

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 726 193**

51 Int. Cl.:

**G10L 19/04** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.08.2014 PCT/FI2014/050658**

87 Fecha y número de publicación internacional: **03.03.2016 WO16030568**

96 Fecha de presentación y número de la solicitud europea: **28.08.2014 E 14761388 (9)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019 EP 3186808**

54 Título: **Cuantificación de parámetros de audio**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**02.10.2019**

73 Titular/es:  
**NOKIA TECHNOLOGIES OY (100.0%)**  
**Karakaari 7**  
**02610 Espoo, FI**

72 Inventor/es:  
**RÄMÖ, ANSSI;**  
**VASILACHE, ADRIANA y**  
**LAAKSONEN, LASSE JUHANI**

74 Agente/Representante:  
**VALLEJO LÓPEZ, Juan Pedro**

ES 2 726 193 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

Cuantificación de parámetros de audio

### 5 **Campo técnico**

Las realizaciones de la presente invención a modo de ejemplo y no limitantes se refieren en general al campo de la codificación de audio y más específicamente al campo de la cuantificación de audio.

### 10 **Antecedentes**

Los codificadores y decodificadores de audio se usan para una amplia diversidad de aplicaciones en sistemas de comunicación, multimedia y almacenamiento. Un codificador de audio se usa para codificar señales de audio, como el habla, en particular para posibilitar una transmisión eficaz o almacenamiento de la señal de audio, mientras que un decodificador de audio construye una señal sintetizada basándose en una señal codificada recibida. Un par de un codificador de audio y un decodificador de audio se denomina como un códec de audio.

20 Cuando se implementa un códec de audio, este por lo tanto tiene por objetivo ahorrar capacidad de transmisión y almacenamiento mientras mantiene una alta calidad de la señal de audio sintetizada. También la robustez con respecto a la transmisión de errores es importante, especialmente con aplicaciones móviles y de voz sobre el protocolo de internet (VoIP). Por otra parte, la complejidad del códec de audio está limitada por la potencia de procesamiento de la plataforma de aplicación.

25 Un códec del habla (incluyendo un codificador del habla y un decodificador del habla) puede observarse como un códec de audio que está adaptado específicamente para codificar y decodificar señales de audio. En un codificador del habla típico, la señal del habla de entrada se procesa en segmentos, que se denominan tramas. Normalmente la longitud de trama es de 10 a 30 ms, mientras que un segmento anticipado que cubre por ejemplo 5–15 ms en el comienzo de la trama inmediatamente siguiente puede estar además disponible para el codificador. La longitud de trama puede fijarse (por ejemplo a 20 ms) o la longitud de trama puede variarse de trama a trama. Una trama puede dividirse adicionalmente en un número de sub tramas. Para cada trama, el codificador del habla determina una representación paramétrica de la señal de entrada. Los parámetros se cuantifican y transmiten a través de un canal de comunicación o se almacenan en un medio de almacenamiento en una forma digital. En el extremo de recepción, el decodificador del habla construye la señal sintetizada basándose en los parámetros recibidos.

35 La construcción de los parámetros y la cuantificación normalmente están basadas en libros de códigos, que contienen vectores de código optimizados para la respectiva tarea de cuantificación. En muchos casos, las relaciones de alta compresión requieren libros de códigos altamente optimizados. A menudo el rendimiento de un cuantificador puede mejorarse para una relación de compresión dada usando predicción de una o más tramas anteriores y/o de una o más tramas siguientes. Una cuantificación de este tipo se denominará a continuación como cuantificación predictiva, en contraste a una cuantificación no predictiva que no se basa en cualquier información de tramas precedentes. Una cuantificación predictiva aprovecha una correlación entre una trama de audio actual y al menos una trama de audio vecina para obtener una predicción para la trama actual de modo que por ejemplo únicamente tengan que codificarse las desviaciones a partir de esta predicción. Esto requiere libros de códigos especializados.

45 La cuantificación predictiva, sin embargo, puede dar como resultado problemas en el caso de errores en la transmisión o almacenamiento. Con cuantificación predictiva, una nueva trama no puede decodificarse perfectamente, incluso cuando se reciba correctamente, si al menos una trama precedente en la que está basada la predicción es errónea o falta. Es por lo tanto útil aplicar una cuantificación no predictiva en lugar de una predictiva de vez en cuando, por ejemplo a intervalos predefinidos (de número de tramas fijo), para evitar largas series de propagación de errores. Para una cuantificación no predictiva ocasional de este tipo, que también se denomina como cuantificación de "seguridad-red", puede aplicarse uno o más criterios de selección para seleccionar una de la cuantificación predictiva y de la cuantificación no predictiva en una base trama a trama para limitar la propagación de errores en el caso de un borrado de trama.

55 El documento US 2008/180307 A1 describe la cuantificación de representación de audio paramétrica que incluye dos modos de cuantificación: un primer modo predictivo y un segundo modo no predictivo. La decisión acerca de qué modo emplear en una trama se determina por el error de cuantificación, multiplicado por el factor de escala. El factor de escala es fijo y una cuantificación no predictiva se desencadena adicionalmente contando la racha de las tramas codificadas predictivamente anteriores.

### 60 **Sumario**

La invención proporciona una solución al problema técnico de acuerdo con las características de las reivindicaciones independientes.

65

**Breve descripción de las figuras**

Las realizaciones de la invención se ilustran a modo de ejemplo, y no por medio de limitación, en las figuras de los dibujos adjuntos.

- 5 La Figura 1 ilustra esquemáticamente algunos componentes de un sistema de acuerdo con una realización de ejemplo.
- La Figura 2 ilustra un método de acuerdo con una realización de ejemplo.
- 10 La Figura 3 ilustra un método de acuerdo con una realización de ejemplo.
- La Figura 4 ilustra un método de acuerdo con una realización de ejemplo.
- 15 La Figura 5 ilustra un método de acuerdo con una realización de ejemplo.
- La Figura 6 ilustra esquemáticamente algunos componentes de un aparato de acuerdo con una realización de ejemplo.

**Descripción de algunas realizaciones**

20 Incluso aunque los enfoques de cuantificación de seguridad-red señalados anteriormente pueden aplicarse para proporcionar una mejora sobre cuantificación predictiva pura en términos del rendimiento de codificación global con y sin errores en almacenamiento o transmisión, puede permanecer aún un considerable problema debido al carácter predictivo de la cuantificación.

25 Normalmente, dependiendo de las características de la señal de audio de entrada, la cuantificación predictiva puede proporcionar rendimiento de cuantificación que supere al de la cuantificación no predictiva hasta el 70 al 90 % de las tramas. El rendimiento superior de la cuantificación predictiva puede ser especialmente pronunciado durante segmentos de la señal del habla que muestran características espectrales estacionarias (por ejemplo habla vocalizada), que puede extenderse a través de decenas de tramas consecutivas, conduciendo posiblemente de esta manera a largas rachas de tramas consecutivas para las que se aplica cuantificación predictiva.

30 Como un ejemplo, un enfoque para mejorar el rendimiento global del enfoque de seguridad-red señalado anteriormente aumentando el uso de la cuantificación no predictiva incluye usar una ganancia de preferencia para favorecer la cuantificación no predictiva sobre la predictiva a pesar del mejor rendimiento de cuantificación proporcionado mediante la cuantificación predictiva. Es decir, la cuantificación predictiva puede requerirse para superar la no predictiva en un margen predefinido fijo (o en un factor predefinido fijo) para que se seleccione la cuantificación predictiva sobre la no predictiva. Como un ejemplo en este sentido, el requisito para seleccionar la cuantificación predictiva puede incluir que la cuantificación predictiva deba ser por ejemplo 1,3 veces mejor en términos de error de cuantificación que la cuantificación no predictiva (por ejemplo de manera que el error de cuantificación que resulta en la cuantificación predictiva multiplicado por 1,3 debe ser menor que el error de cuantificación que resulta en la cuantificación no predictiva para la misma trama), reduciendo por lo tanto el uso de la cuantificación predictiva. Esta opción da como resultado acortar las rachas de tramas consecutivas cuantificadas con la cuantificación predictiva (en dependencia del valor aplicado del margen predefinido fijo) y es por lo tanto adecuada para aumentar la robustez con respecto a los errores en transmisión o almacenamiento, mientras que por otra parte puede reducir el rendimiento de cuantificación en el caso de un canal de transmisión limpio o almacenamiento libre de errores. Además, seleccionar el valor para tal margen fijo predefinido puede no ser una tarea sencilla, corriendo de esta manera un riesgo de dar como resultado rachas de tramas consecutivas más cortas a las deseadas o más largas a las deseadas cuantificadas con la cuantificación predictiva.

50 Como otro ejemplo, el rendimiento del enfoque de seguridad-red implica establecer un valor máximo para una racha de tramas consecutivas cuantificadas con la cuantificación predictiva. Aunque este enfoque es eficaz al limitar la longitud máxima de la propagación del error en el caso de un borrado de trama o error de trama, falla al tener en cuenta diferencias en la mejora de rendimiento proporcionada por la cuantificación predictiva en señales de audio de diferentes características. Por lo tanto, este enfoque también implica un riesgo de dar como resultado rachas de tramas consecutivas más cortas a las deseadas o más largas a las deseadas cuantificadas con la cuantificación predictiva. Además, la terminación forzada de una racha de tramas consecutivas cuantificadas de manera predictiva puede tener lugar en una trama donde el rendimiento de cuantificación de la cuantificación predictiva es superior al de la cuantificación no predictiva, imponiendo de esta manera un riesgo de una seria degradación de calidad de audio a corto plazo.

65 La presente invención continúa a partir de la consideración de que usar el enfoque seguridad-red para interrumpir una racha de tramas cuantificadas de manera predictiva forzando una trama cuantificada de manera no predictiva sirve para evitar preventivamente la posible propagación de errores, mientras por otra parte la interrupción forzada de la racha de tramas cuantificadas de manera predictiva, especialmente en una trama donde la mejora de rendimiento proporcionada mediante la cuantificación predictiva es significativa, es probable que comprometa el rendimiento de

cuantificación global a corto plazo y por lo tanto conduzca a calidad de audio comprometida. Se propone por lo tanto que el criterio de selección aplicado al seleccionar entre cuantificación predictiva y no predictiva para una trama dada esté dispuesto para provocar preferir la cuantificación no predictiva sobre la cuantificación predictiva en un factor que se aumenta con longitud creciente de una racha de tramas consecutivas para las que se ha seleccionado la cuantificación predictiva. En paralelo, puede evaluarse uno o más criterios de selección adicionales para seleccionar entre cuantificaciones predictivas y no predictivas.

