



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 291 877**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **04733391 .9**

86 Fecha de presentación : **17.05.2004**

87 Número de publicación de la solicitud: **1747555**

87 Fecha de publicación de la solicitud: **31.01.2007**

54 Título: **Codificación de audio con diferentes modelos de codificación.**

45 Fecha de publicación de la mención BOPI:
01.03.2008

45 Fecha de la publicación del folleto de la patente:
01.03.2008

73 Titular/es: **Nokia Corporation**
Keilalahdentie 4
02150 Espoo, FI

72 Inventor/es: **Mäkinen, Jari;**
Lakaniemi, Ari y
Ojala, Pasi

74 Agente: **Curell Suñol, Marcelino**

ES 2 291 877 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

Codificación de audio con diferentes modelos de codificación.

5 Campo de la invención

La presente invención se refiere a un método para respaldar una codificación de una señal de audio, en el que hay disponibles por lo menos un primer modo de un codificador y un segundo modo de un codificador para codificar una sección específica de la señal de audio. Por lo menos el primer modo del codificador permite una codificación de una sección específica de la señal de audio basándose en por lo menos dos modelos de codificación diferentes. En el primer modo del codificador, se habilita una selección de un modelo de codificación respectivo para codificar una sección específica de una señal de audio mediante por lo menos una regla de selección la cual se basa en un análisis de características de la señal en una ventana de análisis la cual abarca por lo menos una sección de la señal de audio que precede a la sección específica. La invención se refiere también a un módulo correspondiente, a un dispositivo electrónico correspondiente, a un sistema correspondiente y a un producto de programa de software correspondiente.

Antecedentes de la invención

Se conoce la codificación de señales de audio para posibilitar una transmisión y/o un almacenamiento eficaces de señales de audio.

Una señal de audio puede ser una señal de voz u otro tipo de señal de audio, tal como música, y para tipos diferentes de señales de audio podrían resultar adecuados modelos de codificación diferentes.

Una de las técnicas usadas ampliamente para codificar señales de voz es la codificación de Predicción Lineal con Excitación por Código Algebraico (ACELP). La ACELP modela el sistema de producción de la voz humana, y resulta adecuado de forma muy satisfactoria para codificar la periodicidad de una señal de voz. Como consecuencia, se puede lograr una calidad alta de la voz con velocidades binarias muy bajas. Por ejemplo, la Banda Ancha Multivelocidad Adaptativa (AMR-WB) es un códec de voz que se basa en la tecnología ACELP. El AMR-WB se ha descrito por ejemplo en la especificación técnica 3GPP TS 26.190: "Speech Codec speech processing functions; AMR Wideband speech codec; Transcoding functions", V5.1.0 (2001-12). No obstante, los códecs de voz que se basan en el sistema de producción de voz humana se comportan normalmente de forma bastante deficiente para otros tipos de señales de audio, tales como música.

Una de las técnicas usadas ampliamente para codificar señales de audio que no sean voz es la codificación por transformada (TCX). La superioridad de la codificación por transformada para las señales de audio se basa en el enmascaramiento perceptivo y la codificación en el dominio de la frecuencia. La calidad de la señal de audio resultante se puede mejorar adicionalmente seleccionando una longitud de trama de codificación adecuada para la codificación por transformada. No obstante, aunque las técnicas de codificación por transformada dan como resultado una calidad elevada para señales de audio que no sean voz, su rendimiento no es satisfactorio para señales de voz periódicas. Por esta razón, la calidad de la voz codificada por transformadas es normalmente bastante baja, especialmente con longitudes de tramas TCX elevadas.

El códec AMR-WB ampliado (AMR-WB+) codifica una señal de audio estereofónica como una señal monofónica con una alta velocidad binaria y proporciona cierta información colateral para una ampliación estereofónica. El códec AMR-WB+ utiliza tanto la codificación ACELP como modelos TCX para codificar la señal monofónica esencial en una banda de frecuencias de entre 0 Hz y 6.400 Hz. Para el modelo TCX, se utiliza una longitud de la trama de codificación de 20 ms, 40 ms ó 80 ms.

Como un modelo ACELP puede deteriorar la calidad de audio y la codificación por transformada normalmente se comporta de forma deficiente para la voz, especialmente cuando se utilizan tramas de codificación largas, debe seleccionarse el modelo de codificación respectivamente mejor dependiendo de las propiedades de la señal que se va a codificar. De hecho, la selección del modelo de codificación que se va a utilizar se puede llevar a cabo de varias maneras.

En sistemas que requieren técnicas de complejidad reducida, tales como los servicios multimedia móviles (MMS), normalmente se aprovechan algoritmos de clasificación de música/voz para seleccionar el modelo de codificación óptimo. Estos algoritmos clasifican la señal fuente completa bien como música o bien como voz basándose en un análisis de las propiedades de la energía y de la frecuencia de la señal de audio.

Si una señal de audio consta solamente de voz o solamente de música, el uso del mismo modelo de codificación para toda la señal basándose en dicha clasificación de música/voz resultará satisfactorio. No obstante, en muchos otros casos, la señal de audio que se va a codificar es un tipo combinado de señal de audio. Por ejemplo, la voz puede estar presente al mismo tiempo que la música y/o se puede alternar temporalmente con la música en la señal de audio.

En estos casos, una clasificación de las señales fuente completas en una categoría de música o voz es un planteamiento demasiado limitado. En este caso, la calidad de audio global únicamente se puede maximizar conmutando temporalmente entre los modelos de codificación cuando se codifica la señal de audio. Es decir, el modelo ACELP se

ES 2 291 877 T3

usa parcialmente también para codificar una señal fuente clasificada como señal de audio que no sea voz, mientras que el modelo TCX se usa parcialmente también para una señal fuente clasificada como señal de voz.

5 El códec AMR-WB ampliado (AMR-WB+) está diseñado también para codificar dichos tipos combinados de señales de audio con modelos de codificación combinados basándose en cada trama individual.

La selección de los modelos de codificación en el AMR-WB+ se puede llevar a cabo de varias maneras.

10 En el planteamiento más complejo, la señal en primer lugar se codifica con todas las combinaciones posibles de modelos ACELP y TCX. A continuación, la señal se sintetiza nuevamente para cada combinación. Seguidamente, se selecciona la excitación más satisfactoria basándose en la calidad de las señales de voz sintetizadas. La calidad de la voz sintetizada resultante con una combinación específica se puede medir, por ejemplo, determinando su relación/ruido (SNR). Este tipo de planteamiento de análisis-por-síntesis proporcionará resultados satisfactorios. No obstante, en algunas aplicaciones, el mismo no es viable debido a su complejidad extremadamente elevada. Dichas aplicaciones incluyen, por ejemplo, aplicaciones móviles. La complejidad es en gran medida resultado de la codificación ACELP, que es la parte más compleja de un codificador.

20 En sistemas, por ejemplo, como el MMS, el planteamiento de análisis-por-síntesis de bucle cerrado completo es demasiado complejo para llevarlo a la práctica. Por esta razón, en un codificador MMS, se utiliza un método de bucle abierto de complejidad reducida para determinar si se selecciona un modelo de codificación ACELP ó un modelo TCX para codificar una trama específica.

25 El AMR-WB+ ofrece dos planteamientos diferentes de bucle abierto de complejidad reducida para seleccionar el modelo de codificación respectivo para cada trama. Ambos planteamientos de bucle abierto evalúan las características de la señal fuente y los parámetros de codificación para seleccionar un modelo de codificación respectivo.

30 En el primer planteamiento de bucle abierto, en primer lugar una señal de audio se divide dentro de cada trama en varias bandas de frecuencia, y se analiza la relación entre la energía de las bandas de frecuencia inferiores y la energía de las bandas de frecuencia superiores, así como las variaciones del nivel de energía en dichas bandas. A continuación, el contenido de audio de cada trama de la señal de audio se clasifica como contenido de tipo música o contenido de tipo voz basándose en ambas mediciones realizadas o en diferentes combinaciones de estas mediciones usando diferentes ventanas de análisis y valores de umbral de decisión.

