512** se convierten al dominio logarítmico, o forma exponencial, y entonces se cuantifican. En una realización, se produce un factor de normalización cuantificado, denominado en lo sucesivo en el presente documento como Índice 1. El flujo de control avanza hasta la etapa **516**.

En la etapa **516**, las ganancias normalizadas de cada subgrupo producidas en la etapa **512** se cuantifican. En una realización, se cuantifican dos subgrupos para producir dos valores de ganancia cuantificada, denominados en lo sucesivo en el presente documento como Índice 2 e Índice 3. El flujo de control avanza hasta la etapa **518**.

50 Las etapas **518 a 520** describen las etapas de procedimiento para generar una señal de habla sorda cuantificada aleatoria.

En la etapa **518**, se genera una señal de ruido aleatorio para cada subtrama. Se selecciona un porcentaje predeterminado de los números aleatorios de mayor amplitud generados por subtrama. Los números no

seleccionado se fijan a cero. En una realización, el porcentaje de números aleatorios seleccionado es del 25%. El flujo de control avanza hasta la etapa **520**.

En la etapa **520**, los números aleatorios seleccionados se ajustan a escala mediante las ganancias cuantificadas para cada subtrama producidas en la etapa **516**. El flujo de control avanza hasta la etapa **522**.

- 5 Las etapas **522 a 528** describen etapas de procedimiento para el filtrado perceptivo de la señal aleatoria. El filtrado perceptivo de las etapas **522 a 528** mejora la calidad perceptiva y mantiene la naturalidad de la señal de habla sorda cuantificada aleatoria.

En la etapa **522**, la señal de habla sorda cuantificada aleatoria se filtra paso banda para eliminar las componentes de extremo alto y bajo. El flujo de control avanza hasta la etapa **524**.

- 10 En la etapa **524**, se aplica un filtro de conformación preliminar fijo a la señal de habla sorda cuantificada aleatoria. El flujo de control avanza hasta la etapa **526**.

En la etapa **526**, se analizan las energías de banda baja y alta de la señal aleatoria y la señal residual original. El flujo de control avanza hasta la etapa **528**.

- 15 En la etapa **528**, se compara el análisis de energía de la señal residual original con el análisis de energía de la señal aleatoria, para determinar si es necesario un filtrado adicional de la señal aleatoria. Basándose en el análisis, se selecciona o bien ningún filtro, o bien uno de dos filtros finales predeterminados para filtrar adicionalmente la señal aleatoria. Los dos filtros finales predeterminados son un filtro de conformación final paso alto y un filtro de conformación final paso bajo. Se genera un mensaje de indicación de selección de filtro para indicar a un descodificador qué filtro final (o ningún filtro) se aplicó. En una realización, el mensaje de indicación de selección de filtro es de 2 bits. El flujo de control avanza hasta la etapa **530**.

- 20 En la etapa **530**, se transmiten un índice para el factor de normalización cuantificado producido en la etapa **514**, índices para las ganancias de subgrupo cuantificadas producidas en la etapa **516**, y el mensaje de indicación de selección de filtro generado en la etapa **528**. En una realización, se transmiten Índice 1, Índice 2, Índice 3, y una indicación de selección de filtro final de 2 bits. Incluyendo los bits requeridos para transmitir los índices de parámetros LPC cuantificados, la tasa de transmisión de bits de una realización es de 2 Kilobits por segundo. (La cuantificación de parámetros LPC no está dentro del alcance de las realizaciones dadas a conocer).

La figura 6 es un diagrama de flujo que ilustra las etapas de descodificación de una técnica de codificación a baja tasa de transmisión de bits de alto rendimiento para habla sorda.

- 30 En la etapa **602** se reciben un índice de factor de normalización, índices de ganancia de subgrupo cuantificadas y un indicador de selección de filtro final para una trama de habla sorda. En una realización, se reciben Índice 1, Índice 2, Índice 3, y una indicación de selección de filtro de 2 bits. El flujo de control avanza hasta la etapa **604**.

En la etapa **604** se recupera el factor de normalización a partir de tablas de consulta utilizando el índice de factor de normalización. El factor de normalización se convierte del dominio logarítmico, o forma exponencial, al dominio lineal. El flujo de control avanza hasta la etapa **606**.

- 35 En la etapa **606** se recuperan las ganancias a partir de tablas de consulta utilizando los índices de ganancia. Las ganancias recuperadas se ajustan a escala mediante los factores de normalización recuperados para recuperar las ganancias cuantificadas de cada subgrupo de la trama original. El flujo de control avanza hasta la etapa **608**.

En la etapa **608** se genera una señal de ruido aleatorio para cada subtrama, exactamente como en la codificación. Se selecciona un porcentaje predeterminado de los números aleatorios de mayor amplitud generados por subtrama.

- 40 Los números no seleccionados se fijan a cero. En una realización, el porcentaje de números aleatorios seleccionado es del 25%. El flujo de control avanza hasta la etapa **610**.

En la etapa **610**, los números aleatorios seleccionados se ajustan a escala mediante las ganancias cuantificadas para cada subtrama recuperadas en la etapa **606**.

- 45 Las etapas **612 a 616** describen etapas de procedimiento de descodificación para el filtrado perceptivo de la señal aleatoria.

En la etapa **612**, la señal de habla sorda cuantificada aleatoria se filtra paso banda para eliminar componentes de extremo alto y bajo. El filtro paso banda es idéntico al filtro paso banda utilizado en la codificación. El flujo de control avanza hasta la etapa **614**.

- 50 En la etapa **614** se aplica un filtro de conformación preliminar fijo a la señal de habla sorda cuantificada aleatoria. El filtro de conformación preliminar fijo es idéntico al filtro de conformación preliminar fijo utilizado en la codificación. El flujo de control avanza hasta la etapa **616**.