En consecuencia, las realizaciones de la presente invención proporcionan una posibilidad de aumentar el rendimiento de codificación de audio en caso de errores de canal contribuyendo hacia acortar rachas extensivamente largas de tramas consecutivas en las que se ha aplicado la cuantificación predictiva mientras aún hace uso del rendimiento superior de la cuantificación predictiva siempre que el rendimiento supere claramente el de la cuantificación no predictiva. Aunque un enfoque de este tipo puede dar como resultado aumentar el error de cuantificación promedio objetivo, el criterio de selección puede adaptarse para garantizar mantener el error de cuantificación al nivel que presente cualquier posible imprecisión resultante al modelar de la señal de audio suficientemente pequeña para que el error sea difícilmente audible o no audible en absoluto.

La distorsión espectral (SD) es un ejemplo de una medida comúnmente aplicada para indicar la cantidad de error de cuantificación, y SD también es adecuada para evaluar la audibilidad de un error de cuantificación. Puede suponerse, por ejemplo, que si una SD debida a una cuantificación radica por debajo de 1 dB, la distorsión es normalmente inaudible para un oído humano, en el enfoque de seguridad-red puede hacerse uso de este hecho, por ejemplo, seleccionando la cuantificación no predictiva siempre que produzca una SD que está por debajo de un umbral predefinido, por ejemplo 1 dB. Para ilustrar adicionalmente este aspecto, en general no es necesario cuantificar un segmento de señal de audio particular, por ejemplo con cuantificación predictiva, para obtener por ejemplo una SD muy baja de 0,5 dB, si la cuantificación no predictiva del mismo segmento de señal de audio da como resultado una SD de 0,9 dB, que ya es suficiente desde el punto de vista del oído humano. En un caso de este tipo, aunque el error de cuantificación objetivo resultante de la cuantificación no predictiva es mayor para el segmento de señal de audio individual, el error de cuantificación resultante puede considerarse aún inaudible y por lo tanto puede ser ventajoso seleccionar la cuantificación no predictiva para este segmento de audio particular para facilitar limitar o prevenir la propagación de error de cuantificación debido a un borrado de trama o pérdida de trama: si hubiera un borrado o pérdida de segmento de señal de audio anterior a este segmento de señal de audio, la cuantificación predictiva se realizaría pobremente, pero los parámetros obtenidos de una cuantificación no predictiva podrían decodificarse perfectamente. En un enfoque de este tipo, una mejora debido al uso de la cuantificación no predictiva en lugar de la predictiva se hace audible únicamente para los segmentos de señal de audio con uno o más errores, mientras que para canales despejados normalmente no hay degradación audible. En consecuencia, tal técnica puede aplicarse como parte del enfoque de seguridad-red para contribuir hacia suficiente calidad de cuantificación tanto en condiciones de canal limpias como en presencia de borrados/errores de trama, posiblemente en paralelo con criterios de selección adicionales para selección entre las cuantificaciones predictivas y no predictivas.

Como se hace evidente a partir de la descripción anterior, una medida de error adecuada que puede compararse con un umbral predeterminado puede estar por lo tanto relacionada con una distorsión espectral sobre un intervalo de frecuencia entre el segmento de señal de audio original y un segmento de señal de audio que resulta en una cuantificación. Tal medida de error puede calcularse para tanto la cuantificación predictiva como la cuantificación no predictiva. Calcular la medida de error en términos de distorsión espectral sobre el intervalo de frecuencia es también adecuado, por ejemplo, para parámetros de frecuencia espectral de inmitancia (ISF) o parámetros de frecuencia espectral lineal (LSF) que pertenecen a un segmento de señal de audio.

La distorsión espectral  $SD$  para un respectivo segmento de señal de audio (por ejemplo una trama de la señal de audio) puede representarse por la siguiente ecuación:

$$SD = \frac{1}{\pi} \int_0^{\pi} \left[ \log S(\omega) - \log \hat{S}(\omega) \right]^2 d\omega, \quad (1)$$

Donde  $\hat{S}(\omega)$  y  $S(\omega)$  son los espectros de la trama del habla con y sin cuantificación, respectivamente. Aunque esta distorsión espectral sería, por ejemplo, una medida particularmente exacta para el libro de códigos y selección de cuantificación de los parámetros de codificación predictiva lineal (LPC) en un segmento de audio, el esfuerzo computacional para determinar la distorsión espectral de acuerdo con la ecuación (1) podría reducirse usando métodos computacionalmente más sencillos.

En este sentido, la medida de error considerada puede comprender una medida de error que al menos se aproxima a la distorsión espectral (por ejemplo de acuerdo con la ecuación (1)). Una medida de error de este tipo puede obtenerse, por ejemplo, combinando errores ponderados entre un componente del segmento de señal de audio original y un componente correspondiente del segmento de señal de audio que resulta en la cuantificación. La medida de error puede ser, por ejemplo, una medida de error pisco acústicamente significativa, obtenida por ejemplo combinando

errores cuadráticos medios ponderados, donde la ponderación de errores proporciona una ponderación psico acústicamente significativa. La expresión ponderación psico acústicamente significativa significa que aquellos componentes espectrales en una señal de audio que se reconocen por el oído humano se destacan en comparación con aquellos que aparentemente no son reconocidos por el oído humano. Tal ponderación puede proporcionarse por un conjunto de factores de ponderación que pueden aplicarse para multiplicar componentes respectivos del segmento de señal de audio a ponderar o componentes respectivos del parámetro de audio a ponderar para formar un conjunto de componentes ponderados, componentes ponderados que se combinan a continuación (por ejemplo suman) para formar la medida de error ponderada. Pueden calcularse factores de ponderación adecuados para este fin de diversas maneras.

Un ejemplo de un error psico acústicamente significativo de este tipo puede comprender un error ponderado, por ejemplo un error cuadrático medio ponderado, entre parámetros de ISF originales (no cuantificados) y parámetros de ISF correspondientemente cuantificados. Como otro ejemplo, un error psico acústicamente significativo puede comprender un error ponderado, por ejemplo un error cuadrático medio ponderado entre parámetros de LSF originales (no cuantificados) y parámetros de LSF cuantificados correspondientes.

En general, se ha de entender que la medida de error considerada puede determinarse basándose en el segmento de señal de audio completamente cuantificado o en un segmento de señal de audio parcialmente cuantificado, por ejemplo basándose en uno o más parámetros cuantificados seleccionados en el respectivo segmento de señal de audio, por ejemplo los parámetros de ISF o los parámetros de LSF a los que se ha hecho referencia anteriormente.

La Figura 1 representa un diagrama de bloques esquemático de un sistema a modo de ejemplo, en el que puede implementarse una selección de una cuantificación predictiva o no predictiva de acuerdo con una realización de la invención. En este texto, las expresiones cuantificación no predictiva y cuantificación de seguridad-red se usarán de manera sinónima.

El sistema ilustrado en la Figura 1 comprende un primer dispositivo 100 electrónico y un segundo dispositivo 150 electrónico. El primer dispositivo 100 electrónico está configurado para codificar datos de audio, por ejemplo para una transmisión de banda ancha, y el segundo dispositivo 150 electrónico está configurado para decodificar datos de audio codificados. El primer dispositivo 100 electrónico comprende un componente 111 de entrada de audio, que está enlazado mediante un chip 120 a un componente 112 de transmisión (TX). El componente 111 de entrada de audio puede ser por ejemplo un micrófono, un conjunto de micrófono, una interfaz a otro dispositivo que proporciona datos de audio o una interfaz a una memoria o a un sistema de ficheros desde el que pueden leerse datos de audio.

El chip 120 puede ser por ejemplo un circuito integrado (IC), que incluye circuitería para un codificador 121 de audio, del cual se ilustran esquemáticamente bloques funcionales seleccionados. Incluyen un componente 124 de parametrización y un componente 125 de cuantificación. El componente 112 de transmisión está configurado para posibilitar una transmisión de datos a otro dispositivo, por ejemplo al dispositivo 150 electrónico, mediante un enlace alámbrico o inalámbrico. El codificador 121 o el chip 120 podrían observarse como un aparato a modo de ejemplo de acuerdo con la invención, y el componente de cuantificación como que representa componentes de procesamiento correspondientes.

El dispositivo 150 electrónico comprende un componente 162 de recepción, que se enlaza mediante un chip 170 a un componente 161 de salida de audio. El componente 162 de recepción está configurado para posibilitar una recepción de datos desde otro dispositivo, por ejemplo desde el dispositivo 100 electrónico, mediante un enlace alámbrico o inalámbrico. El chip 170 puede ser por ejemplo un circuito integrado (IC), que incluye circuitería para un decodificador 171 de audio, del cual se ilustra un componente 174 de sintetización. El componente 161 de salida de audio puede ser por ejemplo un altavoz o una interfaz a otro dispositivo, al que se han de reenviar los datos de audio decodificados.

Se ha de entender que las conexiones representadas de la Figura 1 pueden realizarse mediante diversos componentes no mostrados en las mismas.

Una operación en el sistema de la Figura 1 se describirá ahora en más detalle con referencia a las Figuras 2 a 5.

La Figura 2 representa un diagrama de flujo que ilustra la operación en el codificador 121 de audio como etapas de un método 200 a modo de ejemplo. Cuando una señal de audio se introduce en el dispositivo 100 electrónico, por ejemplo mediante el componente 111 de entrada de audio, puede proporcionarse al codificador 121 de audio para codificación. Antes de que la señal de audio se proporcione al codificador 121 de audio, puede someterse a algún pre-procesamiento. En el caso de que una señal de audio de entrada sea una señal de audio analógica, por ejemplo, puede someterse en primer lugar a una conversión de analógico a digital, etc.

El codificador 121 de audio procesa la señal de audio por ejemplo en tramas de audio de 20 ms, usando una anticipación de 10 ms. Cada trama de audio constituye un segmento de señal de audio. El componente 124 de parametrización convierte en primer lugar la trama de audio actual en una representación de parámetros (etapa 201). La representación de parámetros para una trama de audio de la señal de audio puede incluir uno o más parámetros de audio que son descriptivos de la señal de audio en la trama, mientras que un parámetro de audio puede ser un

parámetro escalar (único) o un parámetro vectorial. En el siguiente ejemplo, se describe el procesamiento de acuerdo con diversas realizaciones de la presente invención con referencias a los parámetros de LSF y/o ISF de una manera a modo de ejemplo y no limitante.