35 En el segundo planteamiento de bucle abierto, al cual se hace referencia también como perfeccionamiento de clasificación de los modelos, la selección del modelo de codificación se basa en una evaluación de la periodicidad y de las propiedades estacionarias del contenido de audio en una trama respectiva de la señal de audio. Las periodicidades estacionarias se evalúan más específicamente determinando la correlación, parámetros de Predicción a Largo Plazo (LTP) y mediciones de distancias espectrales.

40 El códec AMR-WB+ permite adicionalmente conmutar, durante la codificación de un flujo continuo de audio, entre modos AMR-WB, los cuales utilizan exclusivamente un modelo de codificación ACELP, y modos ampliados, los cuales utilizan bien un modelo de codificación ACELP ó bien un modelo TCX, siempre que no varíe la frecuencia de muestreo. La frecuencia de muestreo puede ser, por ejemplo, 16 kHz.

45 Los modos ampliados dan salida a una velocidad binaria mayor que los modos AMR-WB. De este modo, una conmutación de un modo ampliado a un modo AMR-WB puede resultar ventajosa cuando las condiciones de transmisión en la red que conecta el extremo de codificación y el extremo de decodificación requieran cambiar de un modo de velocidad binaria superior a un modo de velocidad binaria inferior para reducir la congestión en la red. También podría ser necesario un cambio de un modo de velocidad binaria superior a un modo de velocidad binaria inferior para incorporar nuevos receptores de la gama baja en un Servicio Móvil de Difusión General/Multidifusión (MBMS).

55 Por otro lado, una conmutación de un modo AMR-WB a un modo ampliado puede resultar ventajosa cuando un cambio en las condiciones de transmisión de la red permita un cambio de un modo de velocidad binaria inferior a un modo de velocidad binaria superior. El uso de un modo de velocidad binaria superior posibilita una mejor calidad de audio.

60 Como el códec central usa la misma frecuencia de muestreo de 6,4kHz para los modos AMR-WB y los modos ampliados AMR-WB+ y utiliza técnicas de codificación por lo menos parcialmente similares, un cambio de un modo ampliado a un modo AMR-WB, o viceversa, en esta banda de frecuencias, se puede gestionar uniformemente. No obstante, como el proceso de codificación de la banda central es ligeramente diferente para un modo AMR-WB y un modo ampliado, hay que ocuparse de que todas las variables de estado y las memorias intermedias requeridas se almacenen y copian de un algoritmo a otro cuando se conmuta entre los modos.

65 Además, debe tenerse en cuenta que una selección del modelo de codificación es necesaria únicamente en los modos ampliados. En los planteamientos habilitados de clasificación de bucle abierto, se aprovechan ventanas de análisis y memorias intermedias de datos relativamente largas. La selección del modelo de codificación aprovecha el análisis estadístico con ventanas de análisis que presentan una longitud de hasta 320 ms, lo cual se corresponde con 16 tramas de una señal de audio de 20 ms. Como en el modo AMR-WB no hay por qué almacenar en memoria

intermedia una información correspondiente, la misma no se puede copiar de forma sencilla a los algoritmos del modo ampliado. De este modo, después de una conmutación del AMR-WB al AMR-WB+, las memorias intermedias de datos de algoritmos de clasificación, por ejemplo las correspondientes usadas para un análisis estadístico, no tienen ninguna información válida o las mismas se reinician. De este modo, durante los primeros 320 ms después de una conmutación, puede que el algoritmo de selección del modelo de codificación no esté completamente adaptado o actualizado para la señal de audio en curso. Una selección que se base en unos datos no válidos de las memorias intermedias da como resultado una decisión distorsionada del modelo de codificación. Por ejemplo, a un modelo de codificación ACELP se le puede asignar un valor importante en la selección, incluso aunque la señal de audio requiera una codificación basada en un modelo TCX para mantener la calidad del audio.

De este modo, la selección del modelo de codificación no es óptima ya que la selección del modelo de codificación de complejidad reducida se comporta de forma no satisfactoria después de una conmutación de un modo AMR-WB a un modo ampliado.

El documento US-A-6640209 da a conocer un codificador multimodo en el que se ejecuta una conmutación del modo después de que la señal de entrada cumpla unos ciertos criterios sobre un número predeterminado de tramas.

Sumario de la invención

Uno de los objetivos de la invención es mejorar la selección de un modelo de codificación después de una conmutación de un primer modo de codificación a un segundo modo de codificación.

Se propone un método para respaldar una codificación de una señal de audio, en el que hay disponibles por lo menos un primer modo de un codificador y un segundo modo de un codificador para codificar una sección específica de la señal de audio. Además, por lo menos el primer modo del codificador permite una codificación de una sección específica de la señal de audio basándose en por lo menos dos modelos de codificación diferentes. En el primer modo del codificador, se habilita una selección de un modelo de codificación respectivo para codificar una sección específica de una señal de audio mediante por lo menos una regla de selección la cual se basa en características de la señal que se han determinado por lo menos parcialmente a partir de una ventana de análisis que abarca por lo menos una sección de la señal de audio que precede a la sección específica. Se propone que el método comprenda, después de una conmutación del segundo modo del codificador al primer modo del codificador, la activación de dicha por lo menos una regla de selección en respuesta a la recepción de por lo menos tantas secciones de la señal de audio como las abarcadas por la ventana de análisis.

El primer modo del codificador y el segundo modo del codificador pueden ser, por ejemplo, aunque no de forma exclusiva, un modo ampliado y un modo AMR-WB de un códec AMR-WB+, respectivamente. En ese caso, los modelos de codificación disponibles para el primer modo del codificador pueden ser, por ejemplo, un modelo de codificación ACELP y un modelo TCX.

Por otra parte, se propone un módulo para respaldar una codificación de una señal de audio. El módulo comprende una parte del primer modo del codificador adaptada para codificar una sección específica de una señal de audio en un primer modo del codificador y una parte del segundo modo del codificador adaptada para codificar una sección respectiva de una señal de audio en un segundo modo del codificador. El módulo comprende además medios de conmutación para conmutar entre la parte del primer modo del codificador y la parte del segundo modo del codificador. La parte del modo del codificador incluye una parte de codificación la cual está adaptada para codificar una sección respectiva de la señal de audio basándose en por lo menos dos modelos de codificación diferentes. La parte del primer modo del codificador comprende además una parte de selección adaptada para aplicar por lo menos una regla de selección con vistas a seleccionar un modelo de codificación respectivo, el cual será usado por la parte de codificación para codificar una sección específica de una señal de audio. Dicha por lo menos una regla de selección se basa en características de la señal que se han determinado por lo menos parcialmente a partir de una ventana de análisis que abarca por lo menos una sección de una señal de audio que precede a la sección específica. La parte de selección está adaptada para activar dicha por lo menos una regla de selección después de una conmutación por parte de los medios de conmutación desde la parte del segundo modo del codificador a la parte del primer modo del codificador en respuesta a la recepción de por lo menos tantas secciones de la señal de audio como las abarcadas por la ventana de análisis.

Este módulo puede ser por ejemplo un codificador o una parte de un codificador.

Por otra parte, se propone un dispositivo electrónico el cual comprende dicho módulo.

Por otra parte, se propone un sistema de codificación de audio el cual comprende dicho módulo y adicionalmente un decodificador para decodificar señales de audio que han sido decodificadas por dicho módulo.

Finalmente, se propone un producto de programa de software en el cual se almacena un código de software para respaldar una codificación de una señal de audio. Hay disponibles por lo menos un primer modo del codificador y un segundo modo del codificador para codificar una sección respectiva de la señal de audio. Por lo menos el primer modo del codificador permite una codificación de una sección respectiva de la señal de audio basándose en por lo menos dos modelos de codificación diferentes. En el primer modo del codificador se habilita una sección de un modelo de

codificación respectivo para codificar una sección específica de una señal de audio mediante por lo menos una regla de selección que se basa en características de la señal que han sido determinadas a partir de una ventana de análisis que abarca por lo menos una sección de la señal de audio que precede a la sección específica. Cuando se ejecuta en un componente de procesado de un codificador, el código de software activa dicha por lo menos una regla de selección después de una conmutación desde el segundo modo del codificador al primer modo del codificador en respuesta a la recepción de por lo menos tantas secciones de la señal de audio como las abarcadas por la ventana de análisis.