En la etapa **616**, basándose en el mensaje de indicación de selección de filtro, se selecciona o bien ningún filtro, o

bien uno de dos filtros predeterminados para filtrar adicionalmente la señal aleatoria en un filtro de conformación final. Los dos filtros predeterminados del filtro de conformación final son un filtro de conformación final paso alto (filtro 2) y un filtro de conformación final paso bajo (filtro 3) idénticos al filtro de conformación final paso alto y al filtro de conformación final paso bajo del codificador. La señal aleatoria cuantificada de salida del filtro de conformación final se ajusta a escala para que tenga la misma energía que la salida de señal del filtro paso banda. La señal aleatoria cuantificada se filtra mediante un filtro de síntesis LPC para generar una señal de habla sintetizada. Puede aplicarse un postfiltro posterior a la señal de habla sintetizada para generar el habla de salida descodificada final.

La figura 7A es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro paso bajo en los analizadores (314, 324) de energía de banda utilizados para analizar la energía de banda baja en la señal residual $r(n)$, emitida desde el filtro (304) LPC en el codificador, y en la señal aleatoria ajustada a escala y filtrada, $\hat{r}_3(n)$, emitida desde el filtro (322) de conformación preliminar en el codificador.

La figura 7B es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro paso alto en los analizadores (314, 324) de energía de banda utilizados para analizar la energía de banda alta en la señal residual $r(n)$, emitida desde el filtro (304) LPC en el codificador, y en la señal aleatoria ajustada a escala y filtrada, $\hat{r}_3(n)$, emitida desde el filtro (322) de conformación preliminar en el codificador.

La figura 8A es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro de conformación final paso banda bajo en un filtro (320,407) paso banda utilizado para conformar la señal aleatoria ajustada a escala, $\hat{r}_1(n)$, emitida desde el multiplicador (307,405) en el codificador y en el descodificador.

La figura 8B es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro de conformación final paso banda alto en el filtro (322,409) de conformación preliminar utilizado para conformar la señal aleatoria ajustada a escala, $\hat{r}_2(n)$, emitida desde el filtro (320, 407) paso banda en el codificador y en el descodificador.

La figura 8C es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro de conformación final paso alto, en el filtro (316, 410) de conformación final, utilizado para conformar la señal aleatoria ajustada a escala y filtrada, $\hat{r}_3(n)$, emitida desde el filtro (322,409) de conformación preliminar en el codificador y descodificador.

La figura 8D es un gráfico de la respuesta en frecuencia de frecuencia normalizada frente a amplitud de un filtro de conformación final paso bajo, en el filtro (316, 410) de conformación final, utilizado para conformar la señal aleatoria ajustada a escala y filtrada, $\hat{r}_3(n)$, emitida desde el filtro (322,409) de conformación preliminar en el codificador y descodificador.

La descripción anterior de las realizaciones preferidas se proporciona para permitir a cualquier experto en la técnica llevar a cabo o utilizar las realizaciones dadas a conocer. Las diversas modificaciones a estas realizaciones serán fácilmente evidentes para los expertos en la técnica, y los principios genéricos definidos en el presente documento pueden aplicarse a otras realizaciones sin el uso de la actividad inventiva. Por tanto, no se pretende que las realizaciones dadas a conocer se limiten a las realizaciones mostradas en el presente documento, sino que ha de concedérsele el alcance más amplio acorde con los principios y características novedosas dadas a conocer en el presente documento.

REIVINDICACIONES

1. Un procedimiento de descodificación de segmentos del habla sordos, que comprende:
 - recuperar (606) un grupo de ganancias cuantificadas usando índices recibidos asociados a una pluralidad de subtramas;
 - 5 generar (608) una señal de ruido aleatorio que comprende números aleatorios asociados a cada una de la pluralidad de subtramas;
 - seleccionar (608) un porcentaje predeterminado de los números aleatorios de mayor amplitud de la señal de ruido aleatorio asociados a cada una de la pluralidad de subtramas;
 - ajustar a escala (610) los números aleatorios de mayor amplitud seleccionados por las ganancias recuperadas asociadas a cada subtrama para producir una señal de ruido aleatorio ajustada a escala.
 - 10 filtrar (612) un paso banda y conformar la señal de ruido aleatorio ajustada a escala; y.
 - seleccionar (616) un segundo filtro basado en un indicador de selección de filtros recibidos y conformar además la señal de ruido aleatorio ajustada a escala con el filtro seleccionado.
2. El procedimiento de la reivindicación 1, que comprende además filtrar el ruido aleatorio ajustado a escala.
- 15 3. El procedimiento de la reivindicación 1, en el que la pluralidad de subtramas comprenden divisiones de diez subtramas por trama de habla sorda codificada.
4. El procedimiento de la reivindicación 1, en el que la pluralidad de subtramas comprende divisiones de ganancias de subtrama divididas en subgrupos.
5. El procedimiento de la reivindicación 4, en el que los subgrupos comprenden dividir un grupo de diez ganancias de subtrama en dos grupos de cinco ganancias de subtrama cada una.
- 20 6. El procedimiento de la reivindicación 3, en el que la trama de habla sorda codificada comprende 160 muestras por trama muestreadas a ocho kilohercios por segundo durante 20 milisegundos.
7. El procedimiento según la reivindicación 1, en el que el porcentaje predeterminado de números aleatorios de mayor amplitud es del veinticinco por ciento.
- 25 8. El procedimiento según la reivindicación 4, en el que dos factores de normalización se recuperan para dos subgrupos de cinco ganancias de subtrama cada uno.
9. Un descodificador (214) para descodificar segmentos de habla sorda, que comprende:
 - medios para recuperar un grupo de ganancias cuantificadas usando índices recibidos para una pluralidad de subtramas;
 - 30 medios para generar una señal de ruido aleatorio que comprende números aleatorios para cada una de la pluralidad de subtramas;
 - medios para seleccionar un porcentaje predeterminado de los números aleatorios de mayor amplitud de la señal de ruido aleatorio para cada una de la pluralidad de subtramas; y
 - medios para ajustar a escala los números aleatorios de mayor amplitud seleccionados mediante las ganancias recuperadas para cada subtrama para producir una señal de ruido aleatorio ajustada a escala.
 - 35 medios para filtrado paso banda y conformar la señal de ruido aleatorio ajustada a escala.
 - medios para seleccionar un segundo filtro basado en un indicador de selección de filtros recibidos y además ajustar a escala la señal de ruido aleatorio con el filtro seleccionado.
10. El descodificador (214) de la reivindicación 9, que comprende medios para filtrar adicionalmente el ruido aleatorio ajustado a escala.
- 40 11. El descodificador (214) de la reivindicación 9, en el que los medios para seleccionar un porcentaje predeterminado de los números aleatorios de mayor amplitud de la señal de ruido aleatorio comprenden medios para seleccionar el veinticinco por ciento de los números aleatorios de mayor amplitud.
12. El descodificador (214) de la reivindicación 9, en el que dichos medios se expresan como:
 - 45 un descuantificador (406) de ganancia configurado para recuperar dicho grupo de ganancias cuantificadas,