5 El componente 125 de cuantificación realiza por un lado una cuantificación no predictiva de uno o más parámetros de la trama de audio (etapa 211), por ejemplo usando un libro de códigos no predictivo. El componente 125 de cuantificación puede realizar una cuantificación de parámetros seleccionados únicamente en esta etapa, mientras que los parámetros adicionales pueden cuantificarse en una etapa posterior (por ejemplo, después de la selección de una de las cuantificaciones predictivas y no predictivas basándose en la etapa 203). Además, el componente 125 de  
10 cuantificación deriva un valor de una medida de error que es descriptiva de un error de cuantificación  $E_1$  que resulta en una cuantificación no predictiva del uno o más parámetros de audio de la trama de audio (etapa 212). Usando un vector de LSF que comprende los parámetros de LSF que son descriptivos de características espectrales de la trama de audio como un ejemplo, el error de cuantificación  $E_1$  puede comprender, por ejemplo, un error cuadrático medio entre los parámetros de LSF cuantificados con la cuantificación no predictiva y los parámetros de LSF originales (no cuantificados) para la trama de audio un error cuadrático medio ponderado entre los parámetros de LSF cuantificados con la cuantificación no predictiva y los parámetros de LSF originales (no cuantificados) para la trama de audio, donde la ponderación es una ponderación psico acústicamente significativa.

El componente 125 de cuantificación realiza, por otra parte, una cuantificación predictiva de uno o más parámetros de la trama de audio (etapa 221) por ejemplo usando un libro de códigos predictivo. El componente 125 de cuantificación puede realizar de nuevo una cuantificación de parámetros seleccionados únicamente en esta etapa (por ejemplo después de la selección de una de las cuantificaciones predictivas y no predictivas basándose en la etapa 203), mientras que los parámetros adicionales pueden cuantificarse en una etapa posterior. Además, el componente 125 de  
20 cuantificación deriva un valor de una medida de error que es descriptiva de un error de cuantificación  $E_2$  que resulta en una cuantificación predictiva del uno o más parámetros de audio de la trama de audio (etapa 222). Como en el caso de la etapa 212, usar el vector de LSF como un ejemplo de un parámetro de audio, el error de cuantificación  $E_1$  puede comprender, por ejemplo, un error cuadrático medio o un error cuadrático medio ponderado (psico acústicamente) entre los parámetros de LSF cuantificados con la cuantificación predictiva y los parámetros de LSF originales (no cuantificados) para la trama de audio.

La cuantificación predictiva puede comprender, por ejemplo, usar cualquier método de predicción conocido en la técnica para calcular un valor predicho de un parámetro de audio (por ejemplo, un vector de LSF o un componente del mismo) en la trama de audio actual  $i$  basándose en el valor del respectivo parámetro de audio (por ejemplo el vector de LSF o un componente del mismo) en una o más tramas que preceden a la trama de audio  $i$  (por ejemplo tramas de  
35 audio  $i - j$ , donde  $j = 1, \dots, j_{\max}$ ) y/o basándose en una o más tramas que siguen la trama de audio  $i$  (por ejemplo, las tramas de audio  $i + k$ , donde  $k = 1, \dots, k_{\max}$ ) y usando un cuantificador para cuantificar la diferencia entre el valor original (no cuantificado) del parámetro de audio en la trama de audio actual y el valor predicho (por ejemplo basándose en un libro de códigos predictivo).

40 En este sentido, el componente 125 de cuantificación puede aplicar un modelo de predicción lineal o de predicción no lineal para la cuantificación predictiva. Como un ejemplo ilustrativo y no limitante, la predicción en este sentido puede comprender calcular el valor predicho del parámetro de audio para la trama de audio  $i$  basándose en el valor del respectivo parámetro de audio en la trama de audio más cercana (por ejemplo la más reciente)  $i - 1$  usando uno de un modelo de predicción autorregresivo (AR), un modelo de predicción de media móvil (MA) y un modelo de predicción de media móvil autorregresiva (ARMA).

A continuación, el componente 125 de cuantificación selecciona cualquiera de una cuantificación no predictiva o una cuantificación predictiva para la trama de audio actual basándose en los respectivos errores de cuantificación determinados  $E_1$  y  $E_2$ . En este sentido, el componente 125 de cuantificación puede determinar si el error de cuantificación  $E_2$  supera el error de cuantificación  $E_1$ , en al menos un margen adaptativo  $M$  (etapa 203). El margen adaptativo  $M$  es dependiente del número de tramas consecutivas que preceden a la trama de audio actual en la que se proporciona el uno o más parámetros de audio cuantificados con cuantificación predictiva. En otras palabras, el margen adaptativo  $M$  para la trama actual es dependiente del número de tramas entre la trama de audio precedente más cercana para la que se ha seleccionado la cuantificación no predictiva y la trama actual. Este número de tramas  
55 puede indicarse como la longitud  $L$  de racha de predicción (actual). La determinación del margen adaptativo  $M$  se describe más tarde en este texto.

Si la determinación en la etapa 203 es afirmativa, es decir en el caso de que el error de cuantificación  $E_2$  supere el error de cuantificación  $E_1$  en al menos el margen adaptativo  $M$ , el componente 125 de cuantificación proporciona uno o más parámetros de audio de la trama de audio actual cuantificados con la cuantificación no predictiva (etapa 213) como parte de la señal de audio codificada. En contraste, si la determinación en la etapa 203 no es afirmativa, es decir en el caso de que el error de cuantificación  $E_2$  falle al superar el error de cuantificación  $E_1$  en al menos el margen adaptativo  $M$ , el componente 125 de cuantificación proporciona uno o más parámetros de audio de la trama de audio actual cuantificados con la cuantificación predictiva (etapa 223) como parte de señal de audio codificada.

65 El componente 125 de cuantificación puede aplicar, como alternativa o adicionalmente, uno o más criterios adicionales

que provocan la selección de la cuantificación no predictiva y por lo tanto el método 200 puede variarse, por ejemplo, introduciendo una o más etapas adicionales de determinación o selección antes o después de la etapa 203. Como un ejemplo en este sentido, en una variación del método 200 el componente 125 de cuantificación puede determinar antes de la etapa 203 si el error de cuantificación  $E_1$  es menor que un umbral predefinido  $E_{th}$ , continuar a la etapa 213 en el caso de que esta determinación sea afirmativa, y continuar a la etapa 203 en el caso de que esta determinación no sea afirmativa. El umbral  $E_{th}$  puede ser un umbral por debajo del cual el error de cuantificación  $E_1$  puede considerarse que es inaudible. Un valor apropiado para el umbral  $E_{th}$  es diferente para diferentes parámetros de audio y posibles diferentes funciones de ponderación aplicadas para ponderar el error de cuantificación, y se ha de calcular por ensayo y error fuera de línea. Pero una vez que se ha encontrado un valor apropiado para el umbral  $E_{th}$ , la complejidad computacional aumenta en el codificador debido a que la verificación en la etapa 302 es mínima. Como un ejemplo, el umbral  $E_{th}$  puede establecerse a un valor que corresponde una SD en el intervalo de 0,8 a 1,0 dB, por ejemplo 0,9 dB.

Como un ejemplo de determinación del margen adaptativo  $M$  en dependencia de la longitud de racha de predicción  $L$ , el margen  $M$  puede aumentarse desde su valor inicial  $M_0$  en una cantidad predefinida  $M_s$  para cada trama de audio entre la trama de audio actual y la trama de audio precedente más cercana para la que se ha seleccionado la cuantificación no predictiva.

Como otro ejemplo de determinación del margen adaptativo  $M$  en dependencia de la longitud de racha de predicción  $L$ , el margen  $M$  puede aumentarse desde su valor inicial  $M_0$  en una cantidad predefinida  $M_s$  para cada trama de audio en exceso de un umbral predefinido  $L_0$  entre la trama de audio actual y la trama de audio precedente más cercana para la que se ha seleccionado la cuantificación no predictiva. En otras palabras, el margen  $M$  puede aumentarse desde su valor inicial  $M_0$  en una cantidad predefinida  $M_s$  ( $L - L_0$ ) veces, con la condición de que  $L$  sea mayor que  $L_0$ .

Como un ejemplo, el umbral  $L_0$  puede establecerse a un valor predeterminado fijo, por ejemplo a tres (por ejemplo  $L_0 = 3$ ), pero de igual manera a cualquier otro valor deseado. Como otro ejemplo, el valor del umbral  $L_0$  puede establecerse (o ajustarse) en dependencia de las características de audio de la trama actual y/o una o más tramas inmediatamente precedentes a la trama actual. Como un ejemplo adicional, el valor del umbral  $L_0$  puede establecerse (o ajustarse) en dependencia de un modo de codificación aplicado por el codificador 121 de audio o por el componente 125 de cuantificación para la trama actual y/o para una o más tramas inmediatamente precedentes a la trama actual.

En la estructura del método 200, el margen adaptativo  $M$  se resetea al valor inicial  $M_0$  (etapa 214) para la siguiente trama de audio en el caso de que se haya seleccionado la cuantificación no predictiva para la trama de audio actual o se haya adaptado (etapa 224) por la cantidad predefinida  $M_s$  para la siguiente trama de audio en el caso de que se haya seleccionado la cuantificación predictiva para la trama de audio actual.

Como otro ejemplo, resetear el margen adaptativo  $M$  (etapa 214) y/o la adaptación del margen adaptativo  $M$  (etapa 224) pueden tener lugar, basándose en la cuantificación seleccionada para la trama precedente más cercana (es decir la trama precedente más reciente), después de la recepción de la siguiente trama de audio pero antes de la comparación de los errores de cuantificación  $E_1$  y  $E_2$  (en la etapa 203) en su lugar. Como un ejemplo adicional, en lugar de restear explícitamente el margen adaptativo  $M$  (etapa 214) y ajustar el margen adaptativo  $M$  (etapa 224), el margen adaptativo  $M$  puede calcularse basándose en la longitud de racha de predicción  $L$  o basándose en la longitud de racha de predicción  $L$  y el umbral predefinido  $L_0$ . O el margen adaptativo  $M$  puede obtenerse a partir de una tabla accesible por el componente 125 de cuantificación, tabla que almacena valores del margen adaptativo  $M$  a través de un intervalo deseado de valores de la longitud de racha de predicción  $L$ . Ejemplos en este sentido se describirán más adelante en este texto.

