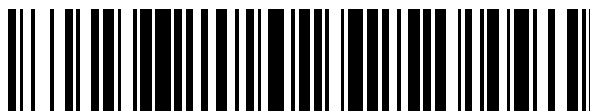


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 685 574**

51 Int. Cl.:

G10L 19/025 (2013.01)

G10L 19/028 (2013.01)

G10L 19/12 (2013.01)

G10L 19/22 (2013.01)

G10L 25/90 (2013.01)

G10L 25/21 (2013.01)

G10L 19/26 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.07.2015 PCT/EP2015/067160**

87 Fecha y número de publicación internacional: **04.02.2016 WO16016190**

96 Fecha de presentación y número de la solicitud europea: **27.07.2015 E 15744175 (9)**

97 Fecha y número de publicación de la concesión europea: **27.06.2018 EP 3175455**

54 Título: **Control dependiente de la armonicidad de una herramienta de filtro de armónicos**

30 Prioridad:

28.07.2014 EP 14178810

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

10.10.2018

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80636 Munich, DE**

72 Inventor/es:

**MARKOVIC, GORAN;
HELMRICH, CHRISTIAN;
RAVELLI, EMMANUEL;
JANDER, MANUEL y
DÖHLA, STEFAN**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 685 574 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).

Control dependiente de la armonicidad de una herramienta de filtro de armónicos

DESCRIPCIÓN

5 La presente solicitud se refiere a la decisión sobre el control de una herramienta de filtro de armónicos como el enfoque pre/post filtro o post-filtro solamente. Esta herramienta es, por ejemplo, aplicable a codificación de audio y voz unificada (USAC) MPEG-D y al próximo códec 3GPP EVS.

10 Los códec de audio basados en transformada como AAC, MP3, o TCX en general introducen ruido de cuantización inter-armónico cuando procesan señales de audio armónicas, particularmente a velocidades de bits bajas.

El efecto es aún peor cuando el códec de audio basado en transformada opera con bajo retardo, debido a la peor resolución de frecuencia y/o selectividad introducida por un tamaño de transformada más corto y/o una respuesta de frecuencia de ventana inferior.

15 Este ruido inter-armónico en general es percibido como un elemento “de gorjeo” muy molesto, que reduce significativamente el rendimiento del códec de audio basado en transformada al evaluarlo subjetivamente en un material de audio altamente tonal como música o voz.

20 Una solución común a este problema es emplear técnicas basadas en predicciones, preferentemente una predicción usando un modelo auto-regresivo (AR) sobre la adición o sustracción de muestras ingresadas o decodificadas basadas, ya sea en el dominio de transformada o en el dominio de tiempo.

25 Sin embargo, el uso de estas técnicas en las señales con la estructura temporal cambiante conduce nuevamente a efectos no deseados como corrimiento temporal de eventos musicales de percusión o discursos fuertes o aún vestigios de impulsos debido a la repetición de un transitorio tipo impulso simple. Por ello debe tenerse especial cuidado en aquellas señales que contienen componentes transitorios y armónicos donde existe ambigüedad entre los transitorios y trenes de pulsos (los últimos se refieren a una señal armónica compuesta por pulsos individuales de muy corta duración; estas señales son conocidas como trenes de pulsos).

30 Existen varias soluciones para mejorar la calidad subjetiva de los códec de audio basados en transformada en señales de audio armónicas. Todas ellas exploran la periodicidad de largo plazo (tono) de formas de onda muy armónicas, estacionarias, y se basan en técnicas de predicción, ya sea en el dominio de transformada o el dominio de tiempo. La mayoría de las soluciones son conocidas como predicción de largo plazo (LTP) o predicción de tono, caracterizadas por un par de filtros que se aplican a la señal: un pre-filtro en el codificador (usualmente como un primer paso en el dominio de tiempo o frecuencia) y un post-filtro en el decodificador (usualmente como un último paso en el dominio de tiempo o frecuencia). Unas otras pocas soluciones, sin embargo, aplican solamente un proceso de post-filtrado simple en el lado del decodificador generalmente conocido como post-filtro de armónicos o post-filtro de bajos. Todos estos enfoques independientemente de ser pares de pre- y post-filtros o post-filtros solamente, se denominan herramienta de filtro de armónicos.

Ejemplos de dominios de transformada son:

45 [1] H. Fuchs, “Improving MPEG Audio Coding by Backward Adaptive Linear Stereo Prediction”, 99th AES Convention, Nueva York, 1995, Preprint 4086.