usando los índices recibidos para dicha pluralidad de subtramas;

un generador (402) de números aleatorios configurado para generar dicha señal de ruido aleatorio, que comprende números aleatorios para cada uno de la pluralidad de subtramas;

5 un selector (404) de números aleatorios configurado para seleccionar dicho porcentaje predeterminado de los números aleatorios de mayor amplitud de la señal de ruido aleatorio para cada una de la pluralidad de subtramas;

un selector de números aleatorios y multiplicador (405) configurado a escala de los números aleatorios de mayor amplitud seleccionados por las ganancias recuperadas para cada subtrama, para producir dicha señal de ruido aleatorio ajustada a escala.

10 un filtrado (407) paso banda y un primer filtro de conformación (409) para filtrar y conformar la señal de ruido aleatorio ajustada a escala; y

un segundo filtro de conformación (410) configurado para seleccionar dicho segundo filtro basado en un indicador de selección de filtros y la configuración adicional de la señal de ruido aleatorio ajustada a escala con el filtro seleccionado.

15 13. El descodificador de la reivindicación 12, que comprende un postfiltro (414) configurado para filtrar adicionalmente el ruido aleatorio ajustado a escala.

14. El descodificador de la reivindicación 12, en el que el selector (404) de números aleatorios configurado para seleccionar un porcentaje predeterminado de números aleatorios de mayor amplitud de la señal de ruido aleatorio está configurado además para seleccionar el 25% de números aleatorios de mayor amplitud.