El valor inicial  $M_0$  para el margen adaptativo  $M$  puede ser cero o sustancialmente cero. Como alternativa, el valor inicial  $M_0$  para el margen adaptativo  $M$  puede ser ligeramente por encima de cero. Usar un valor inicial  $M_0$  ligeramente por encima de cero sirve para asegurar preferir la cuantificación no predictiva sobre la cuantificación predictiva incluso cuando la longitud de racha de predicción  $L$  es cero (o por debajo del umbral  $L_0$ ). La cantidad predefinida  $M_s$  por la que se ha de ajustar el margen adaptativo  $M$  para uso en la siguiente trama de audio puede ser un valor positivo pequeño para aumentar gradualmente el margen adaptativo  $M$  trama a trama para, finalmente, prácticamente forzar el aprovisionamiento del uno o más parámetros de audio de una trama de audio cuantificada con la cuantificación no predictiva como parte de la señal de audio codificada.

La Figura 3 representa un diagrama de flujo que ilustra la operación en el codificador 121 de audio como etapas de un método 300 a modo de ejemplo. El método 300 sirve como una realización de ejemplo en la estructura descrita anteriormente con referencias al método 200. El método 300 comparte las etapas 201, 211 y 221 con el método 300.

En el método 300, el componente 125 de cuantificación puede derivar un error de cuantificación  $E_{s-net}$  que resulta en una cuantificación no predictiva del uno o más parámetros de audio de la trama de audio actual (etapa 312). Como un ejemplo, el error de cuantificación  $E_{s-net}$  puede comprender un error cuadrático medio entre los parámetros de audio cuantificados con la cuantificación no predictiva y los respectivos parámetros de audio original (no cuantificados) en la trama de audio actual. Como otro ejemplo, el error de cuantificación  $E_{s-net}$  puede comprender una medida de error psico acústicamente relevante, tal como una SD o un error cuadrático medio ponderado (psico acústicamente) entre

los parámetros de audio cuantificados con la cuantificación no predictiva y los respectivos parámetros de audio original (no cuantificados) en la trama de audio actual. Usando los parámetros de LSF como un ejemplo de uno o más parámetros de audio, el error de cuantificación  $E_{s-net}$  puede proporcionarse, por ejemplo, como un error cuadrático medio ponderado entre los parámetros de LSF cuantificados con la cuantificación no predictiva y los parámetros de LSF originales para la trama actual  $i$ , por ejemplo, de acuerdo con la ecuación (2).

$$E_{s-net}^i = \sum_{p=0}^{N-1} W_p^i (QLsfs_p^i - Lsf_p^i)^2, \quad (2)$$

donde  $N$  es la longitud del vector cuantificado (por ejemplo el número de elementos en el vector), donde  $QLsfs_p^i$  es un valor  $p$  de vector de LSF óptimo cuantificado de seguridad-red para la trama  $i$ , donde  $Lsf_p^i$  es el valor  $p$  de vector

de LSF no cuantificado original para la trama  $i$ , y donde  $W_p^i$  es un valor  $p$  de vector de ponderación psico acústicamente relevante para la trama  $i$ . En este sentido, ejemplos de un vector de ponderación adecuado  $W$  incluyen la función de ponderación  $w_{end}$  descrita en la sección 6.8.2.4 de la Recomendación de la ITU-T G.718 (06/2008), Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s (donde el acrónimo ITU-T significa la Unión Internacional de Telecomunicaciones, sector de normalización de Telecomunicaciones) y el vector de ponderación  $w_{mid}$  descrito en la sección 6.8.2.6 de dicha recomendación de la ITU-T G.718.

Continuando la descripción del método 300, el componente 125 de cuantificación puede derivar un error de cuantificación  $E_{pred}$  que resulta en una cuantificación no predictiva del uno o más parámetros de audio de la trama de audio actual (etapa 322). Como un ejemplo, el error de cuantificación  $E_{pred}$  puede comprender un error cuadrático medio entre los parámetros de audio cuantificados con la cuantificación predictiva y los respectivos parámetros de audio original (no cuantificados) en la trama de audio actual. Como otro ejemplo, el error de cuantificación  $E_{pred}$  puede comprender una medida de error psico acústicamente relevante, tal como una SD o un error cuadrático medio ponderado (psico acústicamente) entre los parámetros de audio cuantificados con la cuantificación predictiva y los respectivos parámetros de audio original (no cuantificados) en la trama de audio actual. Usando, de nuevo, los parámetros de LSF como un ejemplo de uno o más parámetros de audio, el error de cuantificación  $E_{pred}$  puede proporcionarse por ejemplo como un error cuadrático medio ponderado entre los parámetros de LSF cuantificados con la cuantificación predictiva y los parámetros de LSF originales para la trama actual  $i$  por ejemplo de acuerdo con la ecuación (3).

$$E_{pred}^i = \sum_{p=0}^{N-1} W_p^i (QLsfp_p^i - Lsf_p^i)^2, \quad (3)$$

donde  $N$  es de nuevo la longitud del vector cuantificado (por ejemplo el número de elementos en el vector), donde  $QLsfp_p^i$  es un valor  $p$  de vector de LSF óptimo cuantificado predictivo para la trama  $i$ , donde  $Lsf_p^i$  es de nuevo

el valor  $p$  de vector de LSF no cuantificado original para la trama  $i$ , y donde  $W_p^i$  es de nuevo un valor  $p$  de vector de ponderación psico acústicamente relevante para la trama  $i$ , por ejemplo de acuerdo con la ecuación (3). Además, las consideraciones con respecto a un vector de ponderación adecuado  $W$  proporcionado en contexto de la ecuación (2) son válidas también para la ecuación (3).

Continuando aún con la descripción del método 300, el componente 125 de cuantificación selecciona cualquiera de la cuantificación predictiva o la no predictiva basándose en los errores de cuantificación  $E_{s-net}$  y  $E_{pred}$ . En particular, el componente 125 de cuantificación puede determinar si un valor cambiado de escala del error de cuantificación  $\hat{E}_{s-net}$  es menor que el error de cuantificación  $E_{pred}$ , en el que  $\hat{E}_{s-net}$  es el error de cuantificación  $E_{s-net}$  cambiado de escala por el valor actual de un factor de cambio de escala adaptativo  $m$ , por ejemplo  $\hat{E}_{s-net} = mE_{s-net}$  (etapa 303).

Si la determinación en la etapa 303 es afirmativa, es decir en el caso de que el error de cuantificación  $E_{s-net}$  cambiado de escala por el valor actual de un factor de cambio de escala adaptativo  $m$  sea menor que el error de cuantificación  $E_{pred}$ , el componente 125 de cuantificación proporciona uno o más parámetros de audio de la trama de audio actual, por ejemplo al menos los parámetros de LSF, cuantificados con la cuantificación no predictiva (etapa 213) como parte de la señal de audio codificada. En contraste, si la determinación en la etapa 303 no es afirmativa, es decir en el caso de que el error de cuantificación  $E_{s-net}$  cambiado de escala por el valor actual de un factor de cambio de escala adaptativo  $m$  no sea menor que el error de cuantificación  $E_{pred}$ , el componente 125 de cuantificación proporciona uno o más parámetros de audio de la trama de audio actual, por ejemplo al menos los parámetros de LSF, cuantificados con la cuantificación predictiva (etapa 223) como parte de la señal de audio codificada.

Aún en el método 300, en el caso de que el componente 125 de cuantificación haya seleccionado la cuantificación no



5 predictiva para el uno o más parámetros de audio en la trama de audio actual  $i$ , el componente 125 de cuantificación puede resetear adicionalmente el factor de cambio de escala adaptativo  $m$  para uso mediante el componente 125 de cuantificación en la siguiente trama de audio  $i+1$  estableciendo el factor de cambio de escala adaptativo  $m$  a un valor inicial  $m_0$ , es decir establecer  $m = m_0$  (etapa 314). Esto corresponde a resetear el margen adaptativo  $M$  a su valor inicial  $M_0$  en la etapa 214 del método 200.

10 En contraste, en el caso de que el componente 125 de cuantificación haya seleccionado la cuantificación predictiva para el uno o más parámetros de audio en la trama de audio actual  $i$ , el componente 125 de cuantificación puede ajustar adicionalmente el factor de cambio de escala adaptativo  $m$  para uso mediante el componente 125 de cuantificación en la siguiente trama de audio  $i+1$  multiplicando el factor de cambio de escala  $m$  por un factor de cambio de escala predefinido  $m_s$ , es decir establecer  $m = m * m_s$  (etapa 324). Esto corresponde a ajustar el margen adaptativo  $M$  por la cantidad predefinida  $M_s$  en la etapa 224 del método 200.

15 El valor inicial  $m_0$  para el factor de cambio de escala adaptativo  $m$  puede ser uno (por ejemplo  $m_0 = 1$ ) o sustancialmente uno. Como una variación de este enfoque, el valor inicial  $m_0$  puede ser ligeramente por debajo de uno, por ejemplo en el intervalo de 0,9 a 0,99 para asegurar preferir la cuantificación no predictiva sobre la cuantificación predictiva incluso cuando la longitud  $L$  de la racha sea cero, es decir en una trama que sigue inmediatamente una trama para la que se ha seleccionado la cuantificación no predictiva. Como un ejemplo ilustrativo para asegurar una preferencia constante para selección de la cuantificación no predictiva, la condición en la etapa 303 puede reescribirse como

$$mE_{s-net} < nE_{pred}, \quad (4)$$

25 con un factor de cambio de escala predefinido  $n$  establecido por ejemplo a un valor en el intervalo de 1,01 a 1,1, por ejemplo como  $n = 1,05$  mientras que el valor inicial  $m_0$  para el factor de cambio de escala  $m$  se establece a uno (por ejemplo  $m_0 = 1$ ).

30 El factor de cambio de escala predefinido  $m_s$  puede ser un valor positivo menor que uno para reducir el factor de cambio de escala adaptativo  $m$  para la siguiente trama  $i+1$ . En este sentido, el factor de cambio de escala predefinido  $m_s$  puede establecerse a un valor seleccionado a partir de un intervalo de 0,7 a 0,95, por ejemplo  $m_s = 0,8$ . Esto corresponde a aumentar el margen adaptativo  $M$  trama a trama durante una racha de tramas de audio consecutivas para las que se ha seleccionado la cuantificación predictiva.