[2] L. Yin, M. Suonio, M. Väänänen, “A New Backward Predictor for MPEG Audio Coding”, 103rd AES Convention, Nueva York, 1997, Preprint 4521.

50 [3] Juha Ojanperä, Mauri Väänänen, Lin Yin, “Long Term Predictor for Transform Domain Perceptual Audio Coding”, 107th AES Convention, Nueva York, 1999, Preprint 5036.

Ejemplos de enfoques por dominio de tiempo que aplican pre- y post-filtrado son:

[4] Philip J. Wilson, Harprit Chhatwal, “Adaptive transform coder having long term predictor”, Patente US 5.012.517, 30 de abril de 1991.

55 [5] Jeongook Song, Chang-Heon Lee, Hyen-O Oh, Hong-Goo Kang, “Harmonic Enhancement in Low Bitrate Audio Coding Using an Efficient Long-Term Predictor”, EURASIP Journal on Advances in Signal Processing, August 2010.

[6] Juin-Hwey Chen, “Pitch-based pre-filtering and post-filtering for compression of audio signals”, Patente US 8.738.385, 27 de mayo de 2014.

60 [7] Jean-Marc Valin, Koen Vos, Timothy B. Terriberry, “Definition of the Opus Audio Codec”, ISSN: 2070-1721, IETF RFC 6716, Septiembre 2012.

[8] Rakesh Taori, Robert J. Sluijter, Eric Kathmann “Transmission System with Speech Encoder with Improved Pitch Detection”, Patente US 5.963.895, 5 de octubre de 1999.

Ejemplos de enfoques por dominio de tiempo que aplican solamente post-filtrado son:

[9] Juin-Hwey Chen, Allen Gersho, "Adaptive Postfiltering for Quality Enhancement of Coded Speech", IEEE Trans. on Speech and Audio Proc., vol. 3, Enero 1995.

[10] Int. Telecommunication Union, "Frame error robust variable bit-rate coding of speech and audio from 8–32 kbit/s", Recomendación ITU-T G.718, Junio 2008. www.itu.int/rec/T-REC-G.718/e, sección 7.4.1.

[11] Int. Telecommunication Union, "Coding of speech at 8 kbit/s using conjugate structure algebraic CELP (CS-ACELP)", Recomendación ITU-T G.729, Junio 2012. www.itu.int/rec/T-REC-G.729/e, sección 4.2.1.

[12] Bruno Bessette et al., "Method and device for frequency-selective pitch enhancement of synthesized speech", Patente US 7.529.660, 30 de mayo de 2003.

Un ejemplo de detector de transitorio es:

[13] Johannes Hilpert et al., "Method and Device for Detecting a Transient in a Discrete-Time Audio Signal", Patente US 6.826.525, 30 de noviembre de 2004.

Literatura relevante en psicoacústica:

[14] Hugo Fastl, Eberhard Zwicker, "Psychoacoustics: Facts and Models", 3a Edición, Springer, 14 de diciembre de 2006.

[15] Christoph Markus, "Background Noise Estimation", Patente europea EP 2.226.794, 6 de marzo de 2009.

Todas las técnicas descritas deciden cuando activar el filtro de predicción en base a una decisión de umbral único (por ejemplo, ganancia de predicción [5] o ganancia de tono [4] o armonicidad que es básicamente proporcional a la correlación normalizada [6]). Además, OPUS [7] emplea histéresis que incrementa el umbral si el tono está cambiando y reduce el umbral si la ganancia en el cuadro previo estaba por arriba de un umbral fijo predefinido. OPUS [7] además desactiva el predictor (tono) de largo plazo si se detecta un transitorio en algunas configuraciones específicas de cuadros. La razón de este diseño parece apartarse de la creencia general de que, en una mezcla de componentes de señal armónicos y transitorios, el transitorio domina la mezcla, y la activación de la LTP o predicción de tono causarían, según se ha expuesto, subjetivamente más inconvenientes que mejoras. Sin embargo, para algunas mezclas de formas de onda que serán expuestas en adelante, la activación del predictor de largo plazo o tono en cuadros de audio transitorios significativamente aumenta la calidad o eficacia de codificación y por ello es beneficiosa. Además, puede ser conveniente, al activar el predictor, variar su intensidad en base a características de señales instantáneas además de una ganancia de predicción, el único enfoque en el estado de la técnica.