La invención tiene su origen en la consideración de que se pueden evitar los problemas con el contenido no válido de las memorias intermedias que se usan como base para una selección de un modelo de codificación si dicha selección se activa únicamente después de que el contenido de las memorias intermedias se haya actualizado por lo menos hasta un nivel requerido por el tipo respectivo de selección. Por esta razón se propone que cuando una regla de selección haga uso de características de la señal que se han determinado usando una ventana de análisis sobre una pluralidad de secciones de la señal de audio, la regla de selección se aplica únicamente cuando se han recibido todas las secciones requeridas por la ventana de análisis. Debe entenderse que la activación puede ser parte de la propia regla de selección.

Una de las ventajas de la invención es que permite una selección mejorada del modelo de codificación después de una conmutación del modo del codificador. Más específicamente, permite evitar una clasificación errónea de las secciones de una señal de audio, y por lo tanto evita la selección de un modelo de codificación inadecuado.

Para el periodo de tiempo después de una conmutación en el cual no se hayan activado ciertas reglas de selección, se proporciona de forma ventajosa una regla de selección adicional la cual no hace uso de información sobre secciones de la señal de audio que preceden a la sección en curso. Esta otra regla se puede aplicar inmediatamente después de una conmutación y por lo menos hasta que se hayan activado otras reglas de selección.

Dicha por lo menos una regla de selección que se basa en características de la señal que han sido determinadas en una ventana de análisis puede comprender una regla de selección individual o una pluralidad de reglas de selección. En este último caso, las ventanas de análisis asociadas pueden tener longitudes diferentes. Como consecuencia, la pluralidad de reglas de selección se pueden activar una después de otra.

La sección de una señal de audio puede ser, en particular, una trama de una señal de audio, por ejemplo, una trama de 20 ms de una señal de audio.

Las características de la señal que son evaluadas por dicha por lo menos una regla de selección se pueden basar totalmente o solo de forma parcial en una ventana de análisis. Debe entenderse que también las características de la señal utilizadas por una regla de selección individual pueden basarse en diferentes ventanas de análisis.

Breve descripción de las figuras

A partir de la siguiente descripción detallada considerada conjuntamente con los dibujos adjuntos se pondrán de manifiesto otros objetivos y características de la presente invención.

La Fig. 1 es un diagrama esquemático de un sistema de codificación de audio según una de las formas de realización de la invención; y

la Fig. 2 es un diagrama de flujo que ilustra una de las formas de realización del método según la invención implementado en el sistema de la Figura 1.

Descripción detallada de la invención

La Figura 1 es un diagrama esquemático de un sistema de codificación de audio según una de las formas de realización de la invención, la cual permite una activación flexible de algoritmos de selección usados para seleccionar un modelo de codificación óptimo.

El sistema comprende un primer dispositivo 1 que incluye un codificador AMR-WB+ 2 y un segundo dispositivo 21 que incluye un decodificador AMR-WB+ 22. El primer dispositivo 1 puede ser, por ejemplo, un servidor MMS, mientras que el segundo dispositivo 21 puede ser, por ejemplo, un teléfono móvil o algún otro dispositivo móvil.

El codificador AMR-WB+ 2 comprende una parte de codificación AMR-WB 4 la cual está adaptada para realizar una codificación ACELP pura, y una parte de codificación ampliada 5, la cual está adaptada para realizar una codificación basada bien en un modelo de codificación ACELP o bien en un modelo TCX. De este modo, la parte de codificación ampliada 5 constituye la parte del primer modo del codificador y la parte de codificación AMR-WB 4 constituye la parte del segundo modo del codificador de la invención.

El codificador AMR-WB+ 2 comprende además un conmutador 6 para reenviar tramas de la señal de audio bien hacia la parte de codificación AMR-WB 4 ó bien hacia la parte de codificación ampliada 5.

La parte de codificación ampliada 5 comprende una parte de determinación de características de la señal 11 y un contador 12. El terminal del conmutador 6 que está asociado a la parte de codificación ampliada 5 está unido a una

ES 2 291 877 T3

entrada de ambas partes 11, 12. La salida de la parte de determinación de características de la señal 11 y la salida del contador 12 están unidas dentro de la parte de codificación ampliada 5 a través de una primera parte de selección 13, una segunda parte de selección 14, una tercera parte de selección 15, una parte de verificación 16, una parte de perfeccionamiento 17 y una parte de selección final 18 hacia una parte de codificación ACELP/TCX 19.

Debe entenderse que las partes presentadas 11 a 19 están diseñadas para codificar una señal de audio monofónica, la cual puede haber sido generada a partir de una señal de audio estereofónica. Se puede generar una información estereofónica adicional en partes ampliadas estereofónicas adicionales no mostradas. Por otra parte, debe observarse que el codificador 2 comprende otras partes no mostradas. Debe entenderse también que las partes presentadas 12 a 19 no tienen porque ser partes independientes, sino que también pueden estar interconectadas entre ellas o con otras partes.

La parte de codificación AMR-WB 4, la parte de codificación ampliada 5 y el conmutador 6 se pueden realizar en particular mediante un software SW ejecutado en un componente de procesado 3 del codificador 2, el cual se indica mediante líneas de trazos.

A continuación se describirá más detalladamente el procesado de la parte de codificación ampliada 5 haciendo referencia al diagrama de flujo de la Figura 2.

El codificador 2 recibe una señal de audio, la cual se ha proporcionado al primer dispositivo 1. En primer lugar, el conmutador 6 suministra la señal de audio a la parte de codificación AMR-WB 4 para lograr una velocidad binaria baja de salida, debido por ejemplo a que no se dispone de la suficiente capacidad en la red que conecta el primer dispositivo 1 y el segundo dispositivo 21. No obstante, posteriormente las condiciones de la red cambian y permiten una velocidad binaria mayor. Por esta razón, a continuación la señal de audio es reenviada por el conmutador 6 hacia la parte de codificación ampliada 5.

En el caso de un conmutador de este tipo, un valor *StatClassCount* del contador 12 se reinicializa a 15 cuando se recibe la primera trama de la señal de audio. A continuación, el contador 12 decrementa en uno su valor *StatClassCount*, cada vez que se introduce otra trama de la señal de audio en la parte de codificación ampliada 5.

Por otro lado, la parte de determinación de características de la señal 11 determina para cada trama introducida de la señal de audio varias características de la señal relacionadas con la energía por medio de bancos de filtros del Detector de Actividad Vocal (VAD) AMR-WB.

Para cada trama introducida de la señal de audio de 20 ms, los bancos de filtros producen la energía de la señal $E(n)$ en cada una de entre doce bandas de frecuencias no uniformes que abarcan una gama de frecuencias de entre 0 Hz y 6.400 Hz. A continuación, el nivel de energía $E(n)$ de cada banda de frecuencias n se divide por la anchura de esta banda de frecuencias en Hz, con vistas a producir un nivel de energía normalizado $E_N(n)$ para cada banda de frecuencias.

Seguidamente, se calcula la desviación estándar respectiva de los niveles de energía normalizados $E_N(n)$ para cada una de las doce bandas de frecuencias usando, por un lado, una ventana corta $std_{corta}(n)$ y, por otro lado, una ventana larga $std_{larga}(n)$. La ventana corta tiene una longitud de cuatro tramas de la señal de audio, y la ventana larga tiene una longitud de dieciséis tramas de la señal de audio. Es decir, para cada banda de frecuencias, para obtener los dos valores de la desviación estándar se usan el nivel de energía de la trama en curso y el nivel de energía de las tramas anteriores 4 y 16, respectivamente. Los niveles de energía normalizados de las tramas anteriores se recuperan a partir de memorias intermedias, en las cuales también se almacenan para un uso posterior los niveles de energía normalizados de la trama en curso de la señal de audio.

No obstante, las desviaciones estándar se determinan únicamente en el caso de que un indicador de actividad vocal VAD indique la existencia de voz activa para la trama en curso. Esta opción conseguirá que el algoritmo reaccione más rápido especialmente después de pausas prolongadas de la voz.

