

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 616 434**

51 Int. Cl.:

G10L 19/22 (2013.01)

G10L 19/02 (2013.01)

G10L 19/08 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.01.2014 PCT/EP2014/051557**

87 Fecha y número de publicación internacional: **07.08.2014 WO2014118136**

96 Fecha de presentación y número de la solicitud europea: **28.01.2014 E 14702505 (0)**

97 Fecha y número de publicación de la concesión europea: **07.12.2016 EP 2951820**

54 Título: **Aparato y método para seleccionar uno de un primer algoritmo de codificación de audio y un segundo algoritmo de codificación de audio**

30 Prioridad:

29.01.2013 US 201361758100 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

13.06.2017

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**RAVELLI, EMMANUEL;
DÖHLA, STEFAN;
FUCHS, GUILLAUME;
FOTOPOULOU, ELENI y
HELMRICH, CHRISTIAN**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 616 434 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).

Aparato y método para seleccionar uno de un primer algoritmo de codificación de audio y un segundo algoritmo de codificación de audio

DESCRIPCIÓN

5

La presente invención se refiere a codificación de audio y, en particular, a codificación de audio conmutada, donde, para diferentes porciones de una señal de audio, la señal codificada se genera usando diferentes algoritmos de codificación.

10

Son conocidos los codificadores de audio conmutados que determinan diferentes algoritmos de codificación para diferentes porciones de la señal de audio. En general, los codificadores de audio conmutados proporcionan conmutación entre dos modos diferentes, es decir, algoritmos, tales como ACELP (Predicción Lineal por Excitación de Código Algebraico (Algebraic Code Excited Linear Prediction)) y TCX (Excitación de Codificación por Transformada) Transform Coded Excitation)).

15

El modo LPD de MPEG USAC (Codificación Unificada de Audio y del Habla MPEG) (MPEG Unified Speech Audio Coding) está basado en los dos modos diferentes ACELP y TCX. ACELP proporciona mejor calidad para señales similares al habla y similares a transitorios. TCX proporciona mejor calidad para señales similares a música y similares a ruido. El codificador decide cuál modo usar en una base trama a trama. La decisión tomada por el codificador es crítica para la calidad del códec. Una sola decisión incorrecta puede producir un fuerte artefacto, particularmente a bajas tasas de bits.

20

El enfoque más directo para decidir cuál modo usar, es una selección de modo de bucle cerrado, es decir, realizar una codificación/decodificación completa de ambos modos, a continuación calcular un criterio de selección (por ejemplo, SNR (Relación de Señal a Ruido) segmentaria) para ambos modos basándose en la señal de audio y las señales de audio codificadas/decodificadas, y finalmente elegir un modo basándose en los criterios de selección. Este enfoque generalmente produce una decisión estable y robusta. Sin embargo, también requiere una significativa cantidad de complejidad, porque hay que ejecutar ambos modos en cada trama.

25

30

Para reducir la complejidad, un enfoque alternativo es la selección de modo de bucle abierto. La selección de bucle abierto consiste en no realizar una codificación/decodificación completa de ambos modos sino, en cambio, elegir un modo usando un criterio de selección calculado con baja complejidad. A continuación se reduce la complejidad del peor caso mediante la complejidad del modo menos complejo (normalmente TCX), menos la complejidad necesaria para calcular el criterio de selección. El ahorro en complejidad normalmente es significativo, lo cual hace que este tipo de enfoque sea atractivo cuando se restringe la complejidad de peor caso de códec.

35

La norma AMR-WB+ (definida en la Norma Internacional 3GPP TS 26.290 V6.1.0 2004-12) incluye una selección de modo de bucle abierto, usada para decidir entre todas las combinaciones de ACELP/TCX20/TCX40/TCX80 en una trama de 80 ms. Se describe en la Sección 5.2.4 de 3GPP TS 26.290. También se describe en el artículo de la conferencia "Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al." y en los documentos US 7.747.430 B2 y US 7.739.120 B2 volviendo al autor de este artículo de conferencia.

40

El documento US7.747.430 B2 desvela una selección de modo de bucle abierto basada en un análisis de parámetros de predicción de largo plazo. El documento US 7.739.120 B2 desvela una selección de modo de bucle abierto basada en características de señal que indican el tipo de contenido de audio en respectivas secciones de una señal de audio, en el que, si no es viable una selección de este tipo, la selección se basa adicionalmente en una evaluación estadística llevada a cabo para secciones respectivamente vecinas.

45

La selección de modo de bucle abierto de AMR-WB+ puede describirse en dos etapas principales. En la primera etapa principal, se calculan varios rasgos en la señal de audio, tal como la desviación típica de los niveles de energía, relación de energía de baja frecuencia/alta frecuencia, energía total, distancia ISP (par espectral de inmitancia), retardos de tono y ganancias, inclinación espectral. A continuación se usan estos rasgos para elegir entre ACELP y TCX, usando un simple clasificador basado en umbral. Si se selecciona TCX en la primera etapa principal, entonces la segunda etapa principal decide entre las posibles combinaciones de TCX20/TCX40/TCX80 de una manera de bucle cerrado.

50

55

El documento WO 2012/110448 A1 desvela un enfoque para decidir entre dos algoritmos de codificación que tienen diferentes características basándose en un resultado de detección de transitorios y un resultado de calidad de una señal de audio. Además, se desvela aplicar una histéresis, en la que la histéresis se basa en las selecciones hechas en el pasado, es decir, para porciones anteriores de la señal de audio.

60

En el artículo de la conferencia "Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al.", se compara la selección de modo de bucle cerrado y de bucle abierto de AMR-WB+. Pruebas de audición subjetiva

indican que la selección de modo de bucle abierto rinde significativamente peor que la selección de modo de bucle cerrado. Pero también se muestra que la selección de modo de bucle abierto reduce en 40 % la complejidad del peor caso.

5 Es el objeto de la invención proporcionar un enfoque mejorado que permita la selección entre un primer algoritmo de codificación y un segundo algoritmo de codificación con buen rendimiento y complejidad reducida.

Este objeto se alcanza mediante un aparato de acuerdo con la reivindicación 1, un método de acuerdo con la reivindicación 13 y un programa de computadora de acuerdo con la reivindicación 23.

10 Las realizaciones de la invención están basadas en el reconocimiento que una selección de bucle abierto con rendimiento mejorado puede implementarse estimando una medida de calidad para cada uno del primer y segundo algoritmos de codificación y seleccionar uno de los algoritmos de codificación basándose en una comparación entre la primera y segunda medidas de calidad. Las medidas de calidad se estiman, es decir, la señal de audio no se codifica y decodifica realmente para obtener las medidas de calidad. Por lo tanto, las medidas de calidad pueden
15 obtenerse con complejidad reducida. Entonces se puede realizar la selección de modo usando las medidas de calidad estimadas comparables con una selección de modo de bucle cerrado.

En realizaciones preferidas de la invención, se implementa una selección de modo de bucle abierto donde primero se estiman la SNR segmentaria de ACELP y TCX con baja complejidad. Y entonces se realiza la selección de modo usando estos valores de SNR segmentaria estimados, como en una selección de modo de bucle cerrado.
20

Las realizaciones de la invención no emplean un enfoque clásico de rasgos + clasificador como se hace en la selección de modo de bucle abierto de AMR-WB+. Sino que en cambio, las realizaciones de la invención intentan estimar una medida de calidad de cada modo y seleccionar el modo que da la mejor calidad.
25

Ahora se describirán en mayor detalle realizaciones de la presente invención con referencia a los dibujos adjuntos, en los cuales:

30 La Figura 1 muestra una vista esquemática de una realización de un aparato para seleccionar uno de un primer algoritmo de codificación y un segundo algoritmo de codificación;

La Figura 2 muestra una vista esquemática de una realización de un aparato para codificar una señal de audio;

35 La Figura 3 muestra una vista esquemática de una realización de un aparato para seleccionar uno de un primer algoritmo de codificación y un segundo algoritmo de codificación;

La Figura 4a y la Figura 4b son posibles representaciones de SNR y SNR segmentaria.