En el documento US 8095359 B2 se describe un códec de audio perceptual que hace uso de bancos de filtros y MDCT con el fin de lograr una representación compacta de la señal de audio, eliminando la redundancia y la irrelevancia de la señal de audio original. Durante partes cuasi-estacionarias de la señal de audio, una resolución de alta frecuencia del banco de filtros es ventajosa para lograr una alta ganancia de codificación, pero esta resolución de alta frecuencia está acoplada a una resolución temporal en bruto que se convierte en un problema durante las partes transitorias de señal produciendo efectos de eco previo audibles. Teniendo esto en cuenta, se consigue una calidad de codificación/descodificación mejorada aplicando sobre la salida de un primer banco de filtros un segundo banco de filtros no uniforme, es decir, una MDCT en cascada. En particular, se usa la conmutación a un banco de filtro de extensión adicional (o banco de filtros de resolución múltiple) para volver a agrupar la representación de frecuencia de tiempo durante las secciones de señal de audio transitorias o de cambio rápido. Aplicando un control de conmutación correspondiente, se evitan los efectos de eco previo y se logran una alta ganancia de codificación y un bajo retardo de codificación.

En consecuencia, constituye un objeto de la presente invención proveer un concepto para el control dependiente de la armonicidad de una herramienta de filtro de armónicos de un códec de audio que mejora la eficacia de codificación, por ejemplo, ganancia de codificación objetiva mejorada o mejor calidad perceptual, o similar.

Este objeto se alcanza mediante lo reivindicado en las reivindicaciones independientes de la solicitud.

Constituye un descubrimiento básico de la presente solicitud que la eficacia de codificación de un códec de audio que usa una herramienta de filtro de armónicos controlable – conmutable o aún ajustable pueda mejorar realizando el control dependiente de la armonicidad de esta herramienta usando una medición de estructura temporal además de una medición de armonicidad con el fin de controlar la herramienta de filtro de armónicos. En particular, la estructura temporal de la señal de audio es evaluada de modo que dependa del tono. Esto permite lograr un control de la herramienta de filtro de armónicos adaptada a la situación de modo que cuando el control se realice solamente en base a la medición de armonicidad se decida contra o reduzca el uso de esta herramienta a pesar que el uso de la herramienta de filtro de armónicos, en tal situación, incrementaría la eficacia de codificación, mientras que en otras situaciones en las cuales la herramienta de filtro de armónicos pueda ser ineficaz o aún destructiva, el control reduce la aplicación de la herramienta de filtro de armónicos adecuadamente.

Implementaciones de la presente invención sobre la materia de las reivindicaciones subordinadas y sus realizaciones preferidas se describen a continuación con referencia a las figuras adjuntas, en las que