A continuación, las desviaciones estándar determinadas se promedian sobre las doce bandas de frecuencia para la ventana tanto larga como corta con vistas a crear dos valores de desviación estándar media $stda_{corta}$ y $stda_{larga}$ como una primera y una segunda características de la señal para la trama en curso de la señal de audio.

Por otra parte, para la trama en curso de la señal de audio se calcula una relación entre la energía de las bandas de frecuencia inferiores y la energía de las bandas de frecuencia superiores. Con este fin, la parte de determinación de características de la señal 11 suma las energías $E(n)$ de las bandas de frecuencias inferiores $n = 1$ a 7 para obtener un nivel de energía LevL. El nivel de energía LevL se normaliza dividiéndolo por la anchura total de estas bandas de frecuencias inferiores en Hz. Por otra parte, la parte de determinación de características de la señal 11 suma las energías $E(n)$ de las bandas de frecuencias superiores $n = 8$ a 11 para obtener un nivel de energía LevH. El nivel de energía LevH se normaliza igualmente dividiéndolo por la anchura total de las bandas de frecuencias superiores en Hz. En estos cálculos no se usa la banda de frecuencias más baja 0, ya que la misma habitualmente contiene tanta energía que distorsionaría los cálculos y provocaría que las contribuciones de las otras bandas de frecuencia resultaran demasiado reducidas. Seguidamente, la parte de determinación de características de la señal 11 define la relación $LPH = LevL / LevH$. Adicionalmente, se calcula una media móvil LPHa usando los valores LPH que se han determinado para la trama en curso de la señal de audio y para las tres tramas anteriores de la señal de audio.

ES 2 291 877 T3

5 A continuación se calcula un valor final *LPHaF* de la relación energía para la trama en curso sumando el valor *LPHa* en curso y los anteriores siete valores *LPHa*. En esta suma, los últimos valores de *LPHa* se ponderan con un peso ligeramente superior a los valores antiguos de *LPHa*. Los anteriores siete valores de *LPHa* se recuperan igualmente de las memorias intermedias, en las cuales también se almacena para un uso posterior el valor de *LPHa* para la trama en curso. El valor *LPHaF* constituye la tercera característica de la señal.

10 La parte de determinación de características de la señal 11 calcula adicionalmente un nivel medio de energía de los bancos de filtros *AVL* para la trama en curso de la señal de audio. Para calcular el valor *AVL*, de la energía $E(n)$ se resta un nivel estimado del ruido de fondo en cada una de las doce bandas de frecuencias. A continuación, los resultados se multiplican por la frecuencia más alta en Hz de la banda de frecuencias correspondiente y los mismos se suman. La multiplicación permite un equilibrado de la influencia de las bandas de frecuencia altas, que contienen una energía relativamente menor que las bandas de frecuencia inferiores. El valor *AVL* constituye una cuarta característica de la señal.

15 Finalmente, la parte de determinación de características de la señal 11 calcula para la trama en curso la energía total $TotE_0$ a partir de todos los bancos de filtros, reducida en una estimación del ruido de fondo para cada banco de filtros. La energía total $TotE_0$ se almacena también en una memoria intermedia. El valor $TotE_0$ constituye una quinta característica de la señal.

20 A continuación, las características determinadas de la señal y el valor del contador *StatClassCount* se suministran a la primera parte de selección 13, la cual aplica un algoritmo según el siguiente pseudo-código para seleccionar el mejor modelo de codificación para la trama en curso:

```
25     si (StatClassCount == 0)
           FIJAR MODO_TCX
30     si (stdalarga < 0,4)
           FIJAR MODO_TCX
35     si no, si (LPHaF > 280)
           FIJAR MODO_TCX
40     si no, si (stdalarga >= 0,4)
           si ((5+(1/(stdalarga -0,4))) > LPHaF)
45             FIJAR MODO_TCX
           si no, si ((~90* Stdalarga +120) < LPHaF)
50             FIJAR MODO_ACELP
           si no
55             FIJAR MODO_INCIERTO
           si no
60             headMode = MODO_INCIERTO
```

65 Puede observarse que este algoritmo aprovecha una característica de la señal $stda_{larga}$, la cual se basa en información sobre dieciséis tramas anteriores de la señal de audio. Por esta razón, en primer lugar se comprueba si ya se han recibido por lo menos diecisiete tramas después de la conmutación desde AMR-WB. Esto se cumple en cuanto el contador 12 presenta un valor *StatClassCount* de cero. En cualquier otro caso, a la trama en curso se le asocia inmediatamente un modo incierto. Esta opción garantiza que el resultado no sea falseado por un contenido no válido de las memorias intermedias que dé como resultado valores incorrectos para las características de la señal $stda_{larga}$ y *LPHaF*.

ES 2 291 877 T3

La información sobre las características de la señal y la selección del modelo de codificación materializada hasta el momento es reenviar a continuación por la primera parte de selección 13 hacia la segunda parte de selección 14, la cual aplica un algoritmo según el siguiente pseudo-código para seleccionar el mejor modelo de codificación para la trama en curso:

```
5
    si (MODO_ACELP o MODO_INCIERTO) y (AVL > 2000)

10
    FIJAR MODO_TCX

    si (StatClassCount < 5)

15
    si (MODO_INCIERTO)

    si (stdacorta < 0,2)

20
    FIJAR MODO_TCX

    si no, si (stdacorta >= 0,2)

25
    si ((2,5+(1/(stdacorta -0,2))) > LPHaF)

    FIJAR MODO_TCX

30
    si no, si ((-90* stdacorta+140) < LPHaF)

    FIJAR MODO_ACELP

35
    si no

    FIJAR MODO_INCIERTO
```

40 Puede observarse que la segunda parte de este algoritmo aprovecha una característica de la señal $stda_{corta}$, la cual se basa en información sobre cuatro tramas anteriores de la señal de audio, y además una característica de la señal LPHaF, la cual se basa en información sobre diez tramas anteriores de la señal de audio. Por esta razón, para esta parte del algoritmo, en primer lugar se comprueba si ya se han recibido por lo menos once tramas después de la conmutación desde AMR-WB. Esta opción se cumple en cuanto el contador presenta un valor *StatClassCount* de "4". Esto garantiza que el resultado no queda falseado por un contenido no válido de la memoria intermedia que dé como resultado valores incorrectos para características de la señal LPHaF y $stda_{corta}$. En general, este algoritmo permite una selección de un modelo de codificación ya para la trama undécima a decimosexta, y adicionalmente incluso para las primeras diez tramas en el caso de que el nivel de energía medio AVL supere un valor predeterminado. En la Figura 2 no se indica esta parte del algoritmo. El algoritmo se aplica igualmente para tramas sucesivas a la trama decimosexta con vistas a perfeccionar la primera selección realizada por la primera parte de selección 13.

55 La información sobre las características de la señal y la selección del modelo de codificación materializada hasta el momento es reenviada a continuación por la segunda parte de selección 14 hacia la tercera parte de selección 15, la cual aplica un algoritmo según el siguiente pseudo-código para seleccionar el mejor modelo de codificación para la trama en curso, en caso de que el modo correspondiente a esta trama siga siendo incierto:

```
    si (MODO_INCIERTO)

60
    si (StatClassCount < 15)

    si ((TotE0 / TotE-1) >25)

65
    FIJAR MODO_ACELP
```

ES 2 291 877 T3

Puede observarse que este pseudo-código aprovecha la relación entre la energía total $TotE_0$ de la trama en curso de la señal de audio y la energía total $TotE_{-1}$ de la trama anterior de la señal de audio. Por esta razón, en primer lugar se comprueba si ya se han recibido por lo menos dos tramas después de la conmutación desde AMR-WB. Esta opción se cumple en cuanto el contador presente un valor *StatClassCount* de "14".

Debe indicarse que los valores de umbral utilizados del contador son únicamente ejemplos y se podrían seleccionar de muchas maneras diferentes. En el algoritmo implementado en la segunda parte de selección 14, por ejemplo, en lugar de la característica de la señal LPHaF se podría evaluar la característica de la señal LPH. En este caso, bastaría con comprobar si ya se han recibido por lo menos cinco tramas, en correspondencia con *StatClassCount* < 12.