35 La Figura 4 representa un diagrama de flujo que ilustra la operación en el codificador 121 de audio como etapas de un método 400 a modo de ejemplo. El método 400 se proporciona como una variación del método 300 y sirve como otra realización de ejemplo en la estructura descrita anteriormente con referencias al método 200. El método 400 comparte todas las etapas del método 300, mientras que se introduce una etapa 302 de verificación adicional antes de la determinación de la etapa 303.

40 La etapa 302 proporciona un criterio adicional para seleccionar la cuantificación no predictiva para uno o más parámetros de audio de la trama de audio actual. En particular, el componente 125 de cuantificación puede seleccionar la cuantificación no predictiva en el caso de que el error de cuantificación  $E_{s-net}$  sea menor que un umbral predefinido  $E_{th}$ . En contraste, el componente 125 de cuantificación puede continuar a la etapa 303 de determinación en el caso de que el error de cuantificación  $E_{s-net}$  no sea menor que un umbral predefinido  $E_{th}$ . En el caso de que la verificación en la etapa 302 sea afirmativa, el método 400 continúa a la cuantificación predictiva del uno o más parámetros de la trama de audio (etapa 221) y adicionalmente a la derivación del error de cuantificación  $E_{pred}$  que resulta en una cuantificación no predictiva del uno o más parámetros de audio de la trama de audio actual (etapa 322). En consecuencia, el procesamiento requerido para la cuantificación predictiva (etapa 212) y la derivación del error de cuantificación  $E_{pred}$  (etapa 322) puede omitirse en el caso de que no sean necesarios para ahorrar recursos computacionales.

50 En una variación del método 400, las etapas 221 y 322 pueden llevarse a cabo en paralelo a las etapas 211 y 312 antes de continuar a la etapa 302. En esta variación, en el caso de que la verificación de la etapa 302 sea afirmativa, el método 400 continúa la etapa 213, mientras que en el caso de que la verificación de la etapa 302 no sea afirmativa, el método 400 continúa a la etapa 303.

55 Junto con las líneas anteriormente descritas para el error de cuantificación  $E_1$ , también en el contexto del método 400 se aplican las consideraciones con respecto al umbral  $E_{th}$  proporcionado en el contexto del método 200: un valor apropiado para el umbral  $E_{th}$  es diferente para diferentes parámetros de audio y posibles diferentes funciones de ponderación aplicadas para ponderar el error de cuantificación, y tiene que calcularse por ensayo y error fuera de línea, y, como un ejemplo, el umbral  $E_{th}$  puede establecerse a un valor que corresponde a una SD en el intervalo de 0,8 a 1,0 dB, por ejemplo 0,9 dB.

60 El método 400 puede comprender, opcionalmente, una o más etapas de determinación adicionales para evaluar una o más reglas de selección respectivas que pueden provocar la selección de la cuantificación no predictiva. Como un ejemplo, tal etapa o etapas de determinación pueden proporcionarse antes o después de la etapa 302.

La Figura 5 representa un diagrama de flujo que ilustra la operación en el codificador 121 de audio como etapas de un método 500 a modo de ejemplo. El método 500 se proporciona como una variación del método 400 y sirve como otra realización de ejemplo en la estructura anteriormente descrita con referencias al método 200. En el método 500, las etapas 314 y 324 del método 400 se sustituyen con las respectivas etapas 414 y 424, mientras que el método 500 comparte todas las etapas restantes del método 400. Aunque se describe en el presente documento como una modificación del método 400, la modificación similar puede aplicarse al método 300 también.

En el método 500, en el caso de que el componente 125 de cuantificación haya seleccionado la cuantificación no predictiva para el uno o más parámetros de audio en la trama de audio actual  $i$ , el componente 125 de cuantificación puede resetear adicionalmente el factor de cambio de escala adaptativo  $m$  para uso mediante el componente 125 de cuantificación en la siguiente trama de audio  $i+1$  estableciendo el factor de cambio de escala adaptativo  $m$  a un valor inicial  $m_0$  (como se ha descrito anteriormente en el contexto de la etapa 314) y reseteando adicionalmente un contador indicativo de la longitud de racha de predicción actual  $L$  a cero (etapa 414).

En contraste, en el caso de que el componente 125 de cuantificación haya seleccionado la cuantificación predictiva para el uno o más parámetros de audio en la trama de audio  $i$ , el componente 125 de cuantificación puede aumentar adicionalmente el contador indicativo de la longitud de racha de predicción actual  $L$  en uno y, posteriormente, ajustar el factor de cambio de escala adaptativo  $m$  para uso mediante el componente 125 de cuantificación en la siguiente trama  $i+1$  multiplicando el factor de cambio de escala  $m$  por un factor de cambio de escala predefinido  $m_s$  (como se ha descrito anteriormente en el contexto de la etapa 324) con la condición de que la longitud de racha de predicción actual  $L$  supere el umbral  $L_0$  (etapa 424). Por lo tanto, el factor de cambio de escala adaptativo  $m$  se mantiene en el valor inicial  $m_0$  hasta que la longitud de racha de predicción actual  $L$  supere el umbral  $L_0$ , mientras que la adaptación del factor de cambio de escala adaptativo  $m$  por el factor de cambio de escala  $m_s$  tenga lugar para cada trama de la longitud de racha de predicción en exceso del umbral  $L_0$ .

En el contexto de los métodos 300, 400 y 500 a modo de ejemplo anteriormente descritos la adaptación del factor de cambio de escala adaptativo  $m$  se describe para que tenga lugar reseteando el factor de cambio de escala  $m$  al valor inicial  $m_0$  (etapas 314, 414) y ajustando el factor de cambio de escala  $m$  a un nuevo valor (etapas 324, 424) para procesamiento de la siguiente trama de audio en el componente 125 de cuantificación.

Como un enfoque alternativo en este sentido, en cada uno de los métodos 300, 400 y 500, las etapas de reseteo y ajuste anteriormente mencionadas pueden omitirse y el valor del factor de cambio de escala adaptativo  $m$  puede derivarse basándose en la longitud de racha de predicción actual  $L$ . Para este fin, uno respectivo de los métodos 300, 400 puede implicar adicionalmente el rastreo del valor actual de la longitud de racha de predicción  $L$ , por ejemplo como se describe en este aspecto en las etapas 414 y 424 del método 500.

Como un ejemplo en este sentido, el factor de cambio de escala adaptativo  $m$  puede calcularse basándose en la longitud de racha de predicción  $L$ , por ejemplo de acuerdo con la ecuación (5a), o basándose en la longitud de racha de predicción  $L$  y el umbral predefinido  $L_0$ , por ejemplo de acuerdo con la ecuación (5b).

$$m = m_0 m_s^L \tag{5a}$$

$$\begin{cases} m = m_0 & \text{si } L \leq L_0 \\ m = m_0 m_s^{(L-L_0)} & \text{de otra manera} \end{cases} \tag{5b}$$

Como otro ejemplo en este sentido, el factor de cambio de escala adaptativo  $m$  puede obtenerse indexando una tabla accesible mediante el componente 125 de cuantificación. Tal tabla puede estar dispuesta para almacenar el valor respectivo del factor de cambio de escala adaptativo  $m$  para cada valor en un intervalo predefinido de valores de  $L$ , por ejemplo de 0 a  $L_{\max}$ , donde  $L_{\max}$  es la longitud máxima considerada (o permitida) de la longitud de racha predictiva  $L$ . El cálculo del factor de cambio de escala adaptativo  $m$  o el acceso a la tabla para encontrar el valor del factor de cambio de escala adaptativo  $m$  puede proporcionarse por ejemplo como una etapa adicional que precede a la etapa 303 (en los métodos 300, 400, 500) o que precede a la etapa 302 (en los métodos 400, 500).

Las tramas de audio cuantificadas proporcionadas pueden transmitirse mediante el transmisor 112 como una parte de datos de audio codificados en un flujo de bits junto con información adicional, por ejemplo junto con una indicación de la cuantificación empleada. Como alternativa, las tramas de audio cuantificadas y la posible indicación de la cuantificación empleada pueden almacenarse en una memoria en el dispositivo 100 electrónico para decodificación posterior y/o transmisión posterior mediante el transmisor 112.

En el dispositivo 150 electrónico, el flujo de bits se recibe mediante el componente 162 de recepción y se proporciona al decodificador 171. En el decodificador 171, el componente 174 de sintetización construye una señal de audio

5 sintetizada basándose en los parámetros cuantificados en el flujo de bits recibido. La señal de audio reconstruida puede a continuación proporcionarse al componente 161 de salida de audio, posiblemente después de algún procesamiento adicional, como una conversión de digital a analógico.

- 5 Los bloques de las Figuras 2 a 5 podrían entenderse también como bloques de procesamiento separados representados esquemáticamente del componente 125 de cuantificación.

10 La Figura 6 es un diagrama de bloques esquemático de un dispositivo electrónico 600 a modo de ejemplo, en el que puede implementarse en software una selección de una cuantificación predictiva o no predictiva de acuerdo con una realización de la invención.

15 El dispositivo 600 electrónico puede ser por ejemplo un teléfono móvil. Comprende un procesador 630 y enlazado a este procesador 630 un componente 611 de entrada de audio, un componente 661 de salida de audio, un transceptor (RX/TX) 612 y una memoria 640. Se ha de entender que las conexiones indicadas del dispositivo 600 electrónico pueden realizarse mediante diversos otros elementos no mostrados.

20 El componente 611 de entrada de audio puede ser por ejemplo un micrófono, un conjunto de micrófonos o una interfaz a una fuente de audio. El componente 661 de salida de audio puede ser por ejemplo un altavoz. La memoria 640 comprende una sección 641 para almacenar código de programa informático y una sección 642 para almacenar datos. El código de programa informático almacenado comprende código para codificar señales de audio usando una cuantificación escalable y posiblemente también código para decodificar señales de audio. El procesador 630 está configurado para ejecutar código de programa informático disponible. Siempre que el código disponible se almacene en la memoria 640, el procesador 630 puede recuperar el código para este fin desde la sección 641 de la memoria 640 cada vez que se requiera. Se ha de entender que puede estar disponible también diverso otro código de programa informático para ejecución, como un código de programa operativo y código de programa para diversas aplicaciones.

25 El código de codificación almacenado o el procesador 630 en combinación con la memoria 640 podría observarse también como un aparato a modo de ejemplo de acuerdo con una realización de la presente invención. La memoria 640 que almacena el código de codificación podría observarse como un producto de programa informático a modo de ejemplo de acuerdo con una realización de la presente invención.