- La figura 1 ilustra un diagrama en bloque de un aparato para controlar una herramienta de filtro de armónicos en términos de ganancia de filtro de acuerdo con una realización;
- 5 La figura 2 ilustra un ejemplo de una condición predeterminada posible a cumplir para aplicar la herramienta de filtro de armónicos;
- La figura 3 ilustra un diagrama de flujo que ilustra una implementación posible de una lógica de decisión que, entre otras cosas, podría ser parametrizada para lograr el ejemplo de condición de la figura 2;
- 10 La figura 4 ilustra un diagrama en bloque de un aparato para ejecutar un control dependiente de la armonicidad (y medición temporal) de una herramienta de filtro de armónicos;
- La figura 5 ilustra un diagrama esquemático que ilustra la posición temporal de una región temporal para determinar la medición de estructura temporal de acuerdo con una realización;
- La figura 6 ilustra esquemáticamente un gráfico de muestras de energía que muestrean temporalmente la energía de la señal de audio dentro de la región temporal de acuerdo con una realización;
- 20 La figura 7 ilustra un diagrama en bloque que ilustra el uso del aparato de la figura 4 en un códec de audio que ilustra el codificador y el decodificador del códec de audio, respectivamente, cuando el codificador usa el aparato de la figura 4, de acuerdo con una realización donde se usa una herramienta de filtro pre/post-armónico;
- 25 La figura 8 ilustra un diagrama en bloque que ilustra el uso del aparato de la figura 4 en un códec de audio que ilustra el codificador y el decodificador del códec de audio, respectivamente, cuando el codificador usa el aparato de la figura 4, de acuerdo con una realización donde se usa una herramienta de post-filtro de armónicos;
- 30 La figura 9 ilustra un diagrama en bloque del controlador de la Figura 4 de acuerdo con una realización;
- La figura 10 ilustra un diagrama en bloque de un sistema que ilustra la posibilidad que el aparato de la figura 4 comparta el uso de las muestras de energía de la Figura 6 con un detector de transitorio;
- 35 La figura 11 ilustra un gráfico de una porción de dominio de tiempo (porción de la forma de onda) de una señal de audio como un ejemplo de una señal de bajo tono que además ilustra el posicionamiento de acuerdo con el tono de la región temporal para determinar la por lo menos única medición de estructura temporal;
- 40 La figura 12 ilustra un gráfico de una porción de dominio de tiempo de una señal de audio como un ejemplo de una señal de alto tono que además ilustra el posicionamiento de acuerdo con el tono de la región temporal para determinar la por lo menos única medición de estructura temporal;
- 45 La figura 13 ilustra un ejemplo de espectrograma de transitorio de impulso y paso dentro de una señal armónica;
- La figura 14 ilustra un ejemplo de espectrograma para ilustrar la influencia de la LTP sobre el transitorio de impulso y paso;
- 50 La figura 15 ilustra, una sobre la otra, las porciones de dominio de tiempo de la señal de audio que se ilustra en la figura 14, y su versión con filtro de paso bajo y alto, respectivamente, con el fin de ilustrar el control de acuerdo con las figuras 2, 3, 16 y 17 para el transitorio de impulso y paso;
- La figura 16 ilustra un gráfico de barras de un ejemplo para la secuencia temporal de energías de segmentos - secuencia de muestras de energía - para un transitorio de impulso y la disposición de la región temporal para determinar la por lo menos única medición de estructura temporal de acuerdo con las figuras 2 y 3;
- 55 La figura 17 ilustra un gráfico de barras de un ejemplo para la secuencia temporal de energías de segmentos - secuencia de muestras de energía - para un transitorio de paso y la disposición de la región temporal para determinar la por lo menos única medición de estructura temporal de acuerdo con las figuras 2 y 3;
- 60 La figura 18 ilustra un ejemplo de espectrograma de un tren de pulsos (extracto que usa un espectrograma FFT

corto);

La figura 19 ilustra un ejemplo de forma de onda del tren de pulsos;
 La figura 20 ilustra un espectrograma FFT corto original del tren de pulsos; y

5 La figura 21 ilustra un espectrograma FFT largo original del tren de pulsos.

10 La siguiente descripción comienza con una primera descripción detallada del control de una herramienta de filtro de armónicos. Se describen brevemente las ideas que condujeron a esta primera realización. Estas ideas, sin embargo, también se aplican a las demás realizaciones. Luego se presentan realizaciones generales seguidas por ejemplos concretos específicos de porciones de señales de audio con el fin de delinear de un modo más concreto los efectos que resultan de las realizaciones de la solicitud.

15 El mecanismo de decisión para activar o controlar una herramienta de filtro de armónicos de, por ejemplo, una técnica basada en la predicción se basa en una combinación de una medición de armonicidad como una correlación normalizada o ganancia de predicción y una medición de estructura temporal, por ejemplo una medición de planicidad temporal o cambio de energía.

20 La decisión puede, según lo expuesto, no depender solamente de la medición de armonicidad del cuadro actual, sino además del cuadro previo, y de una medición de estructura temporal del cuadro actual y opcionalmente el cuadro previo.

25 El esquema de decisión puede ser diseñado de modo tal que la técnica basada en la predicción se active también para transitorios, cada vez que su uso sea psicoacústicamente beneficioso según un modelo respectivo.

Los umbrales usados para activar la técnica basada en la predicción pueden ser, en una realización, dependientes del tono actual en lugar del cambio de tono.

30 El esquema de decisión permite, por ejemplo, evitar la repetición de un transitorio específico, pero permite la técnica basada en la predicción para algunos transitorios y para señales con estructuras temporales específicas donde un detector de transitorio normalmente señalaría bloques de transformada cortos (es decir, la existencia de uno o más transitorios).

35 La técnica de decisión presentada a continuación puede ser aplicada a cualquiera de los métodos basados en predicción antes descritos, ya sea en el dominio de transformada o en el dominio específico, ya sea en enfoques de pre-filtro más post-filtro o post-filtro solamente. Por otra parte, puede aplicarse a predictores que operen limitados por bandas (con paso bajo) o en sub-bandas (con características de paso de banda).