A continuación, la información sobre las características de la señal y la selección del modelo de codificación materializada hasta el momento es reenviada por la tercera parte de selección 15 hacia la parte de verificación 16, la cual aplica un algoritmo según el siguiente pseudo-código:

```
si (MODO_TCX || MODO_INCIERTO))
```

```
    si (AVL > 2000 y TotE0 < 60)
```

```
        FIJAR MODO_ACELP
```

Este algoritmo permite seleccionar posiblemente el mejor modelo de codificación para la trama en curso, en caso de que el modo correspondiente a esta trama siga siendo incierto, y verificar si un modo TCX ya seleccionado resulta adecuado.

Además, después del procesado en la parte de verificación 16, el modo asociado a la trama en curso de la señal de audio puede que siga siendo incierto.

A continuación, en un planteamiento rápido, para las restantes tramas de modo INCIERTO se selecciona simplemente un modelo de codificación predeterminado, es decir bien un modelo de codificación ACELP ó bien un modelo de codificación TCX.

En un planteamiento más sofisticado, ilustrado también en la Figura 2, se realiza en primer lugar algún otro análisis.

Con este fin, la información sobre la selección del modelo de codificación materializada hasta el momento es reenviada a continuación por la parte de verificación 16 hacia la parte de perfeccionamiento 17. La parte de perfeccionamiento 17 aplica un perfeccionamiento de la clasificación del modelo. Tal como se ha mencionado anteriormente, esta opción es una selección del modelo de codificación, la cual se basa en la periodicidad y las propiedades estacionarias de la señal de audio. La periodicidad se observa usando parámetros LTP. Las propiedades estacionarias se analizan usando una correlación normalizada y mediciones de las distancias espectrales.

El análisis realizado por las partes 13, 14, 15, 16 y 17 determina, basándose en las características de la señal de audio, si el contenido de una trama respectiva se puede considerar como voz u otro contenido de audio, tal como música, y selecciona un modelo de codificación correspondiente en el caso de que dicha clasificación sea posible. Las partes 13, 14, 15, 16 realizan un primer planteamiento de bucle abierto que evalúa características relacionadas con la energía, mientras que la parte 17 realiza un segundo planteamiento de bucle abierto que evalúa la periodicidad y las propiedades estacionarias de la señal de audio.

En el caso de que se hayan aplicado en vano dos planteamientos diferentes de bucle abierto para seleccionar un modelo TCX ó un modelo de codificación ACELP, en algunos casos el modelo de codificación óptimo resultará difícil de seleccionar por parte de otros algoritmos de bucle abierto existentes. Por esta razón, en la presente forma de realización, para las restantes selecciones de modo no claras se utiliza una clasificación sencilla basada en un recuento.

La parte de selección final 18 selecciona un modelo de codificación específico para las tramas de modo INCIERTO restantes basándose en una evaluación estadística de los modelos de codificación asociados a las tramas vecinas respectivas, en el caso de que se haya activado un indicador de actividad vocal banderaVAD para la trama de modo INCIERTO respectiva.

Para la evaluación estadística, se consideran una supertrama en curso, a la cual pertenece una trama de modo INCIERTO, y una supertrama anterior que precede a esta supertrama en curso. Una supertrama tiene una longitud de 80 ms y comprende cuatro tramas de audio consecutivas de 20 ms cada una de ellas. La parte de selección final 18, por medio de contadores, cuenta el número de tramas de la supertrama en curso y de la supertrama anterior para las cuales una de las partes de selección anteriores 12 a 17 ha seleccionado el modelo de codificación ACELP. Por otra parte, la parte de selección final 18 cuenta el número de tramas de la supertrama anterior para las cuales una de las partes de

ES 2 291 877 T3

selección anteriores 12 a 17 ha seleccionado un modelo TCX con una longitud de trama de codificación de 40 ms ó 80 ms, para las cuales por otra parte se ha activado el indicador de actividad vocal, y para las cuales adicionalmente la energía total supera un valor de umbral predeterminado. La energía total se puede calcular dividiendo la señal de audio en diferentes bandas de frecuencia, determinando el nivel de la señal por separado para todas las bandas de frecuencia, y sumando los niveles resultantes. El valor de umbral predeterminado para la energía total en una trama se puede fijar, por ejemplo, a 60.

La asignación de modelos de codificación se debe completar para una supertrama en curso completa, antes de que se pueda codificar la supertrama n en curso. De este modo, el recuento de tramas para las cuales se ha asignado un modelo de codificación ACELP no se limita a las tramas que preceden a una trama de modo INCIERTO. A no ser que la trama de modo INCIERTO sea la última trama de la supertrama en curso, se tienen en cuenta también los modelos de codificación seleccionados de las tramas que están por llegar.

El recuento de tramas se puede resumir, por ejemplo, mediante el siguiente pseudo-código:

si ((ModoPrev (i) == TCX80 ó ModoPrev (i) == TCX40) y

BanderaVad_{antigua} (i) == 1 y TotE_i > 60)

RecuentoTCX = RecuentoTCX + 1

si (ModoPrev (i) == MODO_ACELP)

RecuentoACELP = RecuentoACELP + 1

si (j != i)

si (Modo (i) == MODO_ACELP)

RecuentoACELP = RecuentoACELP + 1

En este pseudo-código, i indica el número de una trama en una supertrama respectiva, y presenta los valores 1, 2, 3, 4, mientras que j indica el número de la trama en curso en la supertrama en curso. *ModoPrev(i)* es el modo de la trama i-ésima de 20 ms en la supertrama anterior y *Modo(i)* es el modo de la trama i-ésima de 20 ms en la supertrama en curso. TCX80 representa un modelo TCX seleccionado que usa una trama de codificación de 80 ms y TCX40 representa un modelo TCX seleccionado que usa una trama de codificación de 40 ms. *BanderaVad_{antigua}(i)* representa el indicador de actividad vocal VAD para la trama i-ésima en la supertrama anterior. TotE_i es la energía total de la trama i-ésima. El valor del contador RecuentoTCX representa el número de tramas TCX largas seleccionadas en la supertrama anterior, y el valor del contador RecuentoACELP representa el número de tramas ACELP en la supertrama anterior y la supertrama en curso.

A continuación se realiza una evaluación estadística de la manera siguiente:

Si el número contado de tramas largas de modo TCX, con una longitud de trama de codificación de 40 ms ó 80 ms, en la supertrama anterior, es mayor que 3, para la trama de modo INCIERTO se selecciona igualmente un modelo TCX.

Alternativamente, si el número contado de tramas de modo ACELP en la supertrama en curso y anterior es mayor que 1, para la trama de modo INCIERTO se selecciona un modelo ACELP.

En la totalidad del resto de casos, para la trama de modo INCIERTO se selecciona un modelo TCX.

La selección del Modo(j) del modelo de codificación de la trama j-ésima se puede resumir, por ejemplo, mediante el siguiente pseudo-código:

si (*RecuentoTCX* > 3)

Modo(j) = *MODO_TCX*;

5

si no, si (*RecuentoACELP* > 1)

Modo(j) = *MODO_ACELP*

10

si no

Modo(j) = *MODO_TCX*

15

El planteamiento basado en un recuento se realiza únicamente si el valor del contador *StatClassCount* es menor que 12. Esto significa, que después de una conmutación desde AMR-WB a un modo ampliado, el planteamiento de clasificación basado en un recuento no se realiza en las primeras cuatro tramas, lo cual se corresponde con los primeros 4*20 ms.

20

Si el valor del contador *StatClassCount* es igual a o mayor que 12 y el modelo de codificación se ha clasificado todavía como modo INCIERTO, se selecciona el modelo TCX.

25

Si no se ha activado el indicador de actividad local banderaVAD, indicando por lo tanto la bandera un periodo de silencio, el modo seleccionado es TCX por defecto y no se debe realizar ninguno de los algoritmos de selección de modo.

30

De esta manera, las partes 13, 14 y 15 constituyen dicha por lo menos una parte de selección de la invención, mientras que las partes 16, 17 y 18, y parcialmente la parte 14, constituyen dicha por lo menos otra parte de selección de la invención.