30 Cuando un usuario o por ejemplo un proceso que se ejecuta en el dispositivo 600 electrónico selecciona una función del dispositivo 600 electrónico, que requiere una codificación de una señal de audio de entrada, una aplicación que proporciona esta función provoca que el procesador 630 recupere el código de codificación desde la memoria 640. Las señales de audio recibidas mediante el componente 611 de entrada de audio se proporcionan a continuación al procesador 630 - en el caso de las señales de audio analógicas recibidas después de una conversión a señales de audio digitales y posibles etapas de pre-procesamiento adicionales requeridas/aplicadas antes del aprovisionamiento de la señal de audio al procesador 630.

35 El procesador 630 ejecuta el código de codificación recuperado para codificar la señal de audio digital. La codificación puede corresponder a la codificación descrita anteriormente para la Figura 1 con referencia a una de las Figuras 2 a 5. El código de codificación puede por lo tanto observarse como un código de programa informático que provoca realizar, por ejemplo, la codificación descrita anteriormente para la Figura 1 con referencia a una de las Figuras 2 a 5 cuando el código de programa informático se ejecuta mediante el procesador 630 o mediante cualquier otro aparato informático. La señal de audio codificada se almacena en la porción 642 de almacenamiento de datos de la memoria 640 para uso posterior o se transmite mediante el transceptor 612 a otro dispositivo electrónico.

40 El procesador 630 puede recuperar adicionalmente el código de decodificación desde la memoria 640 y ejecutarlo para decodificar una señal de audio codificada que se recibe mediante el transceptor 612 o se recupera desde la porción 642 de almacenamiento de datos de la memoria 640. La decodificación puede corresponder a la decodificación descrita anteriormente para la Figura 1. La señal de audio digital decodificada puede proporcionarse a continuación al componente 661 de salida de audio. En el caso de que el componente 661 de salida de audio comprenda un altavoz, la señal de audio decodificada puede presentarse por ejemplo a un usuario mediante el altavoz después de una conversión en una señal de audio analógica y posibles etapas de post-procesamiento adicionales. Como alternativa, la señal de audio digital decodificada podría almacenarse en la porción 642 de almacenamiento de datos de la memoria 640.

45 Las funciones ilustradas mediante el componente 125 de cuantificación de la Figura 1 o las funciones ilustradas mediante el procesador 630 que ejecuta el código 641 de programa de la Figura 6 pueden observarse también como medios para derivar un primer error de cuantificación que es descriptivo de un error que resulta en una cuantificación no predictiva de un parámetro de audio de un segmento de señal de audio, medios para derivar un segundo error de cuantificación que es descriptivo de un error que resulta en una cuantificación predictiva de dicho parámetro de audio de dicho segmento de señal de audio, medios para determinar si dicho segundo error de cuantificación supera dicho primer error de cuantificación en al menos un margen adaptativo que es dependiente del número de segmentos de señal de audio consecutivos que preceden dicho segmento de señal de audio en el que dicho parámetro de audio se proporciona cuantificado con dicha cuantificación predictiva, medios para proporcionar dicho parámetro de audio de

dicho segmento de audio cuantificado con dicha cuantificación no predictiva como parte de una señal de audio codificada al menos en el caso de que el resultado de dicha determinación sea afirmativo y medios para proporcionar de otra manera dicho parámetro de audio de dicho segmento de audio cuantificado con dicha cuantificación predictiva como parte de una señal de audio codificada. Los códigos 641 de programa pueden observarse también como que  
5 comprenden tales medios en forma de módulos funcionales o componentes de código.

Aunque se han mostrado y descrito y señalado características novedosas fundamentales de la invención según se aplican a realizaciones preferidas de la misma, se entenderá que pueden realizarse por los expertos en la materia  
10 diversas omisiones y sustituciones y cambios en la forma y detalles de los dispositivos y métodos descritos sin alejarse de la presente invención. Por ejemplo, se pretende de manera expresa que todas las combinaciones de estos elementos y/o etapas de método que realizan sustancialmente la misma función de sustancialmente la misma manera para conseguir los mismos resultados estén dentro del alcance de la invención. Además, debería reconocerse que las estructuras y/o elementos y/o etapas de método mostrados y/o descritos en relación con cualquier forma desvelada o  
15 realización de la invención pueden incorporarse en cualquier otra forma desvelada o descrita o sugerida o realización como una materia general de elección de diseño. Es la intención, por lo tanto, que esté únicamente como se indica por el alcance de las reivindicaciones anexadas a la misma. Adicionalmente, en las reivindicaciones los artículos medios-más-función se pretende que cubran las estructuras descritas en el presente documento como que realizan la función indicada y no únicamente equivalentes estructurales, sino también estructuras equivalentes.

**REIVINDICACIONES**

1. Un método para codificar una señal de audio procesando una secuencia de segmentos de señal de audio, comprendiendo el método
  - 5 derivar un primer error de cuantificación que es descriptivo de un error que resulta en una cuantificación no predictiva de un parámetro de audio de un segmento de señal de audio;
  - derivar un segundo error de cuantificación que es descriptivo de un error que resulta en una cuantificación predictiva de dicho parámetro de audio de dicho segmento de señal de audio;
  - 10 determinar si dicho segundo error de cuantificación supera dicho primer error de cuantificación en al menos un margen adaptativo que es dependiente del número de segmentos de señal de audio consecutivos que preceden dicho segmento de señal de audio en el que dicho parámetro de audio se proporciona cuantificado con dicha cuantificación predictiva;
  - 15 proporcionar dicho parámetro de audio de dicho segmento de audio cuantificado con dicha cuantificación no predictiva como parte de una señal de audio codificada al menos en el caso de que el resultado de dicha determinación sea afirmativo; y
  - proporcionar de otra manera dicho parámetro de audio de dicho segmento de audio cuantificado con dicha cuantificación predictiva como parte de una señal de audio codificada.
- 20 2. Un método de acuerdo con la reivindicación 1, en el que dicho margen adaptativo se aumenta desde su valor inicial predefinido en una cantidad predefinida para cada segmento de señal de audio entre dicho segmento de señal de audio y el segmento de señal de audio precedente más cercano en el que dicho parámetro de audio se proporciona cuantificado con dicha cuantificación predictiva.
- 25 3. Un método de acuerdo con la reivindicación 1, en el que dicho margen adaptativo se aumenta desde su valor inicial predefinido en una cantidad predefinida para cada segmento de señal de audio en exceso de un umbral predefinido entre dicho segmento de señal de audio y el segmento de señal de audio precedente más cercano en el que dicho parámetro de audio se proporciona cuantificado con dicha cuantificación no predictiva.
- 30 4. Un método de acuerdo con las reivindicaciones 2 o 3, en el que dicho valor inicial predefinido del margen es cero o sustancialmente cero.
5. Un método de acuerdo con la reivindicación 1, en el que dicha determinación comprende determinar si dicho primer error de cuantificación multiplicado por un factor de cambio de escala adaptativo es menor que dicho segundo error de cuantificación, factor de cambio de escala adaptativo que representa el margen adaptativo para dicho segmento de señal de audio.
- 35 6. Un método de acuerdo con la reivindicación 5, que comprende adicionalmente reducir dicho factor de cambio de escala en una cantidad predeterminada en el caso de que dicho parámetro de audio de dicho segmento de audio se proporcione cuantificado con dicha cuantificación predictiva.
- 40 7. Un método de acuerdo con la reivindicación 5, que comprende adicionalmente reducir dicho factor de cambio de escala en una cantidad predeterminada en el caso de que dicho parámetro de audio de dicho segmento de audio se proporcione cuantificado con dicha cuantificación predictiva, y
- 45 dicho número de segmentos de señal de audio consecutivos supere un umbral predefinido.
8. Un método de acuerdo con cualquiera de las reivindicaciones 5 a 7, que comprende adicionalmente resetear dicho factor de cambio de escala a un valor inicial predefinido en el caso de que dicho parámetro de audio de dicho segmento de audio se proporcione cuantificado con dicha cuantificación no predictiva.
- 50 9. Un método de acuerdo con la reivindicación 8, en el que dicho valor inicial predefinido es uno.
10. Un método de acuerdo con las reivindicaciones 3 o 7, en el que dicho umbral predefinido es tres.
- 55 11. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 10, en el que dicho parámetro de audio comprende uno de un vector de Frecuencia Espectral de Inmitancia y un vector de Frecuencia Espectral Lineal que son representativos de características espectrales de dicho segmento de audio.
- 60 12. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 11, en el que dicho primer error de cuantificación se obtiene combinando errores ponderados entre un componente de dicho parámetro de audio y un componente correspondiente de dicho parámetro de audio que resulta con dicha cuantificación no predictiva, y
- 65 dicho segundo error de cuantificación se obtiene combinando errores ponderados entre un componente de dicho parámetro de audio y un componente correspondiente de dicho parámetro de audio que resulta con dicha

cuantificación predictiva.

13. Un aparato para codificar una señal de audio procesando una secuencia de segmentos de señal de audio, configurado el aparato para:

- 5 derivar un primer error de cuantificación que es descriptivo de un error que resulta con una cuantificación no predictiva de un parámetro de audio en un segmento de señal de audio;  
derivar un segundo error de cuantificación que es descriptivo de un error que resulta con una cuantificación predictiva de dicho parámetro de audio en dicho segmento de señal de audio;  
10 determinar si dicho segundo error de cuantificación excede dicho primer error de cuantificación en al menos un margen adaptativo que es dependiente del número de segmentos de señales de audio consecutivos que preceden dicho segmento de señal de audio proporcionándose dicho parámetro de audio cuantificado con dicha cuantificación predictiva;  
15 proporcionar dicho parámetro de audio de dicho segmento de audio cuantificado de dicha cuantificación no predictiva como parte de una señal de audio al menos en caso de que el resultado de dicha determinación sea afirmativa; y  
proporcionar de otra manera dicho parámetro de audio de dicho segmento de audio cuantificado con dicha cuantificación predictiva como parte de una señal de audio.

20 14. El aparato según la reivindicación 13, en donde el aparato se configura adicionalmente para realizar el método de acuerdo con cualquiera de las reivindicaciones 2 a 12.

25 15. Un programa informático que comprende un código de programa legible por ordenador configurado para provocar la realización del método de cualquiera de las reivindicaciones 1 a 12 cuando dicho código de programa se ejecuta en un aparato informático.