40 El objetivo general relativo a la activación de la LTP, predicción de tono, o post filtrado armónico es el logro de ambas condiciones, a saber:

- obtener un beneficio objetivo o subjetivo activando el filtro,
- la no introducción de artefactos significativos mediante la activación de dicho filtro.

45 La determinación del beneficio de usar el filtro usualmente se realiza mediante mediciones de auto correlación y/o ganancia de predicción en la señal dada y es muy conocida [1-7].

50 La medición de un beneficio subjetivo también es directa al menos para las señales estacionarias, dado que los datos de mejora perceptual obtenidos por los ensayos de escucha son típicamente proporcionales a las correspondientes mediciones objetivas, es decir, la ganancia de correlación y/o predicción antes mencionada.

55 La identificación o predicción de la existencia de artefactos mediante el filtrado, sin embargo, requiere técnicas más sofisticadas que simples comparaciones de mediciones objetivas como tipo de cuadros (transformadas largas para cuadros estacionarios vs. transformadas cortas para cuadros transitorios) o ganancia de predicción para ciertos umbrales, como se hace en la técnica. Esencialmente, con el fin de impedir artefactos debe asegurarse que los cambios provocados por el filtrado en la forma de onda objetivo no excedan significativamente un umbral de espectro-temporal de tiempo variable en cualquier punto en tiempo o frecuencia. El esquema de decisión de acuerdo con algunas de las realizaciones presentadas a continuación, en consecuencia, usa el siguiente esquema de decisión y control que comprende tres bloques algorítmicos a ejecutar en serie para cada cuadro de la señal de audio a codificar y/o someter al filtrado:

60 Un bloque de medición de armonicidad que calcula los datos del filtro de armónicos comúnmente usado como valores de correlación o ganancia normalizados (en adelante "ganancia de predicción"). Según se indicará nuevamente, la palabra "ganancia" es una generalización para cualquier parámetro comúnmente asociado con la

intensidad del filtro, por ejemplo, un factor de ganancia explícito o la magnitud absoluta o relativa de un conjunto de uno o más coeficientes de filtro.

Un bloque de medición de envolvente T/F que computa los datos de amplitud de tiempo-frecuencia (T/F) o energía o planicidad con una resolución espectral y temporal predefinida (esto también puede incluir mediciones de la transitoriedad del cuadro usado para las decisiones por cuadros, según lo indicado). El tono obtenido en el bloque de medición de la armonicidad es ingresado al bloque de medición de envolvente T/F dado que la región de la señal de audio usada para filtrar el cuadro actual, típicamente usando las muestras de señales pasadas, depende del tono (lo propio ocurre con el envolvente T/F computado).

Un bloque de computación de la ganancia de filtro toma la decisión final sobre qué ganancia de filtro usar para el filtrado (y transmitir en el caudal de bits). Idealmente, este bloque debería computar, para cada ganancia de filtro transmisible menor o igual a la ganancia de predicción, un envolvente tipo patrón de excitación espectro-temporal de la señal objetivo luego de filtrar con dicha ganancia de filtro, y comparar este envolvente "actual" con un envolvente de patrón de excitación de la señal original. Luego es posible usar para la codificación/transmisión la ganancia de filtro mayor cuyo envolvente "actual" espectro temporal correspondiente no difiera de la envolvente "original" en más de cierta medición. Esta ganancia de filtro será denominada psicoacústicamente óptima.

En otras realizaciones descritas a continuación, la estructura de tres bloques está un poco modificada.

En otras palabras, las mediciones de armonicidad y envolvente T/F se obtienen en los bloques correspondientes que se usan subsecuentemente para derivar los patrones de excitación psicoacústicos tanto de entrada como salida, y finalmente la ganancia de filtro se adapta de modo que un umbral de enmascaramiento, dado por una relación entre el envolvente "real" y el "original" no se exceda en forma significativa. Para apreciar esto, debería indicarse que un patrón de excitación en este contexto es muy similar a una representación tipo espectrograma de la señal examinada, pero que exhibe una atenuación temporal modelada a partir de ciertas características de la audición humana y que se manifiesta como "post-enmascaramiento".

La figura 1 ilustra la conexión entre los tres bloques antes presentados. Desafortunadamente, una derivación en el sentido de los cuadros de dos patrones de excitación y una búsqueda por fuerza bruta de la mejor ganancia de filtro a menudo es computacionalmente compleja. En consecuencia, se presentan simplificaciones en la siguiente descripción.