35

En este caso, la parte de codificación ACELP/TCX 19 codifica todas las tramas de la señal de audio basándose en el modelo de codificación seleccionado respectivamente. El modelo TCX se basa, a título de ejemplo, en una transformada rápida de Fourier (FFT) que hace uso de la longitud de la trama de codificación seleccionada, y el modelo de codificación ACELP usa, a título de ejemplo, una LTP y parámetros de libro de código fijo para una excitación por coeficientes de predicción lineal (LPC).

40

A continuación, la parte de codificación 19 suministra las tramas codificadas, con vistas a una transmisión, al segundo dispositivo 21. En el segundo dispositivo 21, el decodificador 22 decodifica todas las tramas recibidas con el modelo de codificación ACELP o con el modelo de codificación TCX usando un modo AMR-WB o un modo ampliado, según se requiera. Las tramas decodificadas se suministran, por ejemplo con vistas a su presentación, a un usuario del segundo dispositivo 21.

45

En resumen, la forma de realización presentada permite una activación flexible de algoritmos de selección, en la cual los algoritmos de selección proporcionados se activan en el orden en el que se actualizan completamente las memorias intermedias de análisis que están relacionadas con las reglas de selección. Cuando se deshabiliten uno o más algoritmos de selección, la selección se realiza basándose en otros algoritmos de selección, los cuales no se basan en el contenido de esta memoria intermedia.

50

Debe indicarse que la forma de realización descrita constituye únicamente una de entra una variedad de posibles formas de realización de la invención.

55

60

65

REIVINDICACIONES

5 1. Método para respaldar una codificación de una señal de audio, en el que hay disponibles por lo menos un primer modo del codificador y un segundo modo del codificador para codificar una sección específica de dicha señal de audio, en el que dicho por lo menos un primer modo del codificador permite una codificación de una sección específica de dicha señal de audio basándose en por lo menos dos modelos de codificación diferentes, y en el que en dicho primer modo del codificador, se habilita una selección de un modelo de codificación respectivo para codificar dicha sección específica de una señal de audio mediante por lo menos una regla de selección la cual se basa en características de la señal, habiéndose determinado por lo menos parcialmente dichas características de la señal a partir de una ventana de análisis, abarcando dicha ventana de análisis por lo menos una sección de dicha señal de audio que precede a dicha sección específica, comprendiendo dicho método, después de una conmutación desde dicho segundo modo del codificador a dicho primer modo del codificador, la activación de dicha por lo menos una regla de selección en respuesta a la recepción de por lo menos tantas secciones de dicha señal de audio como las correspondientes abarcadas por dicha ventana de análisis.

20 2. Método según la reivindicación 1, en el que en dicho primer modo del codificador se habilita asimismo una selección de un modelo de codificación respectivo para codificar una sección específica de una señal de audio mediante por lo menos una regla de selección adicional que utiliza información sobre secciones de dicha señal de audio que preceden a dicha sección específica, aplicándose dicha por lo menos una regla de selección adicional por lo menos siempre que el número de secciones recibidas sea menor que el número de secciones abarcadas por una ventana de análisis, en el cual para dicha por lo menos una regla de selección se determinan características de la señal.

25 3. Método según la reivindicación 1 ó 2, en el que dicha por lo menos una regla de selección, la cual se basa en características de la señal que se han determinado a partir de una ventana de análisis, comprende una primera regla de selección, la cual se basa en características de la señal que se han determinado en una ventana de análisis más corta, y una segunda regla de selección, la cual se basa en características de la señal que se han determinado en una ventana de análisis más larga, en el que dicha primera regla de selección se activa en cuanto se han recibido suficientes secciones de dicha señal de audio para dicha ventana de análisis más corta, y en el que dicha segunda regla de selección se activa en cuanto se han recibido suficientes secciones de dicha señal de audio para dicha ventana de análisis más larga.

30 4. Método según la reivindicación 3, en el que una sección respectiva de dicha señal de audio se corresponde con una trama respectiva de la señal de audio que tiene una longitud de 20 ms, en el que dicha ventana más corta abarca una trama de la señal de audio para la cual se va a seleccionar un modelo de codificación y adicionalmente cuatro tramas anteriores de la señal de audio, y en el que dicha ventana más larga abarca una trama de la señal de audio para la cual se va a seleccionar un modelo de codificación y adicionalmente dieciséis tramas anteriores de la señal de audio.

35 5. Método según una de las reivindicaciones anteriores, en el que dichas características de la señal comprenden una desviación estándar de valores relacionados con la energía en una ventana de análisis respectiva.

40 6. Método según una de las reivindicaciones anteriores, en el que dicho primer modo del codificador es un modo ampliado de un códec ampliado de banda ancha multivelocidad adaptativa y permite una codificación basada en un modelo de codificación de predicción lineal con excitación por código algebraico y adicionalmente una codificación basada en un modelo de codificación por transformada, y en el que dicho segundo modo del codificador es un modo de banda ancha multivelocidad adaptativa de dicho códec ampliado de banda ancha multivelocidad adaptativa y permite una codificación basada en un modelo de codificación de predicción lineal con excitación por código algebraico.

45 7. Método según cualquiera de las reivindicaciones anteriores, en el que dicha sección es una trama o una subtrama de dicha señal de audio.

50 8. Módulo (2,3) para respaldar una codificación de una señal de audio, comprendiendo dicho módulo (2,3):

- una parte del primer modo del codificador (5) adaptada para codificar una sección respectiva de una señal de audio en un primer modo del codificador;
- 55 - una parte del segundo modo del codificador (4) adaptada para codificar una sección respectiva de una señal de audio en un segundo modo del codificador;
- unos medios de conmutación (6) para conmutar entre dicha parte del primer modo del codificador (5) y dicha parte del segundo modo del codificador (4);
- 60 - comprendiendo dicha parte del primer modo del codificador (5) una parte de codificación (9) la cual está adaptada para codificar una sección respectiva de dicha señal de audio basándose en por lo menos dos modelos de codificación diferentes; y
- 65 - comprendiendo asimismo dicha parte del primer modo del codificador (5) una parte de selección (13, 14, 15) adaptada para aplicar por lo menos una regla de selección con vistas a seleccionar un modelo de codificación específico, destinado dicho modelo de codificación a ser usado por dicha parte de codificación (9) para codificar dicha sección específica de una señal de audio, en el que dicha por lo menos una regla de

selección se basa en características de la señal, las cuales se han determinado por lo menos parcialmente a partir de una ventana de análisis que abarca por lo menos una sección de una señal de audio que precede a dicha sección específica, y en el que dicha parte de selección (13, 14, 15) está adaptada para activar dicha por lo menos una regla de selección después de una conmutación realizada por dichos medios de conmutación (6) desde dicha parte del segundo modo del codificador (4) hacia dicha parte del primer modo del codificador (5) en respuesta a la recepción de por lo menos tantas secciones de dicha señal de audio como las correspondientes abarcadas por dicha ventana de análisis.

9. Módulo (2,3) según la reivindicación 8, que comprende asimismo un contador (12) adaptado para contar el número de secciones de dicha señal de audio, las cuales se suministran a dicha parte del primer modo del codificador (5) después de una conmutación desde dicha parte del segundo modo del codificador (4) hacia dicha parte del primer modo del codificador (5).

10. Módulo (2,3) según la reivindicación 8 ó 9, en el que dicha parte del primer modo del codificador (5) comprende asimismo por lo menos una parte adicional de selección (16, 17, 18), la cual está adaptada para aplicar por lo menos una regla de selección adicional para seleccionar un modelo de codificación respectivo, destinado dicho modelo de codificación a ser usado por dicha parte de codificación (9) para codificar una sección específica de una señal de audio, en el que dicha por lo menos una regla de selección adicional no usa información sobre secciones de dicha señal de audio que preceden a dicha sección específica, y en el que dicha por lo menos una regla de selección adicional se aplica después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5) por lo menos siempre que el número de secciones recibidas por dicha parte del primer codificador (5) sea menor que el número de secciones abarcadas por una ventana de análisis utilizada para dicha por lo menos una regla de selección la cual se basa en un análisis de características de la señal en una ventana de análisis.