40 En la siguiente descripción, a los elementos/etapas similares en los diferentes dibujos, se les hace referencia mediante los mismos signos de referencia. Se ha de indicar que en los dibujos, se han omitido algunos rasgos, tales como conexiones de señal y similares, que no son necesarios para la comprensión de la invención.

45 La Figura 1 muestra un aparato 10 para seleccionar uno de un primer algoritmo de codificación, tal como un algoritmo TCX, y un segundo algoritmo de codificación, tal como un algoritmo ACELP, tal como el codificador para codificar una porción de una señal de audio. El aparato 10 comprende un primer estimador 12 para estimar una primera medida de calidad para la porción de señal. La primera medida de calidad está asociada con el primer algoritmo de codificación. En otras palabras, el primer estimador 12 estima una primera medida de calidad que tendría la porción de la señal de audio si se codificara y decodificara usando el primer algoritmo de decodificación, sin codificar y decodificar realmente la porción de la señal de audio usando el primer algoritmo de codificación. El
50 aparato 10 comprende un segundo estimador 14 para estimar una segunda medida de calidad para la porción de señal. La segunda medida de calidad está asociada con el segundo algoritmo de codificación. En otras palabras, el segundo estimador 14 estima la segunda medida de calidad que tendría la porción de la señal de audio si se codificara y decodificara usando el segundo algoritmo de decodificación, sin codificar y decodificar realmente la porción de la señal de audio usando el segundo algoritmo de codificación. Asimismo, el aparato 10 comprende un controlador 16 para seleccionar el primer algoritmo de codificación o el segundo algoritmo de codificación basándose en una comparación entre la primera medida de calidad y la segunda medida de calidad. El controlador puede comprender una salida 18 que indica el algoritmo de codificación seleccionado.
55

60 En una realización, la primera característica asociada con el primer algoritmo de codificación es más adecuada para señales similares a música y similares a ruido, y la segunda característica de codificación asociada con el segundo algoritmo de codificación es más adecuada para señales similares al habla y similares a transitorios. En realizaciones de la invención, el primer algoritmo de codificación es un algoritmo de codificación de audio, tal como un algoritmo de codificación de transformada, por ejemplo, un algoritmo de codificación de MDCT (transformada de

coseno discreta modificada (modified discrete cosine transform)), tal como un algoritmo de codificación de TCX (excitación de codificación por transformada). Otros algoritmos de codificación por transformada pueden basarse en una transformada FFT o cualquier otra transformación o banco de filtros. En las realizaciones de la invención, el segundo algoritmo de codificación es un algoritmo de codificación del habla, tal como un algoritmo de codificación CELP (predicción lineal por excitación de código (code excited linear prediction)), tal como un algoritmo de codificación ACELP (predicción lineal por excitación de código algebraico (algebraic code excited linear prediction)).

En las realizaciones la medida de calidad representa una medida de calidad perceptual. Se puede calcular un único valor que es una estimación de la calidad subjetiva del primer algoritmo de codificación y un único valor que es una estimación de la calidad subjetiva del segundo algoritmo de codificación. El algoritmo de codificación que da la mejor calidad subjetiva estimada puede elegirse simplemente basándose en la comparación de estos dos valores. Esto es diferente de lo que se hace en la AMR-WB+ convencional donde se calcula muchos rasgos que representan diferentes características de la señal y, a continuación, se aplica un clasificador para decidir cuál algoritmo elegir.

En algunas realizaciones, se estima la respectiva medida de calidad basándose en una porción de la señal de audio ponderada, es decir, una versión ponderada de la señal de audio. En las realizaciones, la señal de audio ponderada puede definirse como una señal de audio filtrada por una función de ponderación, donde la función de ponderación es un filtro LPC ponderado $A(z/g)$ con $A(z)$ un filtro LPC y g una ponderación entre 0 y 1 tal como 0,68. Resultó que de esta manera se pueden obtener buenas medidas de calidad perceptual. Obsérvese que el filtro LPC $A(z)$ y el filtro LPC ponderado $A(z/g)$ se determinan en una etapa de pre-procesamiento y que también se usan en ambos algoritmos de codificación. En otras realizaciones, la función de ponderación puede ser un filtro lineal, un filtro FIR o un filtro de predicción lineal.

En las realizaciones, la medida de calidad es la SNR (relación de señal a ruido) segmentaria en el dominio de la señal ponderada. Resultó que la SNR segmentaria en el dominio de la señal ponderada representa una buena medida de la calidad perceptual y, por lo tanto, se puede usar como la medida de calidad de una manera beneficiosa. Esta también es la medida de calidad usada en ambos algoritmos de codificación, ACELP y TCX para estimar los parámetros de codificación.

Otra medida de calidad puede ser la SNR en el dominio de la señal ponderada. Otras medidas de calidad pueden ser la SNR segmentaria, la SNR de la correspondiente porción de la señal de audio en el dominio de la señal no ponderada, es decir, no filtrada por los coeficientes de la LPC (ponderados). Otras medidas de calidad pueden ser la distorsión cepstral o la relación de ruido a máscara (NMR).

En general, la SNR compara las señales de audio original y procesada (tales como señales del habla) muestra a muestra. Su objetivo es medir la distorsión de los codificadores de forma de onda que reproducen la forma de onda de entrada. Se puede calcular la SNR como se muestra en la Figura 5a, donde $x(i)$ e $y(i)$ son las muestras original y procesada indexadas mediante i y N es el número total de muestras. La SNR segmentaria, en lugar de funcionar sobre la señal completa, calcula el promedio de los valores de SNR de segmentos cortos, tales como de 1 a 10 ms, tal como 5 ms. Se puede calcular la SNR como se muestra en la Figura 5b, donde N y M son la longitud del segmento y el número de segmentos, respectivamente.

En algunas realizaciones de la invención, la porción de la señal de audio representa una trama de una señal de audio que se obtiene generando en ventanas la señal de audio y se realiza la selección de un algoritmo de codificación apropiado para una pluralidad de tramas sucesivas obtenidas generando en ventanas una señal de audio. En la siguiente memoria descriptiva, en relación con la señal de audio, los términos "porción" y "trama" se usan de manera intercambiable. En las realizaciones, cada trama se divide en subtramas y se estima la SNR segmentaria para cada trama calculando la SNR para cada subtrama, y se convierte en dB y se calcula el promedio de las SNR de la subtrama en dB.

Por tanto, en las realizaciones, no es la SNR (segmentaria) entre la señal de audio de entrada y la señal de audio decodificada la que se estima, sino que se estima la SNR (segmentaria) entre la señal de audio de entrada ponderada y la señal de audio decodificada ponderada. En lo que se refiere a esta SNR (segmentaria), se puede hacer referencia al capítulo 5.2.3 de la norma AMR-WB+ (Norma Internacional 3GPP TS 26.290 V6.1.0 2004-12).

En las realizaciones de la invención, la respectiva medida de calidad se estima basándose en la energía de una porción de la señal de audio ponderada y basándose en una distorsión estimada introducida cuando se codifica la porción de señal mediante el respectivo algoritmo, en el que el primer y segundo estimadores se configuran para determinar las distorsiones estimadas dependientes de la energía de una señal de audio ponderada.