En orden a evitar computaciones costosas de los patrones de excitación en el esquema de decisión de activación de filtro propuesto, se usan mediciones de envolvente de baja complejidad como estimaciones de las características de los patrones de excitación. Se ha descubierto que en el bloque de medición de envolvente T/F, datos como las energías segmentales (SE), medición de planicidad temporal (TFM), cambio de energía máximo (MEC) o la información de configuración de cuadros tradicional como el tipo de cuadro (largo/fijo o corto/transitorio) es suficiente para derivar estimaciones de los criterios psicoacústicos. Estas estimaciones luego pueden ser utilizadas en el bloque de computación de la ganancia de filtro para determinar, con alta precisión, una ganancia de filtro óptima a ser empleada para la codificación o transmisión. En orden a evitar una búsqueda computacionalmente intensa de la ganancia globalmente óptima, un bucle de distorsión de tasa sobre todas las ganancias posibles (o un subconjunto de las mismas) puede sustituirse por operadores condicionales únicos. Estos operadores "baratos" sirven para decidir si la ganancia de filtro, computada usando datos de los bloques de medición de armonicidad y envolvente T/F, será cero (decisión de no usar el filtrado armónico) o no (decisión de usar el filtrado armónico). Adviértase que el bloque de medición de la armonicidad puede permanecer sin cambios. Una puesta en práctica paso a paso de esta realización de baja complejidad se describe seguidamente.

Según se ha indicado, la ganancia de filtro "inicial" sometida a los operadores condicionales únicos se deriva usando datos de los bloques de medición de armonicidad y envolvente T/F. Más específicamente, la ganancia de filtro "inicial" puede ser igual al producto de la ganancia de predicción variable con el tiempo (a partir del bloque de medición de la armonicidad) y un factor de escala variable por tiempo (a partir de los datos de envolvente psicoacústico del Bloque de medición de envolvente T/F). En orden a reducir aún más la carga computacional, un factor de escala constante, fijo como 0,625 puede usarse en lugar de la variable con el tiempo adaptable a la señal. Esto típicamente conserva la calidad suficiente y se considera en la siguiente realización.

Una descripción paso a paso de una realización concreta para controlar la herramienta de filtro se expone seguidamente.

1. Detección de transitorios y mediciones temporales

La señal de entrada ingresa al detector de transitorio por dominio de tiempo. La señal de entrada es sometida a un filtro de paso alto. La función de transferencia del filtro de paso alto de detección de transitorios es

$$H_{TD}(z) = 0,375 - 0,5z^{-1} + 0,125z^{-2} \quad (1)$$

La señal, filtrada por el filtro de paso alto de detección de transitorio, se indica como $s_{TD}(n)$. La señal sometida a un filtrado de paso alto $s_{TD}(n)$ se segmenta en ocho segmentos consecutivos de la misma longitud. La energía de la señal sometida a un filtrado de paso alto $s_{TD}(n)$ para cada segmento se calcula de la siguiente forma:

$$E_{TD}(i) = \sum_{n=0}^{L_{segmento}-1} (s_{TD}(iL_{segmento} + n))^2, \quad i = 0, \dots, 7 \quad (2)$$

donde $L_{segmento} = L/8$ es el número de muestras en un segmento de 2,5 milisegundos a la frecuencia de muestreo de entrada.

La energía acumulada se calcula usando:

$$E_{Acc} = \max(E_{TD}(i-1), 0.8125E_{Acc}) \quad (3)$$

Se detecta un ataque si la energía de un segmento $E_{TD}(i)$ excede la energía acumulada por un factor constante *Índice de ataque* = 8,5 y el Índice de ataque se establece en i :

$$E_{TD}(i) > \text{Índice de ataque } E_{Acc} \quad (4)$$

En caso de no detectarse ataque en base al criterio anterior, pero se detecta un fuerte incremento de la energía en el segmento i , el Índice de ataque se fija en i sin indicar la presencia de un ataque. El Índice de ataque básicamente se fija en la posición del último ataque en un cuadro con algunas restricciones adicionales.