20

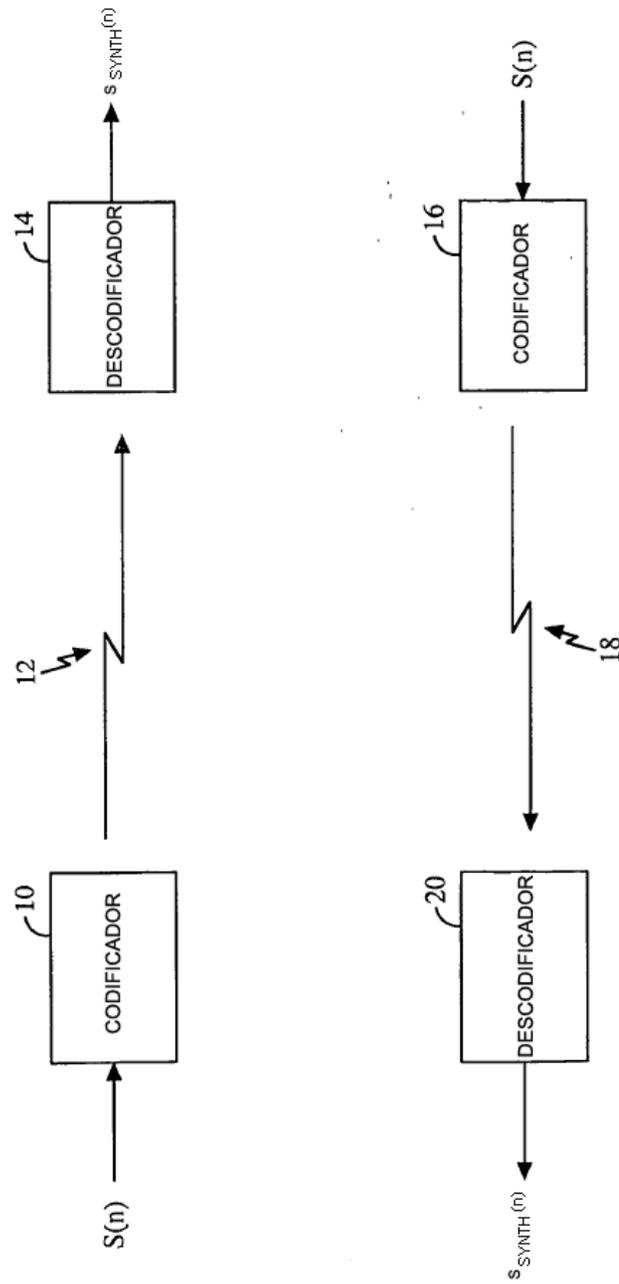


FIG. 1

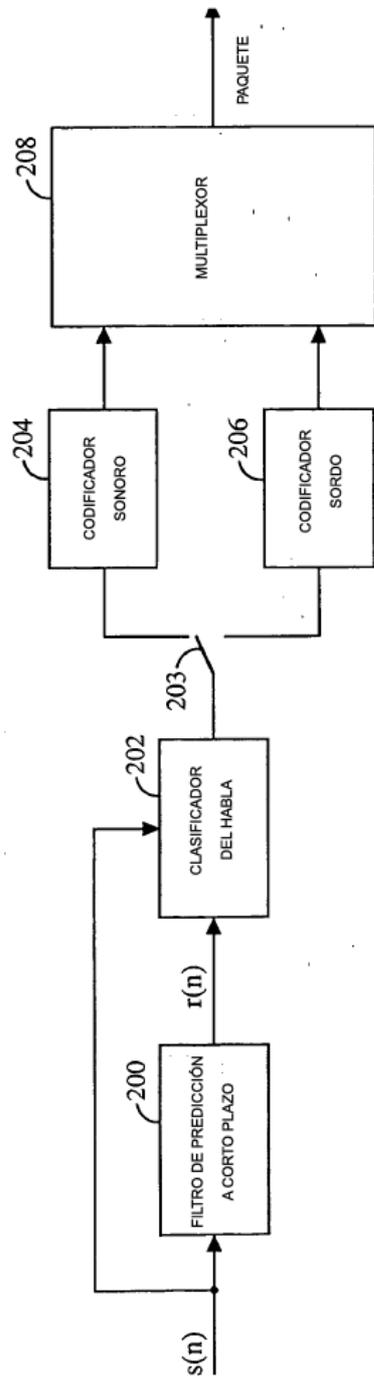


FIG. 2A

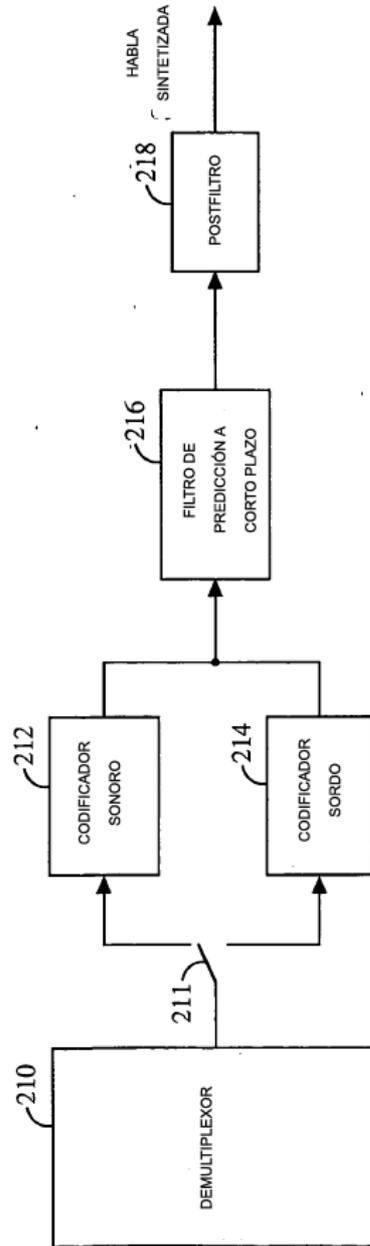
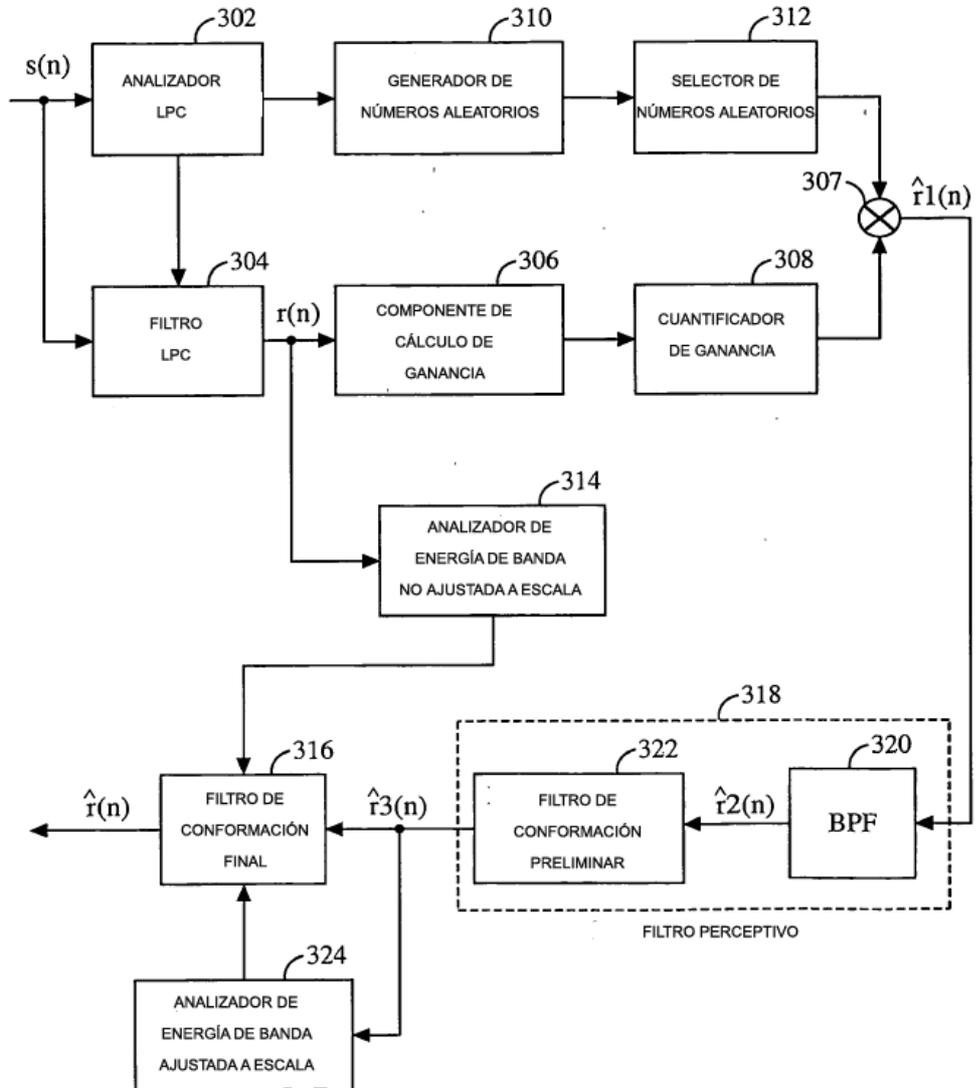