11. Módulo (2,3) según una de las reivindicaciones 8 a 10, en el que dicha por lo menos una parte de selección (13, 14, 15) comprende una primera parte de selección (14) adaptada para aplicar una primera regla de selección la cual se basa en características de la señal que se han determinado en una ventana de análisis más corta y una segunda parte de selección (13) adaptada para aplicar una segunda regla de selección, la cual se basa en características de la señal que se han determinado en una ventana de análisis más larga, en el que dicha primera regla de selección se activa en cuanto dicha parte del primer modelo del codificador (5) ha recibido las suficientes secciones de dicha señal de audio para dicha ventana de análisis más corta después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5), y en el que dicha segunda regla de selección se activa en cuanto dicha parte del primer modelo del codificador (5) ha recibido suficientes secciones de dicha señal de audio para dicha ventana de análisis más larga después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5).

12. Dispositivo electrónico (1) que respalda una codificación de una señal de audio, comprendiendo dicho dispositivo electrónico (2,3):

- una parte del primer modo del codificador (5) adaptada para codificar una sección respectiva de una señal de audio en un primer modo del codificador;
- una parte del segundo modo del codificador (4) adaptada para codificar una sección respectiva de una señal de audio en un segundo modo del codificador;
- unos medios de conmutación (6) para conmutar entre dicha parte del primer modo del codificador (5) y dicha parte del segundo modo del codificador (4);
- comprendiendo dicha parte del primer modo del codificador (5) una parte de codificación (9) la cual está adaptada para codificar una sección respectiva de dicha señal de audio basándose en por lo menos dos modelos de codificación diferentes; y
- comprendiendo asimismo dicha parte del primer modo del codificador (5) una parte de selección (13, 14, 15) adaptada para aplicar por lo menos una regla de selección con vistas a seleccionar un modelo de codificación específico, destinado dicho modelo de codificación a ser usado por dicha parte de codificación (9) para codificar dicha sección específica de una señal de audio, en el que dicha por lo menos una regla de selección se basa en características de la señal, las cuales se han determinado por lo menos parcialmente a partir de una ventana de análisis que abarca por lo menos una sección de una señal de audio que precede a dicha sección específica, y en el que dicha parte de selección (13, 14, 15) está adaptada para activar dicha por lo menos una regla de selección después de una conmutación realizada por dichos medios de conmutación (6) desde dicha parte del segundo modo del codificador (4) hacia dicha parte del primer modo del codificador (5) en respuesta a la recepción de por lo menos tantas secciones de dicha señal de audio como las correspondientes abarcadas por dicha ventana de análisis.

13. Dispositivo electrónico (1) según la reivindicación 12, que comprende asimismo un contador (12) adaptado para contar el número de secciones de dicha señal de audio, las cuales se suministran a dicha parte del primer modo del codificador (5) después de una conmutación desde dicha parte del segundo modo del codificador (4) hacia dicha parte del primer modo del codificador (5).

14. Dispositivo electrónico (1) según la reivindicación 12 ó 13, en el que dicha parte del primer modo del codificador (5) comprende asimismo por lo menos una parte adicional de selección (16, 17, 18), la cual está adaptada para aplicar por lo menos una regla de selección adicional para seleccionar un modelo de codificación respectivo, destinado dicho modelo de codificación a ser usado por dicha parte de codificación (9) para codificar una sección específica de una señal de audio, en el que dicha por lo menos una regla de selección adicional no usa información sobre secciones de dicha señal de audio que preceden a dicha sección específica, y en el que dicha por lo menos una regla de selección adicional se aplica después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5) por lo menos siempre que el número de secciones recibidas por dicha parte del primer codificador (5) sea menor que el número de secciones abarcadas por una ventana de análisis utilizada para dicha por lo menos una regla de selección la cual se basa en un análisis de características de la señal en una ventana de análisis.

15. Dispositivo electrónico (1) según una de las reivindicaciones 12 a 14, en el que dicha por lo menos una parte de selección (13, 14, 15) comprende una primera parte de selección (14) adaptada para aplicar una primera regla de selección la cual se basa en características de la señal que se han determinado en una ventana de análisis más corta y una segunda parte de selección (13) adaptada para aplicar una segunda regla de selección, la cual se basa en características de la señal que se han determinado en una ventana de análisis más larga, en el que dicha primera regla de selección se activa en cuanto dicha parte del primer modelo del codificador (5) ha recibido las suficientes secciones de dicha señal de audio para dicha ventana de análisis más corta después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5), y en el que dicha segunda regla de selección se activa en cuanto dicha parte del primer modelo del codificador (5) ha recibido suficientes secciones de dicha señal de audio para dicha ventana de análisis más larga después de una conmutación desde dicha parte del segundo modo del codificador (4) a dicha parte del primer modo del codificador (5).

16. Dispositivo electrónico (1) según la reivindicación 15, en el que una sección respectiva de dicha señal de audio se corresponde con una trama respectiva de la señal de audio que tiene una longitud de 20 ms, en el que dicha ventana más corta abarca una trama de la señal de audio para la cual se va a seleccionar un modelo de codificación y adicionalmente cuatro tramas anteriores de la señal de audio, y en el que dicha ventana más larga abarca una trama de la señal de audio para la cual se va a seleccionar un modelo de codificación y adicionalmente dieciséis tramas anteriores de la señal de audio.

17. Dispositivo electrónico (1) según una de las reivindicaciones 12 a 16, en el que dicha parte del primer modo del codificador (5) comprende asimismo una parte de determinación de características de la señal (11), la cual determina características de la señal de dicha señal de audio en una ventana de análisis respectiva y la cual suministra dichas características de la señal a dicha parte de selección (13, 14, 15), incluyendo dichas características de la señal una desviación estándar de valores relacionados con la energía en una ventana de análisis respectiva.

18. Dispositivo electrónico (1) según una de las reivindicaciones 12 a 17, en el que dicho primer modo del codificador es un modo ampliado de un códec ampliado de banda ancha multivelocidad adaptativa, estando adaptada dicha parte de codificación (9) de dicha parte del primer modo del codificador (5) para codificar secciones de una señal de audio basándose en un modelo de codificación de predicción lineal con excitación por código algebraico y adicionalmente basándose en un modelo de codificación por transformada, y en el que dicho segundo modo del codificador es un modo de banda ancha multivelocidad adaptativa correspondiente a dicho códec ampliado de banda ancha multivelocidad adaptativa, estando adaptada dicha parte del segundo modo del codificador (4) para codificar secciones de una señal de audio basándose en un modelo de codificación de predicción lineal con excitación por código algebraico.

19. Sistema de codificación de audio (1,2) que comprende un módulo (2,3) según una de las reivindicaciones 8 a 11 y un decodificador (20) para decodificar señales de audio, las cuales han sido codificadas por dicho módulo (2,3).

20. Sistema de codificación de audio (1,2) según la reivindicación 19, que comprende asimismo una parte del primer modo del codificador (5) adaptada para codificar una sección respectiva de una señal de audio en un primer modo del codificador.

21. Sistema de codificación de audio (1,2) según por lo menos una de las reivindicaciones 19 y 20, que comprende asimismo una parte del segundo modo del codificador (4) adaptada para codificar una sección respectiva de una señal de audio en un segundo modo del codificador.

22. Sistema de codificación de audio (1,2) según por lo menos una de las reivindicaciones 19 a 21, que comprende asimismo unos medios de conmutación (6) para conmutar entre dicha parte del primer modo del codificador (5) y dicha parte del segundo modo del codificador (4).

23. Producto de programa de software, en el que está almacenado un código de software para respaldar una codificación de una señal de audio, en el que hay disponibles por lo menos un primer modo del codificador y un segundo modo del codificador para codificar una sección respectiva de dicha señal de audio, en el que por lo menos dicho primer modo del codificador permite una codificación de una sección respectiva de dicha señal de audio basándose en por lo menos dos modelos de codificación diferentes, y en el que en dicho primer modo del codificador se habilita una sección de un modelo de codificación respectivo para codificar una sección específica de una señal de audio mediante por lo menos una regla de selección, la cual se basa en características de la señal que se han determinado a partir de

ES 2 291 877 T3

una ventana de análisis, la cual abarca por lo menos una sección de dicha señal de audio que precede a dicha sección específica, realizando dicho código de software la siguiente etapa cuando se ejecuta en un componente de procesado (3) de un codificador (2):

- 5 - activa dicha por lo menos una regla de selección después de una conmutación desde dicho segundo modo del codificador a dicho primer modo del codificador en respuesta a la recepción de por lo menos tantas secciones de dicha señal de audio como las correspondientes abarcadas por dicha ventana de análisis.