En las realizaciones de la invención, se determina una distorsión de cuantificador estimada introducida por un cuantificador usado en el primer algoritmo de codificación cuando se cuantifica la porción de la señal de audio, y se determina la primera medida de calidad basándose en la energía de la porción de la señal de audio ponderada y la distorsión de cuantificador estimada. En tales realizaciones, se puede estimar una ganancia global para la porción

de la señal de audio de manera que la porción de la señal de audio produciría una tasa de bits dada objetivo cuando se codifique con un cuantificador y un codificador de entropía usado en el primer algoritmo de codificación, en el que la distorsión de cuantificador estimada se determina basándose en la ganancia global estimada. En tales realizaciones, la distorsión de cuantificador estimada puede determinarse basándose en una potencia de la ganancia
 5 estimada. Cuando el cuantificador usado en el primer algoritmo de codificación es un cuantificador escalar uniforme, el primer estimador puede configurarse para determinar la distorsión de cuantificador estimada usando la fórmula $D = G \cdot G / 12$, en la que D es la distorsión de cuantificador estimada y G es la ganancia global estimada. En caso de que el primer algoritmo de codificación use otro cuantificador, la distorsión del cuantificador puede determinarse a partir de la ganancia global de una manera diferente.

10 Los inventores reconocieron que se puede estimar una medida de calidad, tal como una SNR segmentaria, que se obtendría cuando se codifica y decodifica la porción de la señal de audio usando el primer algoritmo de codificación, tal como el algoritmo TCX, de una manera apropiada, usando los rasgos anteriores en cualquier combinación de los mismos.

15 En las realizaciones de la invención, la primera medida de calidad es una SNR segmentaria y la SNR segmentaria se estima calculando una SNR estimada asociada con cada una de una pluralidad de subporciones de la porción de la señal de audio, basándose en una energía de la correspondiente subporción de la señal de audio ponderada y la distorsión de cuantificador estimada y calculando un promedio de las SNR asociadas con las subporciones de la
 20 porción de la señal de audio ponderada para obtener la SNR segmentaria estimada para la porción de la señal de audio ponderada.

25 En las realizaciones de la invención, se determina una distorsión de libro de códigos adaptativos introducida por un libro de códigos adaptativos usado en el segundo algoritmo de codificación cuando se usa el libro de códigos adaptativos para codificar la porción de la señal de audio, y se estima la segunda medida de calidad basándose en una energía de la porción de la señal de audio ponderada y la distorsión de libro de códigos adaptativos estimada.

30 En tales realizaciones, para cada una de una pluralidad de subporciones de la porción de la señal de audio, se puede aproximar el libro de códigos adaptativos basándose en una versión de la subporción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, se puede estimar una ganancia de libro de códigos adaptativos de manera que se minimice un error entre la subporción de la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y se puede determinar una distorsión de libro de códigos adaptativos estimada basándose en la energía de un error entre la subporción de la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante
 35 la ganancia de libro de códigos adaptativos.

40 En las realizaciones de la invención, se puede reducir la distorsión de libro de códigos adaptativos estimada para cada subporción de la porción de la señal de audio en un factor constante para tener en cuenta una reducción de la distorsión que se logra mediante un libro de códigos innovador en el segundo algoritmo de codificación.

45 En las realizaciones de la invención, la segunda medida de calidad es una SNR segmentaria y la SNR segmentaria se estima calculando una SNR estimada asociada con cada subporción, basándose en la energía de la correspondiente subporción de la señal de audio ponderada y la distorsión de libro de códigos adaptativos estimada y calculando un promedio de las SNR asociadas con las subporciones para obtener la SNR segmentaria estimada.

50 En las realizaciones de la invención, se aproxima el libro de códigos adaptativos basándose en una versión de la porción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, se estima una ganancia de libro de códigos adaptativos de manera que se minimiza un error entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y se determina la distorsión de libro de códigos adaptativos estimada basándose en la energía entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante la ganancia de libro de códigos adaptativos. Por tanto, se puede determinar la distorsión de libro de códigos adaptativos estimada con baja complejidad.

55 Los inventores reconocieron que se puede estimar la medida de calidad, tal como una SNR segmentaria, que se obtendría cuando se codifica y decodifica la porción de la señal de audio usando el segundo algoritmo de codificación, tal como un algoritmo ACELP, de una manera apropiada, usando los rasgos anteriores en cualquier combinación de los mismos.

60 En las realizaciones de la invención, se usa un mecanismo de histéresis para comparar las medidas de calidad estimadas. Esto se puede hacer más estable la decisión de cuál algoritmo se ha de usar. El mecanismo de histéresis puede depender de las medidas de calidad estimadas (tal como la diferencia entre ellas) y otros parámetros, tal como estadísticas acerca de decisiones previas, el número de tramas temporalmente estacionarias, transitorios en las tramas. En lo que concierne a tales mecanismos de histéresis, se puede hacer referencia al documento WO

2012/110448 A1, por ejemplo.

En las realizaciones de la invención, un codificador para codificar una señal de audio comprende el aparato 10, una etapa para realizar el primer algoritmo de codificación y una etapa para realizar el segundo algoritmo de codificación, en el que el codificador está configurado para codificar la porción de la señal de audio usando el primer algoritmo de codificación o el segundo algoritmo de codificación dependiendo de la selección por el controlador 16. En las realizaciones de la invención, un sistema para codificar y decodificar comprende el codificador y un decodificador configurado para recibir la versión codificada de la porción de la señal de audio y una indicación del algoritmo usado para codificar la porción de la señal de audio y para decodificar la versión codificada de la porción de señal de audio usando el algoritmo indicado.

Antes de describir una realización del primer estimador 12 y el segundo estimador 14 en detalle con referencia a la Figura 3, se describe una realización de un codificador 20 con referencia a la Figura 2.

El codificador 20 comprende el primer estimador 12, el segundo estimador 14, el controlador 16, una unidad de pre-procesamiento 22, un conmutador 24, una primera etapa de codificador 26 configurada para realizar un algoritmo TCX, una segunda etapa de codificador 28 configurada para realizar un algoritmo ACELP y una interfaz de salida 30. La unidad de pre-procesamiento 22 puede ser parte de un codificador USAC común y puede estar configurada para emitir los coeficientes LPC, los coeficientes LPC ponderados, la señal de audio ponderada y un conjunto de retardos de paso. Se ha de indicar que todos estos parámetros se usan para ambos algoritmos de codificación, es decir, el algoritmo TCX y el algoritmo ACELP. Por tanto, tales parámetros no tienen que calcularse adicionalmente para decisión de modo de bucle abierto. La ventaja de usar parámetros ya calculados para la decisión de modo de bucle abierto, es ahorro de complejidad.

Se proporciona una señal de audio de entrada 40 en una línea de entrada. La señal de audio de entrada 40 se aplica al primer estimador 12, la unidad de pre-procesamiento 22 y a ambas etapas de codificador 26, 28. La unidad de pre-procesamiento 22 procesa la señal de audio de entrada de una manera convencional para derivar los coeficientes LPC y los coeficientes LPC ponderados 42 y para filtrar la señal de audio 40 con los coeficientes LPC ponderados 42 para obtener la señal de audio ponderada 44. La unidad de pre-procesamiento 22 emite los coeficientes LPC ponderados 42, la señal de audio ponderada 44 y un conjunto de retardos de paso 48. Como entienden los expertos en la materia, los coeficientes LPC ponderados 42 y la señal de audio ponderada 44 pueden segmentarse en tramas o subtramas. Se puede obtener la segmentación generando en ventanas la señal de audio de una manera apropiada.

En las realizaciones de la invención, se pueden usar coeficientes LPC cuantificados o coeficientes LPC ponderados cuantificados. Por tanto, se debe entender que la expresión "coeficientes LPC" se pretende que abarque "coeficientes LPC cuantificados" también, la expresión "coeficientes LPC ponderados" se pretende que abarque "coeficientes LPC ponderados cuantificados" también. Con relación a esto, es conveniente indicar que el algoritmo TCX de USAC usa los coeficientes LPC ponderados cuantificados para formar el espectro MCDT.

El primer estimador 12 recibe la señal de audio 40, los coeficientes LPC ponderados 42 y la señal de audio ponderada 44, estima la primera medida de calidad 46 basándose en los mismos y emite la primera medida de calidad al controlador 16. El segundo estimador 16 recibe la señal de audio ponderada 44 y el conjunto de retardos de paso 48, estima la segunda medida de calidad 50 basándose en los mismos y emite la segunda medida de calidad 50 al controlador 16. Como es conocido para los expertos en la materia, los coeficientes LPC ponderados 42, la señal de audio ponderada 44 y el conjunto de retardos de paso 48 ya están calculados en un módulo previo (esto es, la unidad de pre-procesamiento 22) y por lo tanto, están disponibles sin coste.