FIG. 2B



206 ↗

FIG. 3

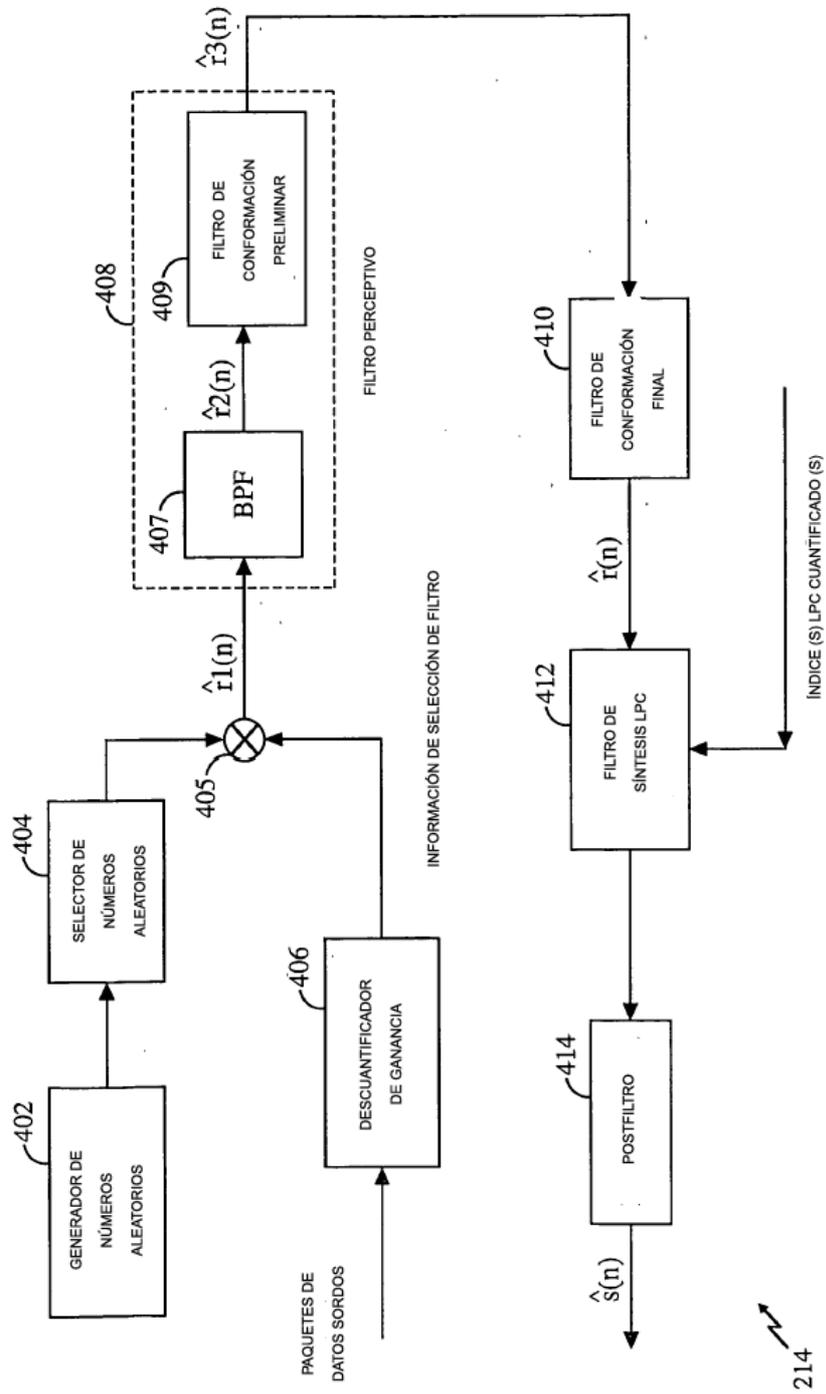


FIG. 4

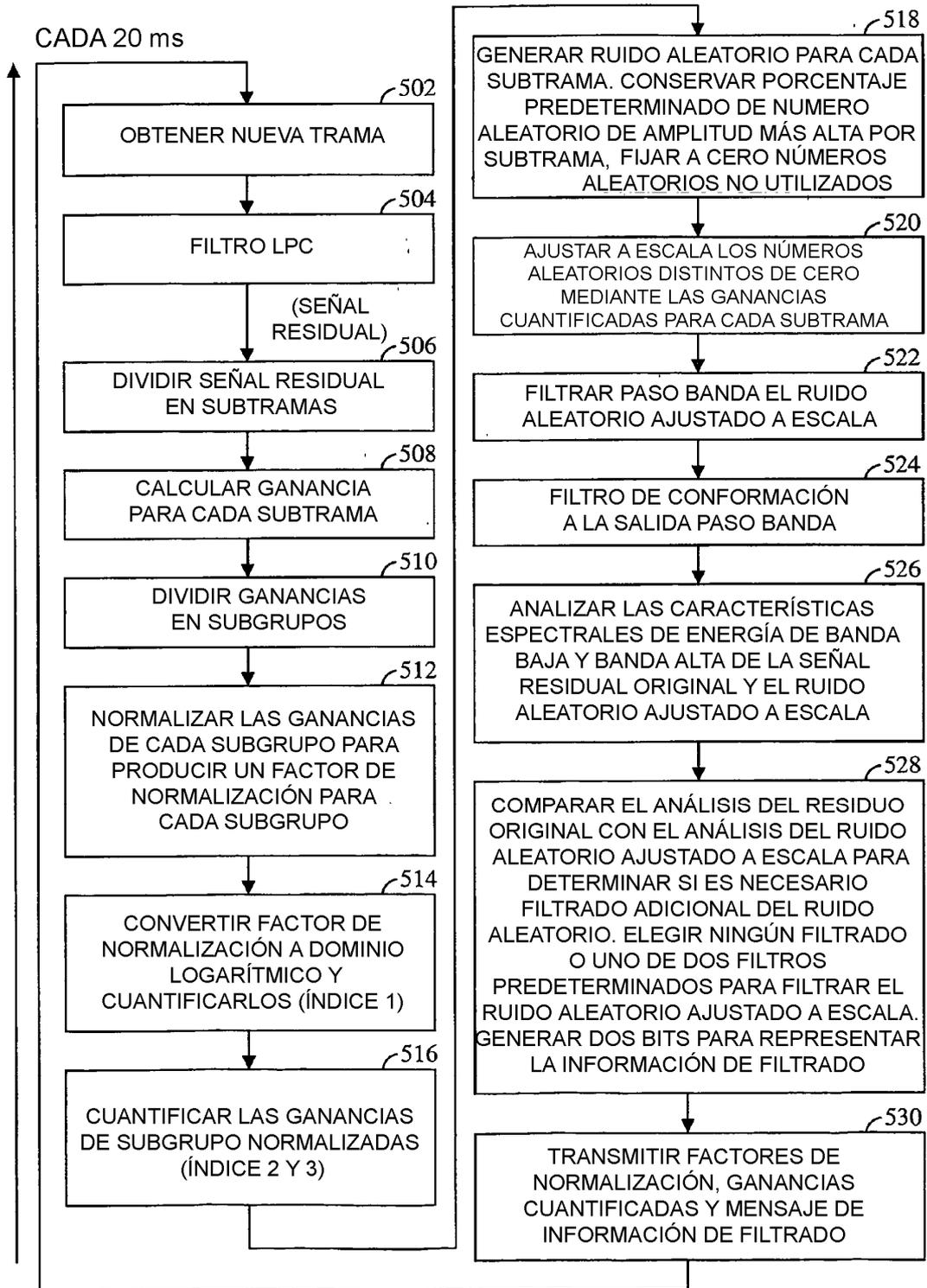


FIG. 5

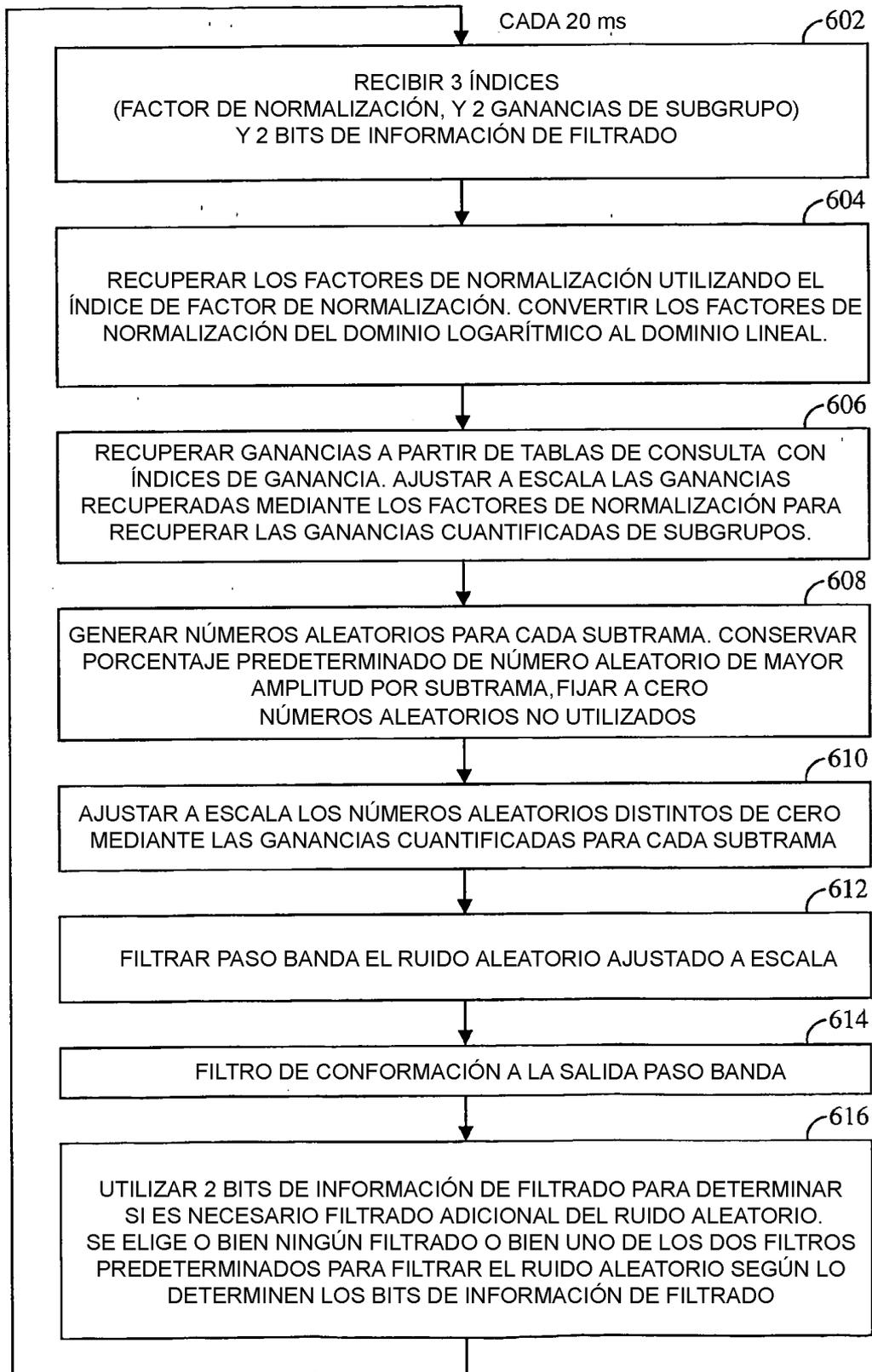