10

15

20

25

30

35

40

45

50

55

60

65

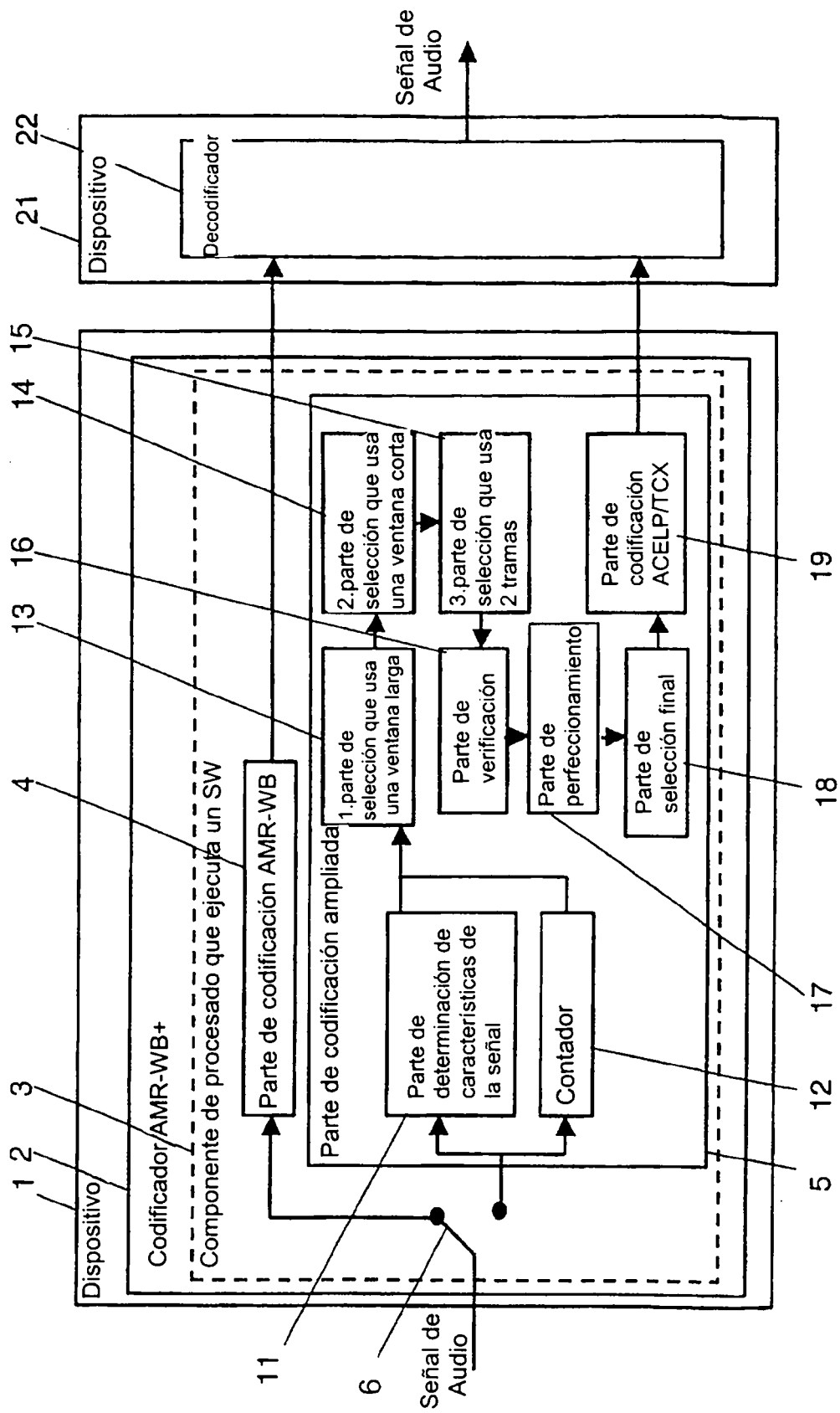


Fig. 1

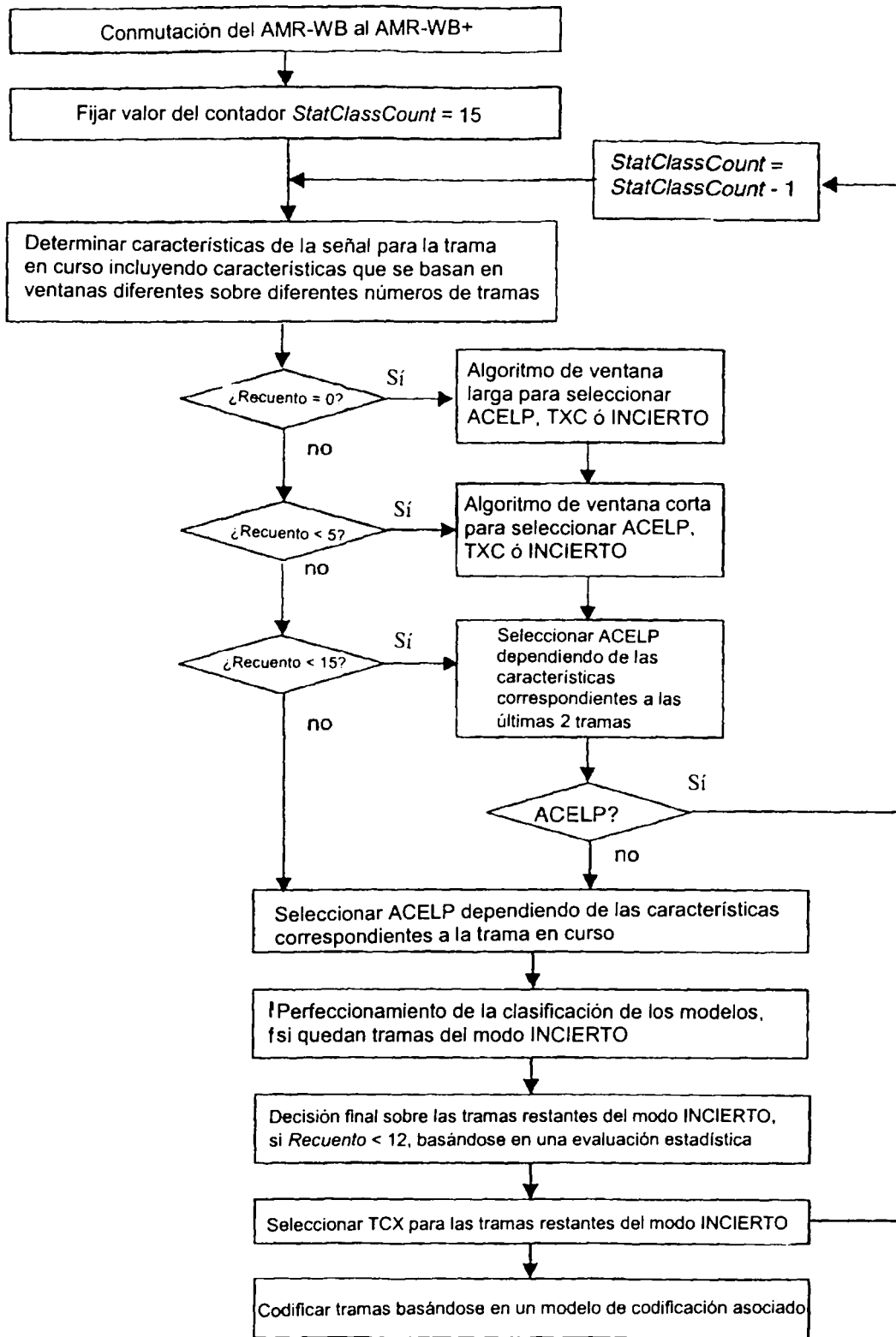


Fig. 2