El controlador toma una decisión para seleccionar el algoritmo TCX o el algoritmo ACELP basándose en una comparación de las medidas de calidad recibidas. Como se ha indicado anteriormente, el controlador puede usar un mecanismo de histéresis para decidir cuál algoritmo usar. La selección de la primera etapa de codificador 26 o la segunda etapa de codificador 28 se muestra esquemáticamente en la Figura 2 por medio del conmutador 24 que está controlado por una señal de control 52 emitida por el controlador 16. La señal de control 52 indica si se ha de usar la primera etapa de codificador 26 o la segunda etapa de codificador 28. Basándose en la señal de control 52, las señales requeridas esquemáticamente indicadas mediante la flecha 54 en la Figura 2 y que incluyen al menos los coeficientes LPC, los coeficientes LPC ponderados, la señal de audio, la señal de audio ponderada, el conjunto de retardos de paso, se aplican, o bien a la primera etapa de codificador 26 o bien a la segunda etapa de codificador 28. La etapa de codificador seleccionada aplica el algoritmo de codificación asociado y emite la representación codificada 56 o 58 a la interfaz de salida 30. La interfaz de salida 30 puede estar configurada para emitir una señal de audio codificada que puede comprender entre otros datos, la representación codificada 56 o 58, los coeficientes LPC o los coeficientes LPC ponderados, parámetros para el algoritmo de codificación seleccionado e información acerca del algoritmo de codificación seleccionado.

Ahora se describen con referencia a la Figura 3, realizaciones específicas para estimar la primera y segunda

medidas de calidad, en las que la primera y segunda medidas de calidad son SNR segmentarias en el dominio de la señal ponderada. La Figura 3 muestra el primer estimador 12 y el segundo estimador 14 y las funcionalidades de los mismos en forma de diagramas de flujo que muestran la respectiva estimación etapa a etapa.

5 Estimación de la SNR segmentaria de TCX

El primer estimador (TCX) recibe la señal de audio 40 (señal de entrada), los coeficientes LPC ponderados 42 y la señal de audio ponderada 44 como entradas.

10 En la etapa 100, la señal de audio 40 se genera en ventanas. La generación en ventanas puede tener lugar con una ventana senoidal de baja superposición de 10 ms. Cuando la trama pasada es ACELP, el tamaño de bloque puede incrementarse en 5 ms, el lado izquierdo de la ventana puede ser rectangular y la respuesta de impulso cero en ventanas del filtro de síntesis ACELP puede eliminarse de la señal de entrada en ventanas. Esto es similar a lo que se hace en el algoritmo TCX. Una trama de la señal de audio 40, que representa una porción de la señal de audio, se emite desde la etapa 100.

15 En la etapa 102, la señal de audio en ventanas, es decir, la trama resultante, se transforma con una MDCT (transformada de coseno discreta modificada). En la etapa 104 se realiza conformación de espectro conformando el espectro de MDCT con los coeficientes LPC ponderados.

20 En la etapa 106 se estima una ganancia global G de manera que el espectro ponderado cuantificado con la ganancia G produciría un objetivo dado R, cuando se codifica con un codificador de entropía, por ejemplo, un codificador aritmético. Se usa la expresión "ganancia global" ya que se determina una ganancia para todo la trama.

25 Ahora se explica un ejemplo de una implementación de la estimación de ganancia global. Se ha de indicar que esta estimación de ganancia global es apropiada para realizaciones en las que el algoritmo de codificación TCX usa un cuantificador escalar con un codificador aritmético. Se supone un cuantificador escalar de este tipo con un codificador aritmético en la norma MPEG USAC.

30 Inicialización

En primer lugar se inicializan las variables usadas para la estimación de ganancia mediante:

- 35 1. Establecer $en[i] = 9,0 + 10,0 \cdot \log_{10}(c[4 \cdot i + 0] + c[4 \cdot i + 1] + c[4 \cdot i + 2] + c[4 \cdot i + 3])$, donde $0 \leq i < L/4$, $c[]$ es el vector de coeficientes a cuantificar, y L es la longitud de $c[]$.
 2. Establecer $fac = 128$, desplazamiento = fac y objetivo = cualquier valor (por ejemplo 1000)

Iteración

40 A continuación, el siguiente bloque de operaciones se realiza NITER veces (por ejemplo, en ese punto, NITER = 10).

1. $fac = fac/2$
 2. $desplazamiento = desplazamiento - fac$
 3. $ener = 0$
 45 4. para todo i donde $0 \leq i < L/4$ hacer lo siguiente:
 si $en[i] - desplazamiento > 3,0$, entonces $ener = ener + en[i] - desplazamiento$
 5. si $ener > objetivo$, entonces $desplazamiento = desplazamiento + fac$
 50

El resultado de la iteración es el valor de desplazamiento. Después de la iteración, se estima la ganancia global según $G = 10^{(desplazamiento/20)}$.

55 La manera específica en la que se estima la ganancia global puede variar dependiendo del cuantificador y del codificador de entropía usados. En la norma MPEG USAC se supone un cuantificador escalar con un codificador aritmético. Otros enfoques TCX pueden usar un cuantificador diferente y los expertos en la materia entienden cómo estimar la ganancia global para tales cuantificadores diferentes. Por ejemplo, la norma AMR-WB+ supone que se usa un cuantificador de red RE8. Para un cuantificador de este tipo, la estimación de la ganancia global podría estimarse como se describe en el capítulo 5.3.5.7 en la página 34 de 3GPP TS 26.290 V6.1.0 2004-12, en la que se supone una tasa de bits objetivo fija.

60 Después de haber estimado la ganancia global en la etapa 106, tiene lugar la estimación de distorsión en la etapa 108. Para ser más específico, la distorsión de cuantificador se aproxima basándose en la ganancia global estimada. En el presente documento realización se supone que se usa un cuantificador escalar uniforme. Por tanto, se

determina la distorsión de cuantificador con la fórmula sencilla $D=G*G/12$, en la cual D representa la distorsión del cuantificador determinada y G representa la ganancia global estimada. Esto corresponde a la aproximación de alta tasa de una distorsión de cuantificador escalar uniforme.

5 Basándose en la distorsión de cuantificador determinada, en la etapa 110 se realiza el cálculo de la SNR segmentaria. Se calcula la SNR en cada subtrama como la relación de la energía de señal de audio ponderada y la distorsión D que se supone constante en las subtramas. Por ejemplo se divide la trama en cuatro subtramas consecutivas (véase la Figura 4). La SNR segmentaria a continuación es el promedio de las SNR de las cuatro subtramas y puede indicarse en dB.

10 Este enfoque permite la estimación de la primera SNR segmentaria que se obtendría cuando se codifica y decodifica realmente la trama objeto usando el algoritmo TCX, sin embargo, sin tener que realmente codificar y decodificar la señal de audio y, por lo tanto, con una complejidad fuertemente reducida y reducido tiempo de cálculo.

15 Estimación de la SNR segmentaria de ACELP

El segundo estimador 14 recibe la señal de audio ponderada 44 y el conjunto de retardos de paso 48 que ya está calculado en la unidad de pre-procesamiento 22.

20 Como se muestra en la etapa 112, en cada subtrama, se aproxima el libro de códigos adaptativos simplemente usando la señal de audio ponderada y el retardo de paso T. El libro de códigos adaptativos se aproxima mediante

$$xw(n-T), n = 0, \dots, N$$

25 en la que xw es la señal de audio ponderada, T es el retardo de paso de la correspondiente subtrama y N es la longitud de la subtrama. Por consiguiente, el libro de códigos adaptativos se aproxima usando una versión de la subtrama desplazada al pasado por T. Por tanto, en las realizaciones de la invención, el libro de códigos adaptativos se aproxima de una manera muy sencilla.