FIG. 6

RESPUESTA EN FRECUENCIA DE FILTRO PASO BAJO
PARA ANÁLISIS DE ENERGÍA DE BANDA

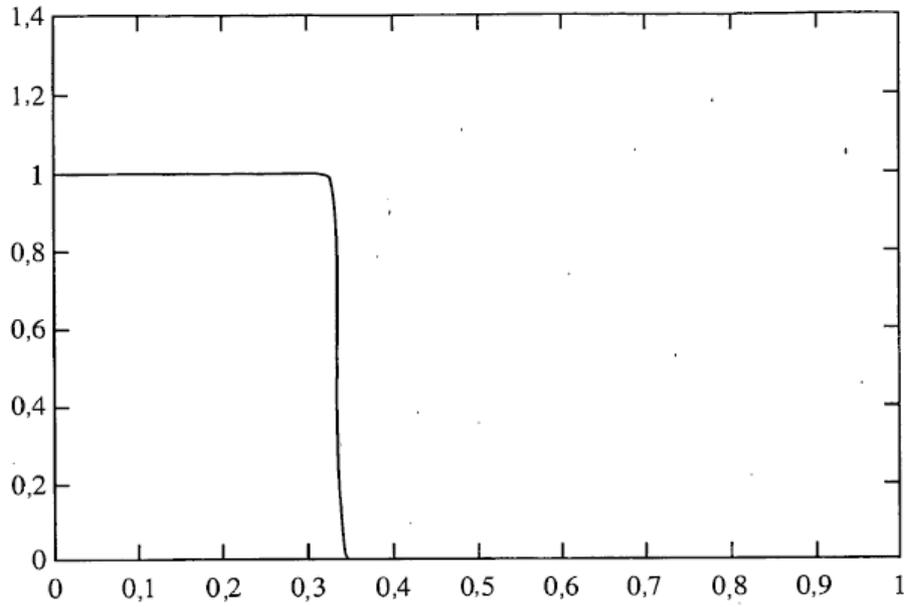


FIG. 7A