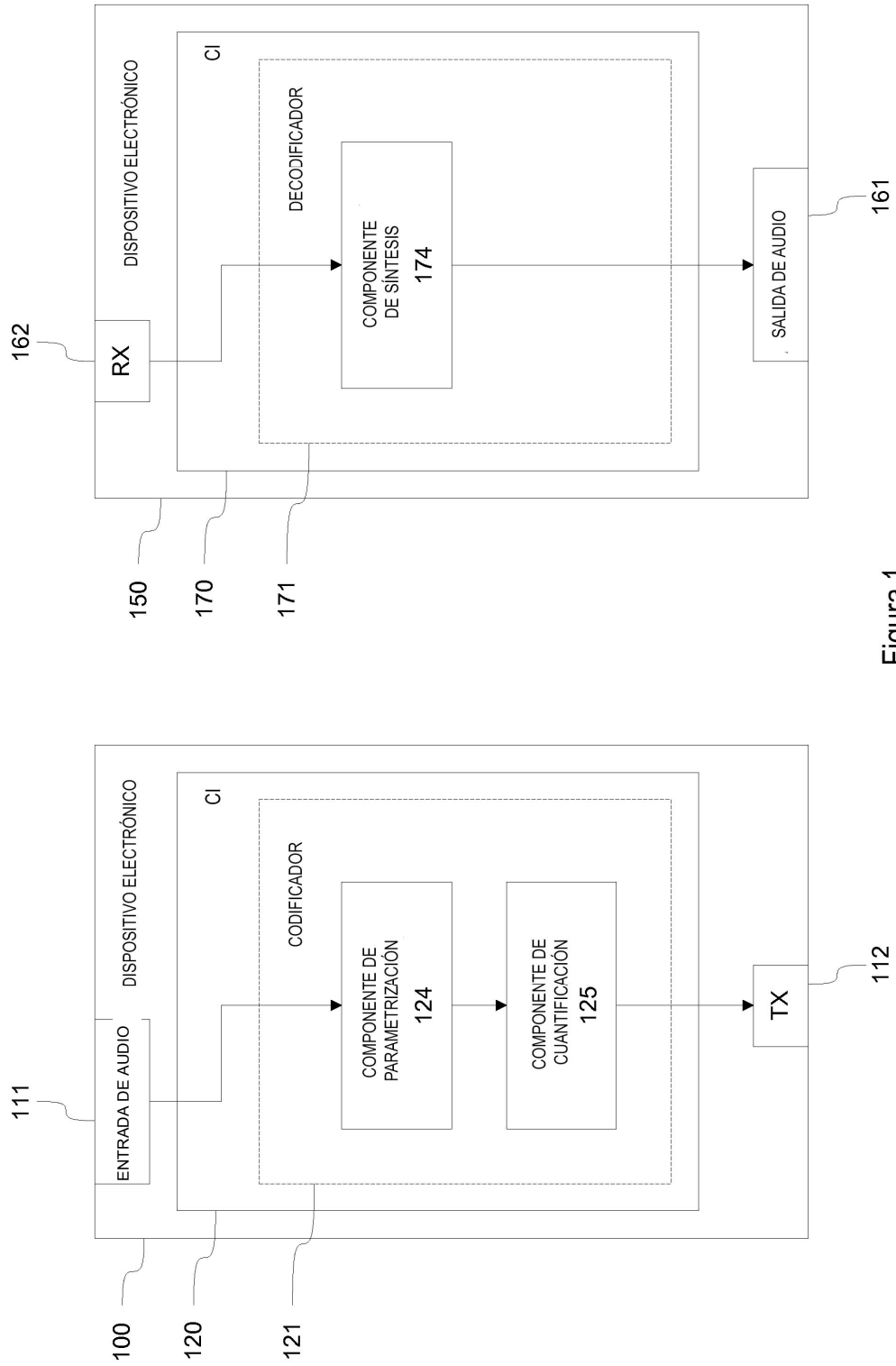


Figura 1

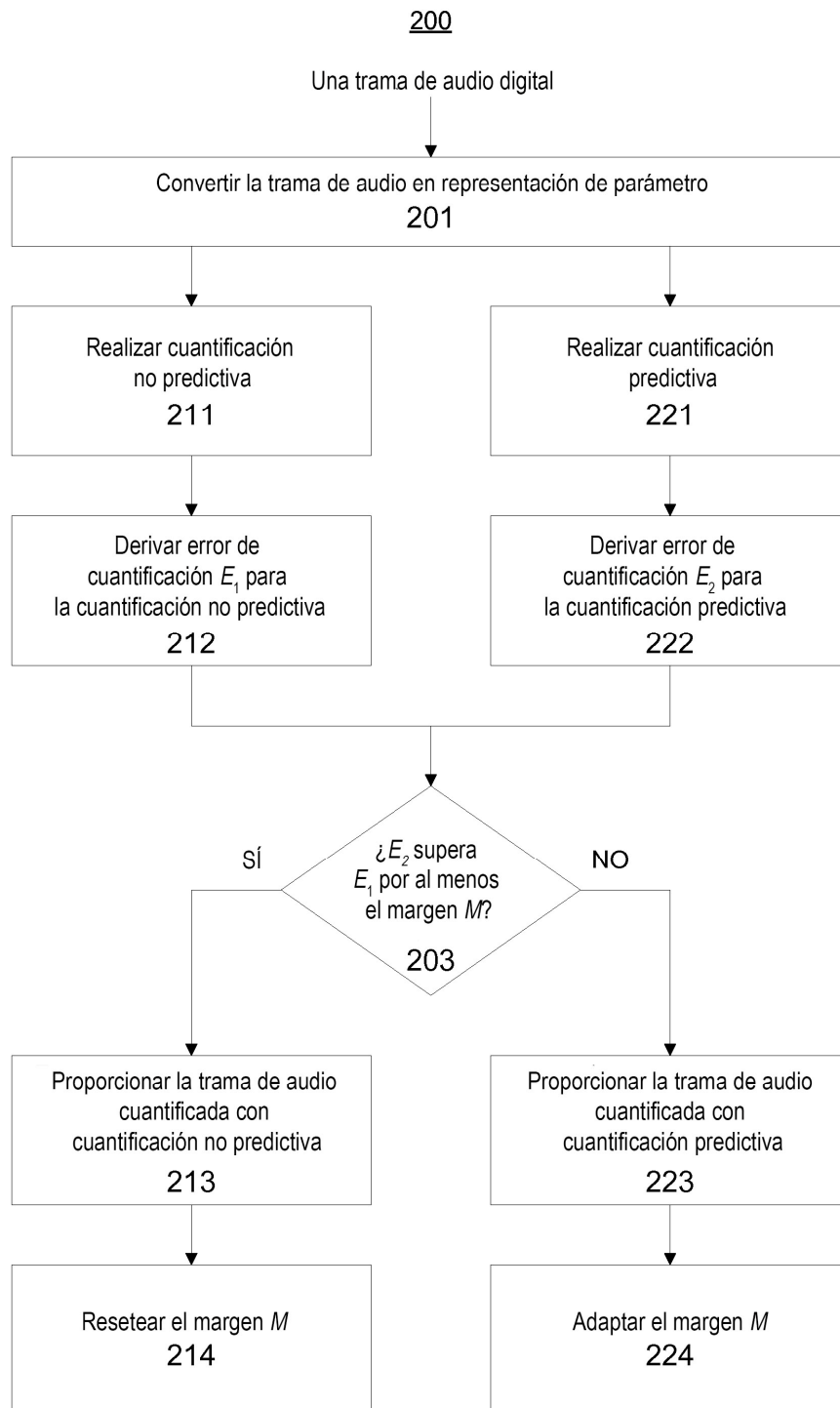


Figura 2



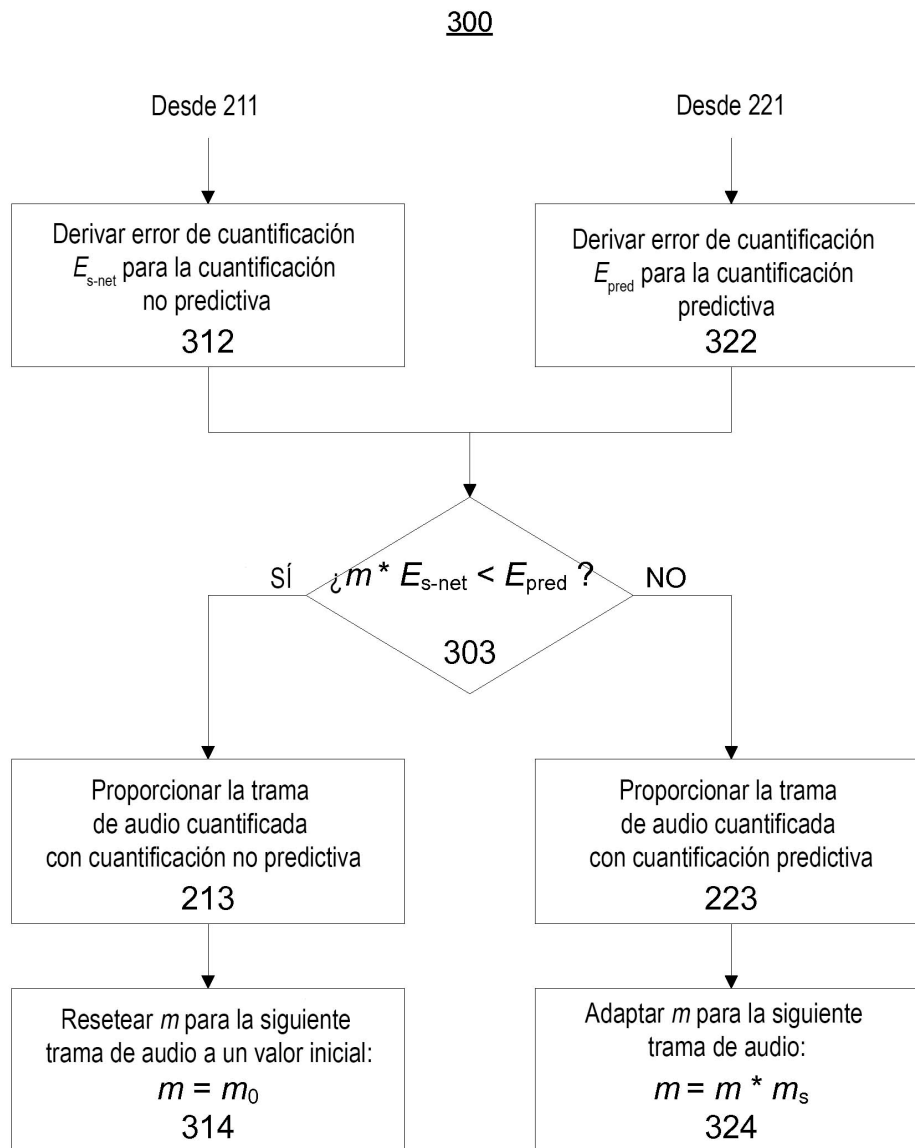


Figura 3

400