30 En la etapa 114 se determina una ganancia de libro de códigos adaptativos para cada subtrama. Para ser más específicos, en cada subtrama, se estima la ganancia de libro de códigos G de manera que minimice el error entre la señal de audio ponderada y el libro de códigos adaptativos aproximado. Esto se puede hacer simplemente comparando las diferencias entre ambas señales para cada muestra y hallando una ganancia de manera que la suma de estas diferencias sea mínima.

35 En la etapa 116 se determina la distorsión de libro de códigos adaptativos para cada subtrama. En cada subtrama, la distorsión D introducida por el libro de códigos adaptativos es simplemente la energía del error entre la señal de audio ponderada y el libro de códigos adaptativos aproximado puesto a escala por la ganancia G.

40 Las distorsiones determinadas en la etapa 116 pueden ajustarse en una etapa opcional 118 para tener en cuenta el libro de códigos innovador. La distorsión del libro de códigos innovador usada en los algoritmos ACELP puede estimarse simplemente como un valor constante. En la realización de la invención descrita, simplemente se supone que el libro de códigos innovador reduce la distorsión D en un factor constante. Por tanto, las distorsiones obtenidas en la etapa 116 para cada subtrama pueden multiplicarse en la etapa 118 por un factor constante, tal como un factor constante del orden de 0 a 1, tal como 0,055.

45 En la etapa 120 tiene lugar el cálculo de la SNR segmentaria. En cada subtrama, se calcula la SNR como la relación de la energía de señal de audio ponderada y la distorsión D. La SNR segmentaria entonces es la media de la SNR de las cuatro subtramas y puede indicarse en dB.

50 Este enfoque permite la estimación de la segunda SNR que se obtendría cuando se codifica y decodifica realmente la trama del caso usando el algoritmo ACELP, sin embargo, sin tener que realmente codificar y decodificar la señal de audio y, por lo tanto, con una complejidad fuertemente reducida y reducido tiempo de cálculo.

55 El primer y segundo estimadores 12 y 14 emiten las SNR segmentarias estimadas 46, 50 al controlador 16 y el controlador 16 toma una decisión de cuál algoritmo ha de usarse para la porción asociada de la salida de audio basándose en las SNR segmentarias estimadas 46, 50. El controlador puede usar opcionalmente un mecanismo de histéresis para hacer que la decisión sea más estable. Por ejemplo, se puede usar el mismo mecanismo de histéresis que en la decisión de bucle cerrado con parámetros de ajuste levemente diferentes. Un mecanismo de histéresis de este tipo puede calcular un valor "dsnr" que puede depender de las SNR segmentarias estimadas (tal como la diferencia entre ellas) y otros parámetros, tal como estadísticas acerca de decisiones previas, el número de tramas temporalmente estáticas y componentes transitorios en las tramas.

60 Sin un mecanismo de histéresis, el controlador puede seleccionar el algoritmo de codificación que tiene la SNR estimada más alta, es decir, se selecciona ACELP si la segunda SNR estimada es menos alta que la primera SNR

estimada y se selecciona TCX si la primera SNR estimada es más alta que la segunda SNR estimada. Con un mecanismo de histéresis, el controlador puede seleccionar el algoritmo de codificación de acuerdo con la siguiente regla de decisión, en la que $acelp_snr$ es la segunda SNR estimada y tcx_snr es la primera SNR estimada:

5 si $acelp_snr + dsnr > tcx_snr$ entonces seleccionar ACELP, en otro caso seleccionar TCX.

Por consiguiente, las realizaciones de la invención permiten estimar SNR segmentarias y selección de un algoritmo de codificación apropiado de una manera sencilla y precisa.

10 En las realizaciones anteriores, las SNR segmentarias se estiman calculando un promedio de SNR estimadas para respectivas subtramas. En realizaciones alternativas, la SNR de una trama entera podría estimarse sin dividir la trama en subtramas.

15 Las realizaciones de la invención permiten una fuerte reducción en el tiempo de cálculo cuando se compara con una selección de bucle cerrado ya que se omite un número de etapas requeridas en la selección de bucle cerrado. Por consiguiente, se puede ahorrar un gran número de etapas y el tiempo de cálculo asociado con ellas mediante el enfoque inventivo mientras que aún se permite la selección de un algoritmo de codificación apropiado con buen rendimiento.

20 A pesar de que se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o a un rasgo de una etapa de método. Análogamente, aspectos descritos en el contexto de una etapa de método también representan una descripción de un correspondiente bloque o elemento o rasgo de un aparato correspondiente.

25 Las realizaciones de los aparatos descritos aquí y los rasgos de los mismos pueden implementarse mediante una computadora, uno o más procesadores, uno o más microprocesadores, campos de matrices de puertas programables (FPGA), circuitos integrados específicos de la aplicación (ASIC) y similares o combinaciones de los mismos, que están configurados o programados para proporcionar las funcionalidades descritas.

30 Algunas o todas las etapas del método pueden ejecutarse por (o usando) un aparato de hardware, como por ejemplo, un microprocesador, una computadora programable o un circuito electrónico. En algunas realizaciones, alguna o más de las etapas del método más importantes pueden ejecutarse mediante un aparato de este tipo.

35 Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un CD, un Blu-Ray, una ROM, una PROM y EPROM, una EEPROM o una memoria FLASH, teniendo señales de control legibles electrónicamente almacenadas en las mismas, que cooperan (o son capaces de cooperar) con un sistema de computadora programable de manera que se realice el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por computadora.

40 Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema de computadora programable, de manera que se realice uno de los métodos descrito en el presente documento.

45 En general, las realizaciones de la presente invención pueden implementarse como un producto de programa de computador con un código de programa, siendo código de programa operativo para realizar uno de los métodos cuando el producto de programa de computadora se ejecuta en una computadora. El código de programa puede almacenarse, por ejemplo, en un portador legible por una máquina.

Otras realizaciones comprenden el programa de computadora para realizar uno de los métodos descritos en el presente documento, almacenado en un portador legible por una máquina.

55 En otras palabras, una realización del método inventivo es, por lo tanto, un programa de computadora que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa de computadora se ejecuta en una computadora.

60 Una realización adicional del método inventivo es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa de computadora para realizar uno de los métodos descritos en el presente documento. El portador de datos, el medio de almacenamiento digital o el medio grabado típicamente son tangibles y/o no transitorios.

Una realización adicional del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales

que representan el programa de computadora para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden configurarse, por ejemplo, para transferirse mediante una conexión de comunicación de datos, por ejemplo, mediante Internet.

5 Una realización adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo de lógica programable, configurado para o programado para, realizar uno de los métodos descritos en el presente documento.

10 Una realización adicional comprende una computadora que tiene instalado en la misma el programa de computadora para realizar uno de los métodos descritos en el presente documento.

15 Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa de computadora para realizar uno de los métodos descritos en este punto, a un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de computadora al receptor.

20 En algunas realizaciones puede usarse un dispositivo de lógica programable (por ejemplo, un campo de matrices de puertas programables) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, el campo de matrices de puertas programables puede cooperar con un microprocesador para realizar uno de los métodos descritos en el presente documento. En general, los métodos preferentemente se realizan mediante algún aparato de hardware.