RESPUESTA EN FRECUENCIA DE FILTRO PASO ALTO
PARA ANÁLISIS DE ENERGÍA DE BANDA

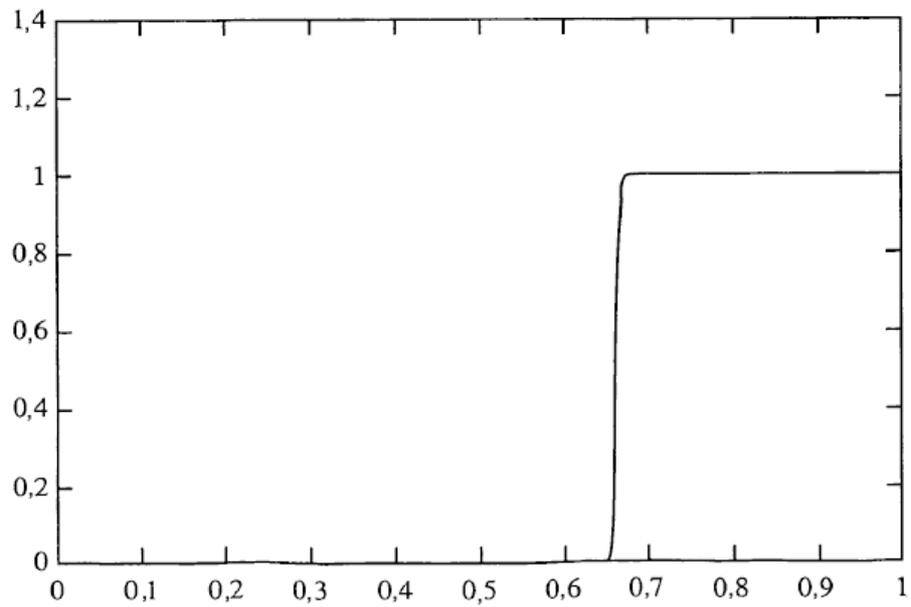


FIG. 7B

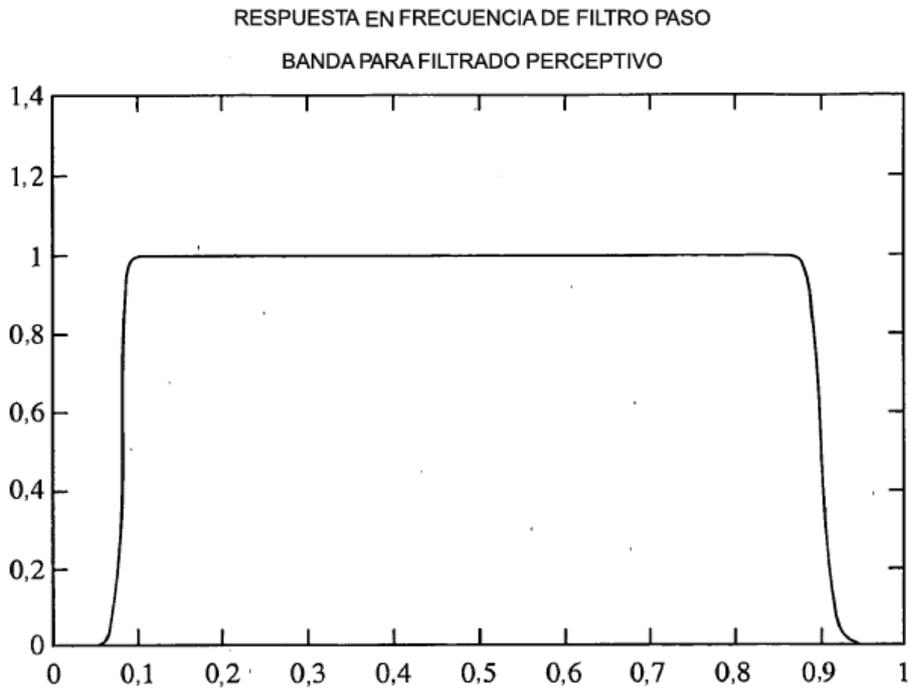