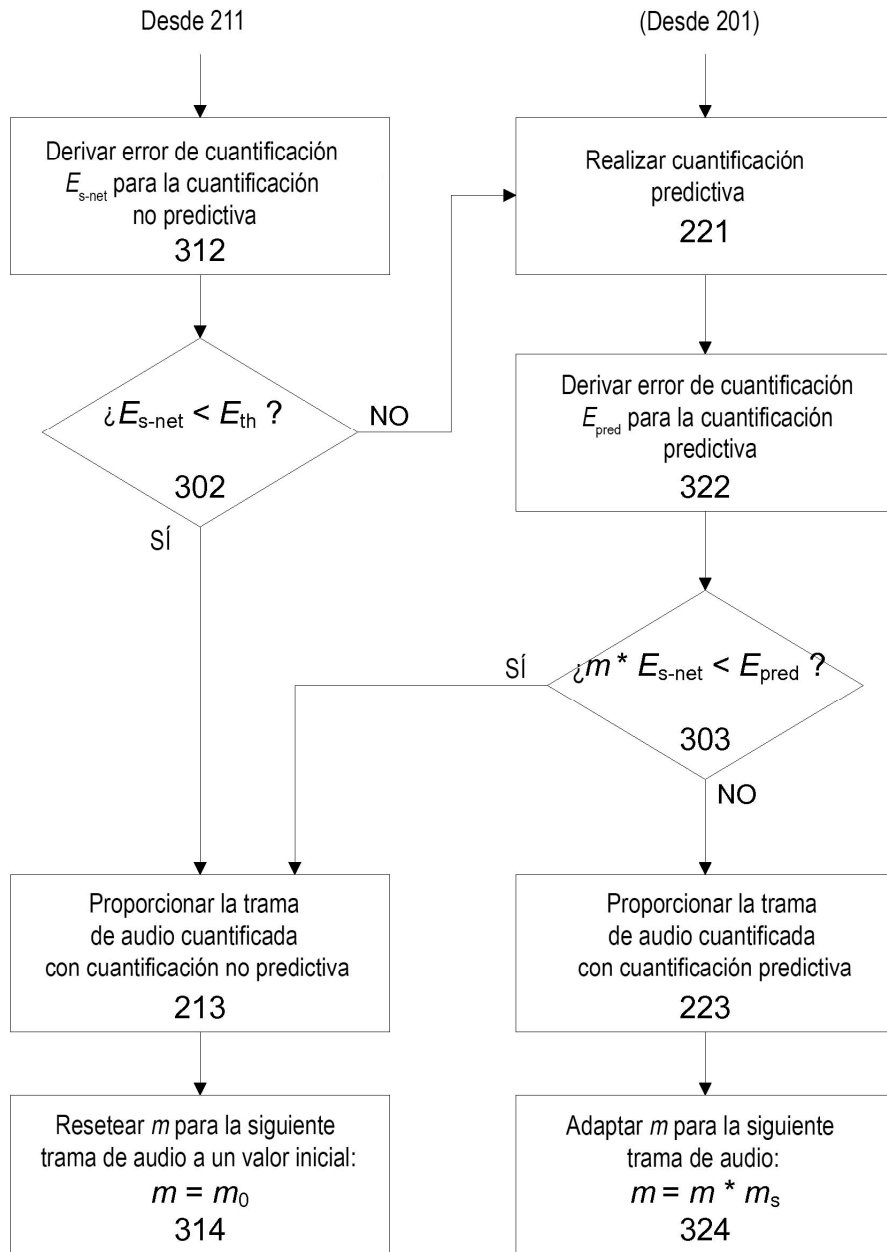


Figura 4

500

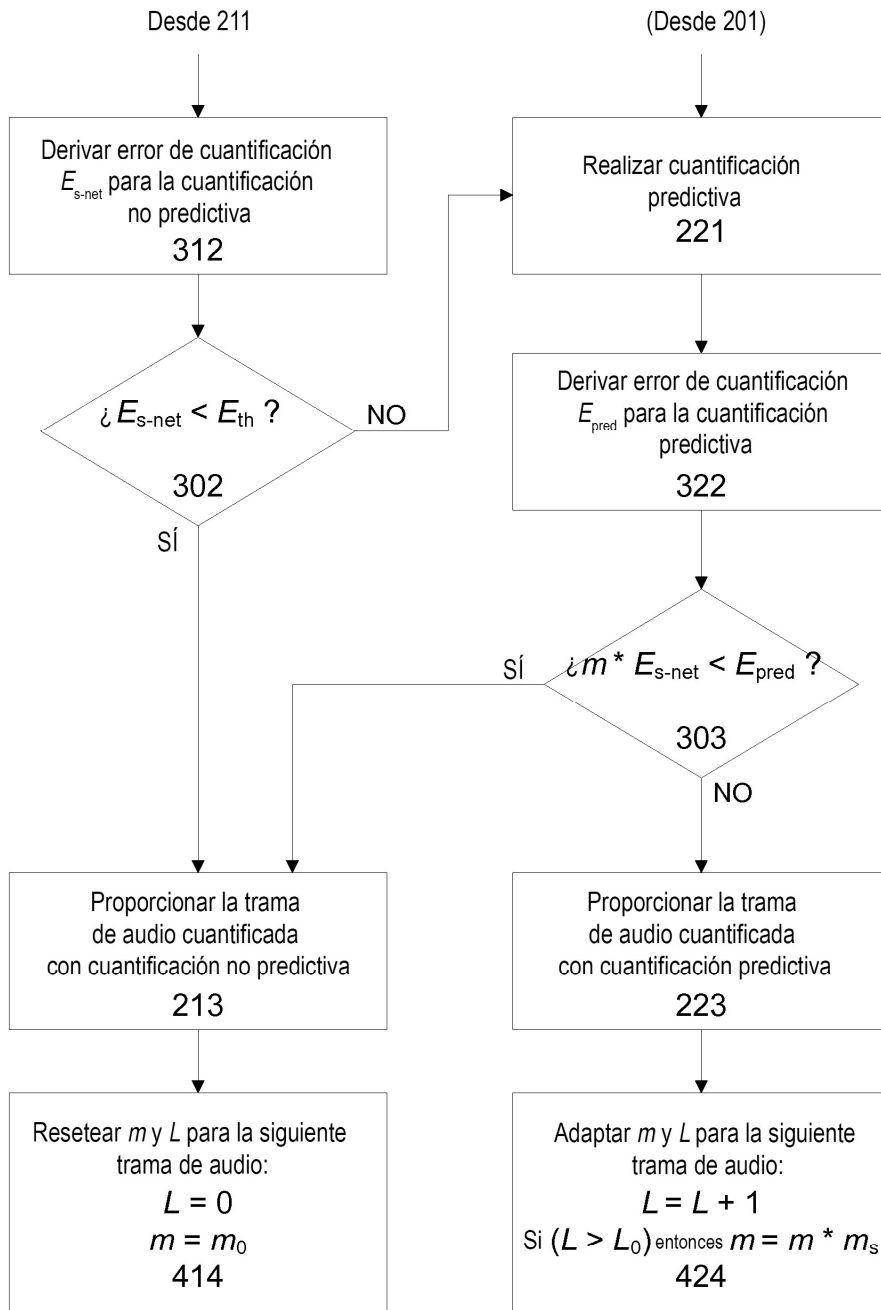


Figura 5

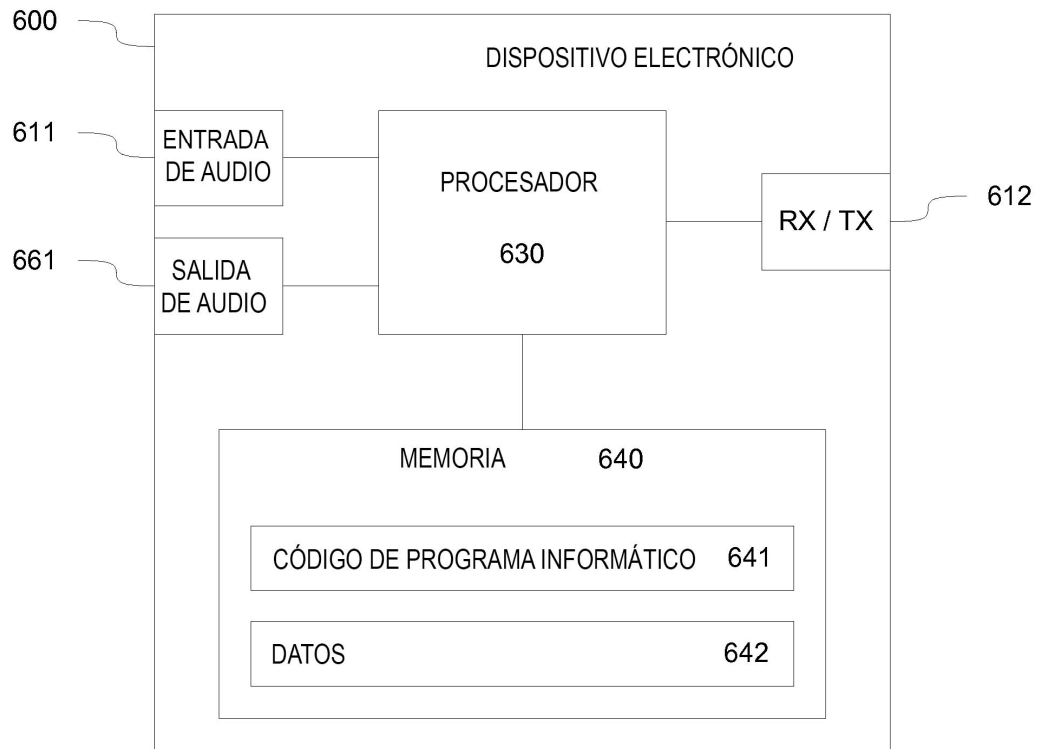


Figura 6