25 Las realizaciones anteriormente descritas son meramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y de los detalles descritos en el presente documento serán evidentes para los expertos en la materia. Por lo tanto, es la intención que la invención esté limitado solo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

REIVINDICACIONES

1. Aparato (10) para seleccionar uno de un primer algoritmo de codificación que tiene una primera característica, y un segundo algoritmo de codificación que tiene una segunda característica para codificar una porción de una señal de audio (40) para obtener una versión codificada de la porción de la señal de audio (40), que comprende:
- 5 un primer estimador (12) para estimar una primera medida de calidad para la porción de la señal de audio, estando asociada la primera medida de calidad con el primer algoritmo de codificación, sin codificar y decodificar realmente la porción de la señal de audio usando el primer algoritmo de codificación;
- 10 un segundo estimador (14) para estimar una segunda medida de calidad para la porción de la señal de audio, estando asociada la segunda medida de calidad con el segundo algoritmo de codificación, sin codificar y decodificar realmente la porción de la señal de audio usando el segundo algoritmo de codificación; y
- 15 un controlador (16) para seleccionar el primer algoritmo de codificación o el segundo algoritmo de codificación basándose en una comparación entre la primera medida de calidad y la segunda medida de calidad, en el que la primera y segunda medidas de calidad son relaciones de señal a ruido, SNR, o SNR segmentarias de la porción correspondiente de una versión ponderada de la señal de audio.
2. Aparato (10) de acuerdo con la reivindicación 1, en el que el primer algoritmo de codificación es un algoritmo de codificación por transformada, un algoritmo de codificación basado en transformada de coseno discreta modificada, MDCT, o un algoritmo de codificación de excitación de codificación por transformada, TCX, y en el que el segundo algoritmo de codificación es un algoritmo de codificación de predicción lineal por excitación de código, CELP, o un algoritmo de codificación de predicción lineal por excitación de código algebraico, ACELP.
3. Aparato (10) de acuerdo con una de las reivindicaciones 1 o 2, en el que el primer estimador (12) está configurado para determinar una distorsión de cuantificador estimada que introduciría un cuantificador usado en el primer algoritmo de codificación cuando se cuantifica la porción de la señal de audio y para estimar la primera medida de calidad basándose en una energía de una porción de una versión ponderada de la señal de audio y la distorsión de cuantificador estimada.
- 30 4. Aparato (10) de acuerdo con la reivindicación 3, en el que el primer estimador (12) está configurado para estimar una ganancia global para la porción de la señal de audio de manera que la porción de la señal de audio produciría una tasa de bits objetivo dada cuando se codifica con un cuantificador y un codificador de entropía usado en el primer algoritmo de codificación, en el que el primer estimador (12) además está configurado para determinar la distorsión de cuantificador estimada basándose en una potencia de la ganancia global estimada, en el que el cuantificador usado en el primer algoritmo de codificación es un cuantificador escalar uniforme y en el que el primer estimador (12) está configurado para determinar la distorsión de cuantificador estimada usando la fórmula $D = G \cdot G / 12$, en la que D es la distorsión de cuantificador estimada y G es la ganancia global estimada.
- 40 5. Aparato (10) de acuerdo con una de las reivindicaciones 3 o 4, en el que la primera medida de calidad es una SNR segmentaria de una porción de la señal de audio ponderada y en el que el primer estimador (12) está configurado para estimar la SNR segmentaria calculando una SNR estimada asociada con cada una de una pluralidad de subporciones de la porción de la señal de audio ponderada, basándose en una energía de las correspondientes subporciones de la señal de audio ponderada y la distorsión de cuantificador estimada y calculando un promedio de las SNR asociadas con las subporciones de la porción de la señal de audio ponderada para obtener la SNR segmentaria estimada para la porción de la señal de audio ponderada.
- 45 6. Aparato (10) de acuerdo con una de las reivindicaciones 1 a 5, en el que el segundo estimador (14) está configurado para determinar una distorsión de libro de códigos adaptativos estimada que introduciría un libro de códigos adaptativos usado en el segundo algoritmo de codificación cuando se usa un libro de códigos adaptativos para codificar la porción de la señal de audio, y en el que el segundo estimador (14) está configurado para estimar la segunda medida de calidad basándose en una energía de una porción de una versión ponderada de la señal de audio y la distorsión de libro de códigos adaptativos estimada, en el que, para cada una de una pluralidad de subporciones de la porción de la señal de audio, el segundo estimador (14) está configurado para aproximar el libro de códigos adaptativos basándose en una versión de la subporción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, para estimar una ganancia de libro de códigos adaptativos de manera que se minimice un error entre la subporción de la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y para determinar la distorsión de libro de códigos adaptativos estimada basándose en la energía de un error entre la subporción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante la ganancia de libro de códigos adaptativos.
- 50 55 60 7. Aparato (10) de acuerdo con la reivindicación 6, en el que el segundo estimador (14) además está configurado para reducir la distorsión de libro de códigos adaptativos estimada determinada para la subporción de la porción de la señal de audio, en un factor constante.

8. Aparato (10) de acuerdo con la reivindicación 6 o 7, en el que la segunda medida de calidad es una SNR segmentaria de la porción de la señal de audio ponderada, y en el que el segundo estimador (14) está configurado para estimar la SNR segmentaria calculando una SNR estimada asociada con cada subporción basándose en la energía de la correspondiente subporción de la señal de audio ponderada, y la distorsión de libro de códigos adaptativos estimada y calculando un promedio de las SNR asociadas con las subporciones para obtener la SNR segmentaria estimada para la porción de la señal de audio ponderada.
9. Aparato (10) de acuerdo con la reivindicación 6, en el que el segundo estimador (14) está configurado para aproximar el libro de códigos adaptativos basándose en una versión de la porción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, para estimar una ganancia de libro de códigos adaptativos de manera que se minimice un error entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y para determinar la distorsión de libro de códigos adaptativos estimada basándose en la energía de un error entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante la ganancia de libro de códigos adaptativos.
10. Aparato (10) de acuerdo con una de las reivindicaciones 1 a 9, en el que el controlador (16) está configurado para utilizar una histéresis para comparar las medidas de calidad estimadas.
11. Aparato (20) para codificar una porción de una señal de audio, que comprende el aparato (10) de acuerdo con una de las reivindicaciones 1 a 10, una primera etapa de codificador (26) para realizar el primer algoritmo de codificación y una segunda etapa de codificador (28) para realizar el segundo algoritmo de codificación, en el que el aparato para codificar (20) está configurado para codificar la porción de la señal de audio usando el primer algoritmo de codificación o el segundo algoritmo de codificación dependiendo de la selección por el controlador (16).
12. Sistema para codificar y decodificar que comprende un aparato (20) para codificar de acuerdo con la reivindicación 11 y un decodificador configurado para recibir la versión codificada de la porción de la señal de audio y una indicación del algoritmo usado para codificar la porción de la señal de audio y para decodificar la versión codificada de la porción de señal de audio usando el algoritmo indicado.
13. Método para seleccionar uno de un primer algoritmo de codificación que tiene una primera característica, y un segundo algoritmo de codificación que tiene una segunda característica para codificar una porción de una señal de audio para obtener una versión codificada de la porción de la señal de audio, que comprende:
- estimar una primera medida de calidad para la porción de la señal de audio, estando asociada la primera medida de calidad con el primer algoritmo de codificación, sin codificar y decodificar realmente la porción de la señal de audio usando el primer algoritmo de codificación;
 - estimar una segunda medida de calidad para la porción de la señal de audio, estando asociada la segunda medida de calidad con el segundo algoritmo de codificación, sin codificar y decodificar realmente la porción de la señal de audio usando el segundo algoritmo de codificación; y
 - seleccionar el primer algoritmo de codificación o el segundo algoritmo de codificación basándose en una comparación entre la primera medida de calidad y la segunda medida de calidad, en el que la primera y segunda medidas de calidad son relaciones de señal a ruido, SNR, o SNR segmentarias de la porción correspondiente de una versión ponderada de la señal de audio.
14. Método de acuerdo con la reivindicación 13, en el que el primer algoritmo de codificación es un algoritmo de codificación por transformada, un algoritmo de codificación basado en transformada de coseno discreta modificada, MDCT, o un algoritmo de codificación de excitación de codificación por transformada, TCX, y en el que el segundo algoritmo de codificación es un algoritmo de codificación de predicción lineal por excitación de código, CELP, o un algoritmo de codificación de predicción lineal por excitación de código algebraico, ACELP.
15. Método de acuerdo con una de las reivindicaciones 13 o 14, que comprende determinar (108) una distorsión de cuantificador estimada que introduciría un cuantificador usado en el primer algoritmo de codificación cuando se cuantifica la porción de la señal de audio y determinar la medida de calidad basándose en una energía de una porción de una versión ponderada de la señal de audio y la distorsión de cuantificador estimada.
16. Método de acuerdo con la reivindicación 15, que comprende estimar (106) una ganancia global para la porción de la señal de audio de manera que la porción de la señal de audio produciría una tasa de bits objetivo dada cuando se codifica con un cuantificador y un codificador de entropía usando en el primer algoritmo de codificación, y determinar (108) la distorsión de cuantificador estimada basándose en una potencia de la ganancia global estimada, en el que el cuantificador es un cuantificador escalar uniforme, en el que la distorsión de cuantificador estimada se determina usando la fórmula $D = G^2/12$, en la que D es la distorsión de cuantificador estimada y G es la ganancia global estimada.