FIG. 8A

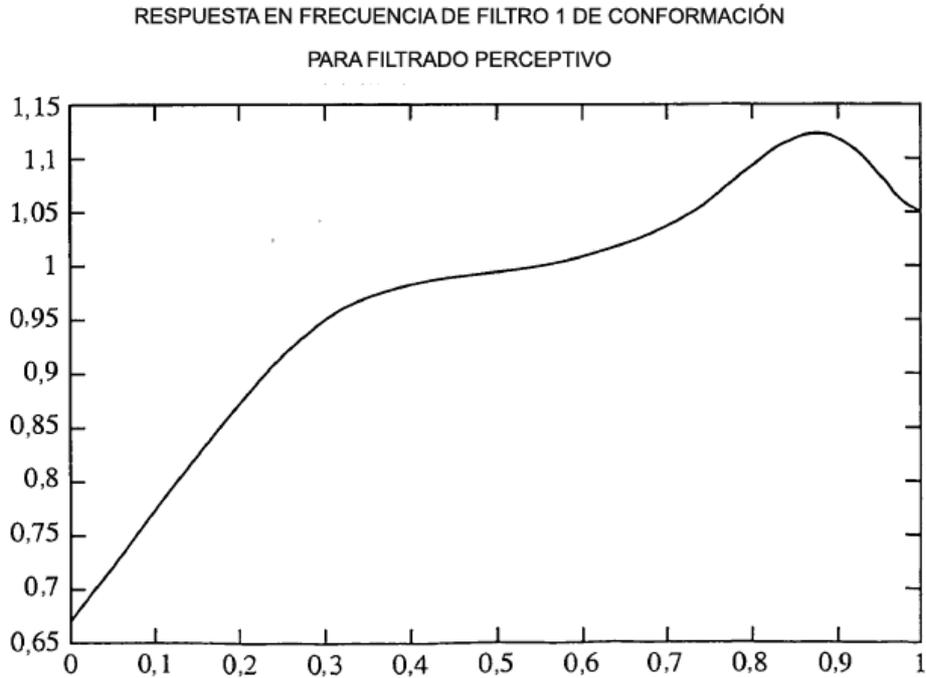


FIG. 8B

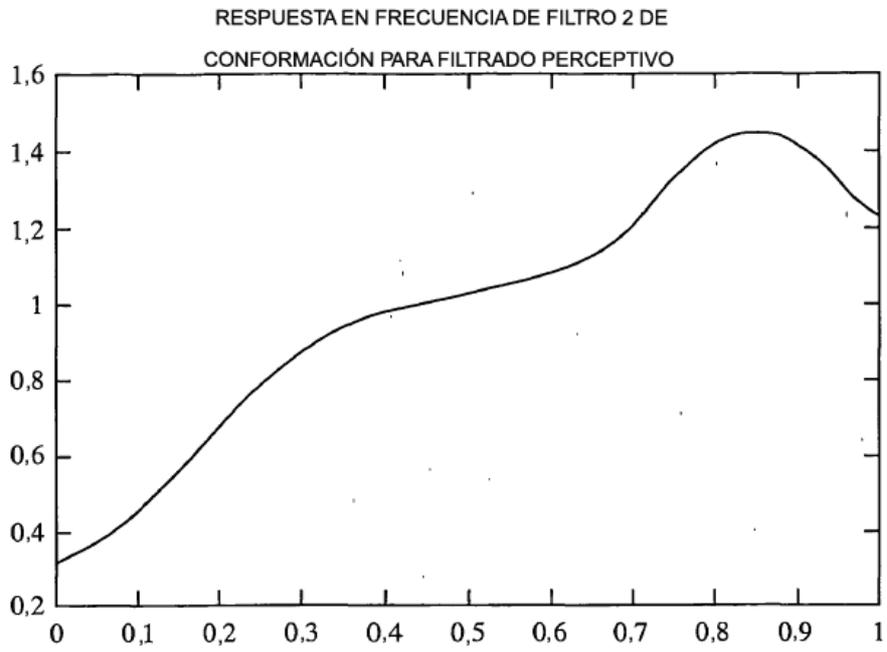


FIG. 8C

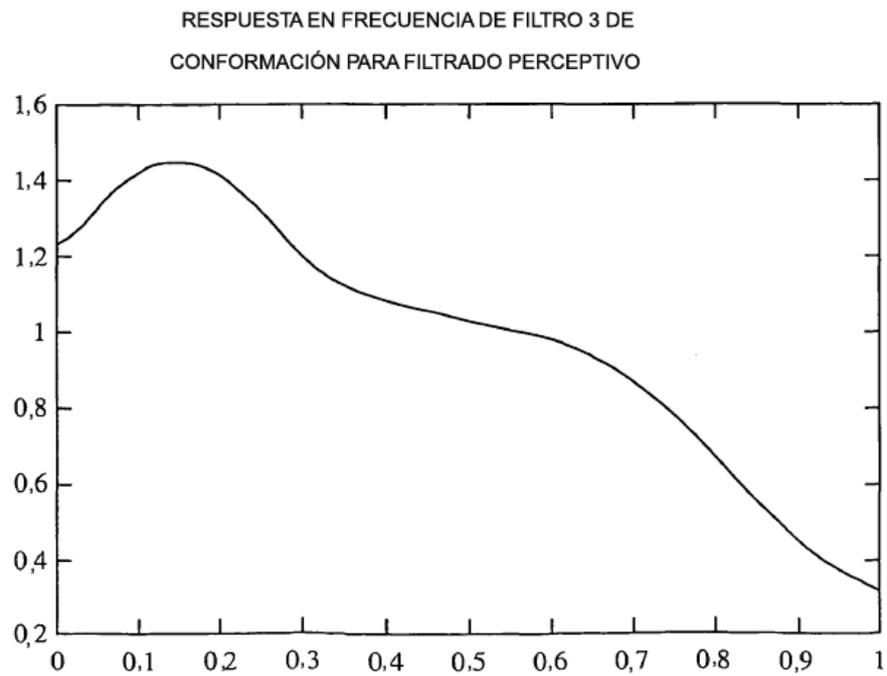


FIG. 8D