17. Método de acuerdo con una de las reivindicaciones 15 o 16, en el que la primera medida de calidad es una SNR segmentaria de la versión filtrada por codificación de predicción lineal, LPC, de una porción de la señal de audio ponderada, y que comprende estimar la primera SNR segmentaria calculando una SNR estimada asociada con cada una de una pluralidad de subporciones de la porción de la señal de audio ponderada, basándose en una energía de las correspondientes subporciones de la señal de audio ponderada y la distorsión de cuantificador estimada y calculando un promedio de las SNR asociadas con las subporciones de la porción de la señal de audio ponderada para obtener la SNR segmentaria estimada para la porción de la señal de audio ponderada.
18. Método de acuerdo con una de las reivindicaciones 13 a 17, que comprende determinar una distorsión de libro de códigos adaptativos estimada (116) que introduciría un libro de códigos adaptativos usado en el segundo algoritmo de codificación cuando se usa el libro de códigos adaptativos para codificar la porción de la señal de audio, y estimar la segunda medida de calidad basándose en una energía de una porción de una versión ponderada de la señal de audio y la distorsión de libro de códigos adaptativos estimada, y que comprende para cada una de una pluralidad de subporciones de la porción de la señal de audio, aproximar (112) el libro de códigos adaptativos basándose en una versión de la subporción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, estimar (114) una ganancia de libro de códigos adaptativos de manera que se minimice un error entre la subporción de la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y determinar (116) la distorsión de libro de códigos adaptativos estimada basándose en la energía de un error entre la subporción de la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante la ganancia de libro de códigos adaptativos.
19. Método de acuerdo con la reivindicación 18, que comprende reducir (118) la distorsión de libro de códigos adaptativos estimada determinada para la subporción de la porción de la señal de audio, en un factor constante.
20. Método de acuerdo con la reivindicación 18 o 19, en el que la segunda medida de calidad es una SNR segmentaria de la porción de la señal de audio ponderada, y que comprende estimar la SNR segmentaria calculando una SNR estimada asociada con cada subporción basándose en la energía de la correspondiente subporción de la señal de audio ponderada, y la distorsión de libro de códigos adaptativos estimada y calculando un promedio de las SNR asociadas con las subporciones para obtener la SNR segmentaria estimada para la porción de la señal de audio ponderada.
21. Método de acuerdo con la reivindicación 18, que comprende aproximar el libro de códigos adaptativos basándose en una versión de la porción de la señal de audio ponderada desplazada al pasado por un retardo de paso determinado en una etapa de pre-procesamiento, estimar una ganancia de libro de códigos adaptativos de manera que se minimice un error entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado, y determinar la distorsión de libro de códigos adaptativos estimada basándose en la energía de un error entre la porción de la señal de audio ponderada y el libro de códigos adaptativos aproximado cambiado de escala mediante la ganancia de libro de códigos adaptativos.
22. Método de acuerdo con una de las reivindicaciones 13 a 21, que comprende utilizar histéresis para comparar las medidas de calidad estimadas.
24. Programa de computadora que tiene un código de programa para realizar, cuando se ejecuta en una computadora, el método de una de las reivindicaciones 13 a 22.

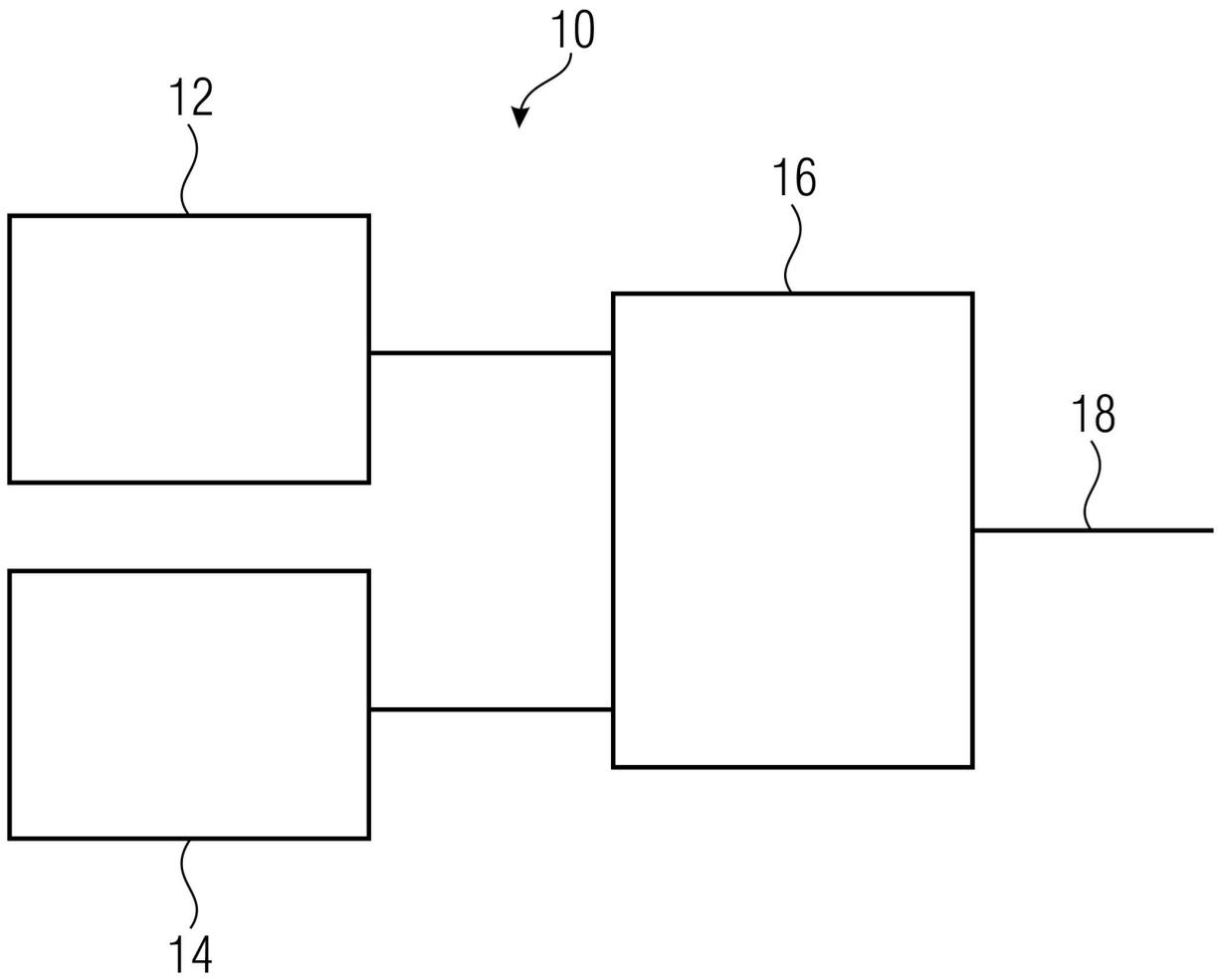


FIG 1

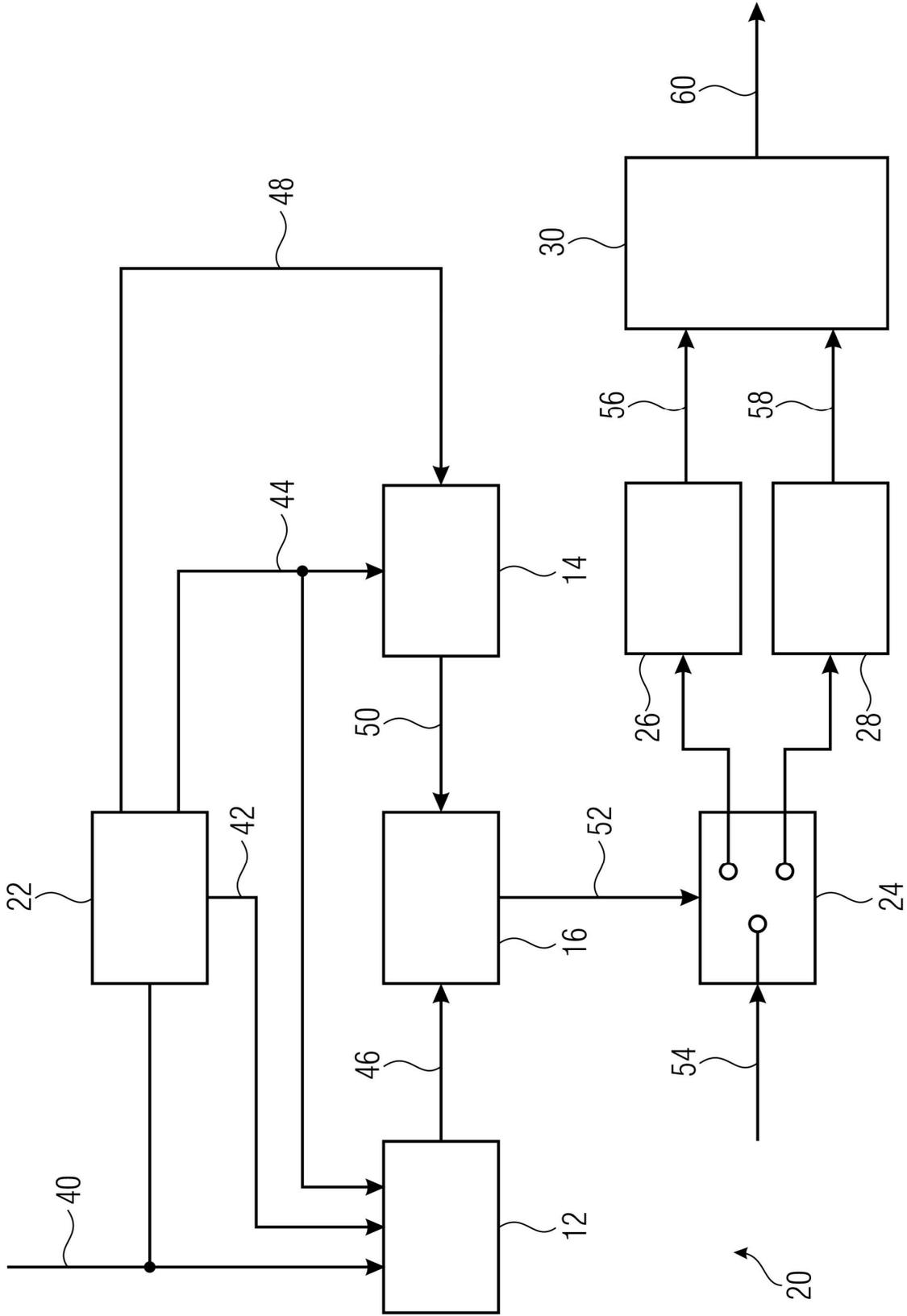


FIG 2

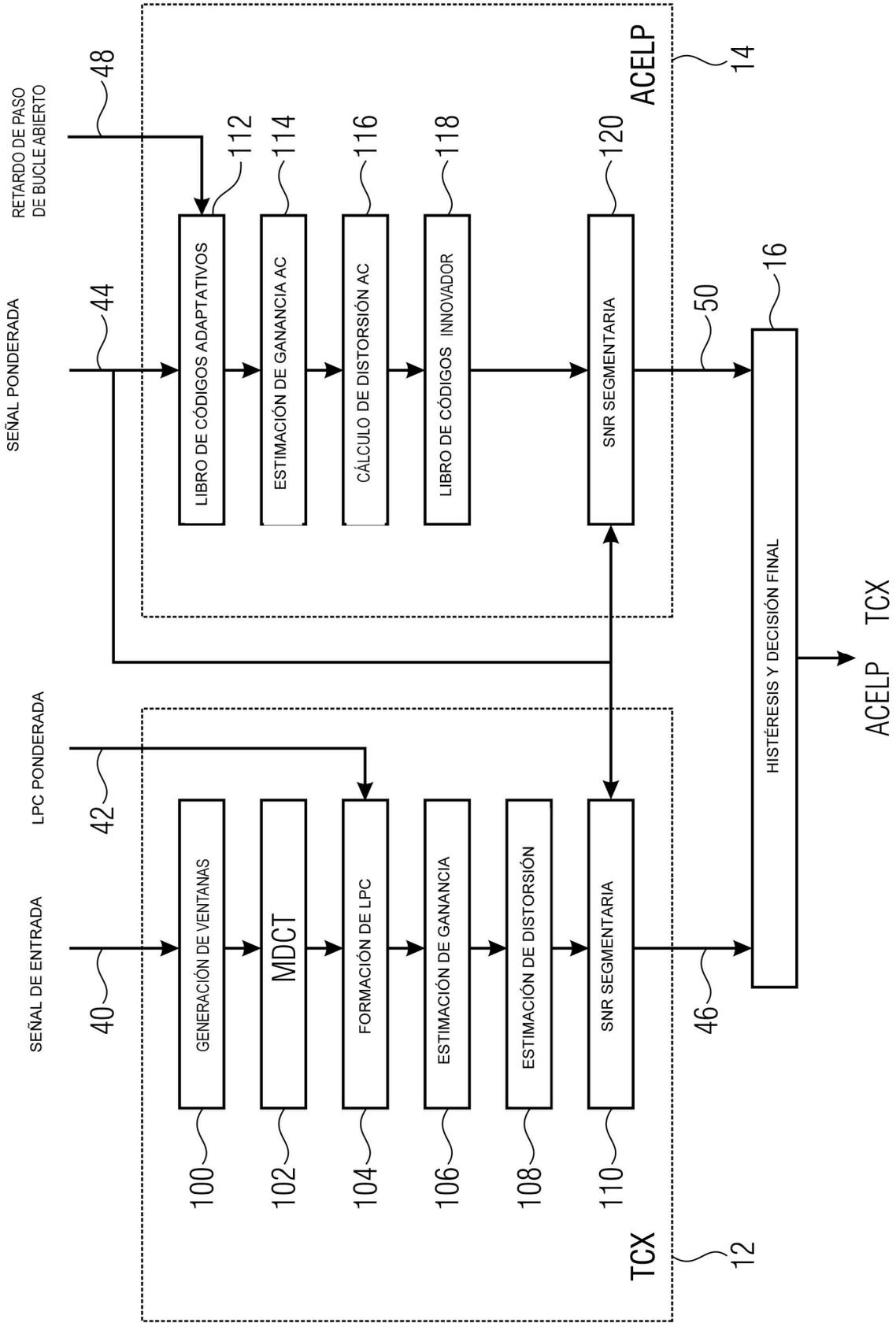


FIG 3

$$\text{SNR} = 10 \log_{10} \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2}$$

FIG 4A

$$\text{SNR}_{\text{seg}} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \sum_{i=Nm}^{Nm+N-1} \left(\frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2} \right)$$

FIG 4B