El cambio de energía para cada segmento se calcula de la siguiente forma:

$$E_{chng}(i) = \begin{cases} \frac{E_{TD}(i)}{E_{TD}(i-1)}, & E_{TD}(i) > E_{TD}(i-1) \\ \frac{E_{TD}(i-1)}{E_{TD}(i)}, & E_{TD}(i-1) > E_{TD}(i) \end{cases} \quad (5)$$

La medición de planicidad temporal se calcula de la siguiente forma:

$$TFM(N_{past}) = \frac{1}{8 + N_{past}} \sum_{i=-N_{past}}^7 E_{chng}(i) \quad (6)$$

El cambio de energía máximo se calcula de la siguiente forma:

$$MEC(N_{past}, N_{new}) = \max(E_{chng}(-N_{past}), E_{chng}(-N_{past} + 1), \dots, E_{chng}(N_{new} - 1)) \quad (7)$$

Si el índice de $E_{chng}(i)$ o $E_{TD}(i)$ es negativo luego indica un valor del segmento previo, avanzando el segmento con relación al cuadro actual.

N_{past} es el número de los segmentos de los cuadros pasados. Es igual a 0 si la medición de planicidad temporal se calcula para el uso en la decisión ACELP/TCX. Si la medición de planicidad temporal es calculada para la decisión TCX LTP luego es igual a:

$$N_{past} = 1 + \min\left(8, \left\lceil 8 \frac{tono}{L} + 0,5 \right\rceil\right) \quad (8)$$

N_{new} es el número de segmentos del cuadro actual. Es igual a 8 para los cuadros no transitorios. Para los cuadros transitorios primero se establecen las ubicaciones de los segmentos con el máximo y mínimo de energía:

$$i_{\max} = \arg \max_{i \in \{-N_{\text{past}}, \dots, 7\}} E_{TD}(i) \quad (9)$$

$$i_{\min} = \arg \min_{i \in \{-N_{\text{past}}, \dots, 7\}} E_{TD}(i) \quad (10)$$

Si $E_{TD}(i_{\min}) > 0,375E_{TD}(i_{\max})$ luego N_{new} se fija en $i_{\max} - 3$, de lo contrario N_{new} se fija en 8.

2. Cambio de longitud del bloque de transformada

5 La longitud de superposición y la longitud del bloque de transformada del TCX dependen de la existencia de un transitorio y su ubicación.

Tabla 1: Codificación de la superposición y longitud de transformada en base a la posición de transitorio

10 Tabla 1: Codificación de la superposición y longitud de transformada en base a la posición de transitorio

Índice de ataque	Superposición con la primera ventana del siguiente cuadro	Decisión de Transformada Corta/Larga (codificación binaria) 0 - Larga, 1 - Corta	Código binario para el ancho de superposición	Código de superposición
Ninguno	ALDO	0	0	00
-2	TOTAL	1	0	10
-1	TOTAL	1	0	10
0	TOTAL	1	0	10
1	TOTAL	1	0	10
2	MÍNIMA	1	10	110
3	INTERMEDIA	1	11	111
4	INTERMEDIA	1	11	111
5	MÍNIMA	1	10	110
6	MÍNIMA	0	10	010
7	INTERMEDIA	0	11	011

15 El detector de transitorio antes descrito básicamente devuelve el índice del último ataque con la restricción de que si hay múltiples transitorios luego se prefiere una superposición MÍNIMA a una INTERMEDIA que se prefiere por sobre una superposición TOTAL. Si el ataque en la posición 2 o 6 no es suficientemente fuerte luego se elige la superposición INTERMEDIA en lugar de la MÍNIMA.

3. Estimación de Tono

20 Se estima un retardo de tono (parte entera + parte fracción) por cuadro (tamaño de cuadro por ejemplo 20 ms). Esto se realiza en tres etapas para reducir la complejidad y mejorar la precisión de la estimación.

a. Primera estimación en la parte entera del retardo de tono

25 Se usa un algoritmo de análisis de tono que produce un perfil de evolución de tono suave (por ejemplo, el análisis de tono de bucle abierto descrito en Rec. ITU-T G.718, sec. 6.6). Este análisis en general se realiza por subcuadros (tamaño de subcuadro por ejemplo 10 ms), y produce una estimación de retardo de tono por subcuadro. Advuértase que estas estimaciones de retardo de tono no tienen una parte fraccionada y en general se estiman sobre una señal muestreada hacia abajo (tasa de muestreo, por ejemplo, 6400 Hz). La señal usada puede ser cualquier señal de
30 audio, por ejemplo, una señal de audio compensada LPC como la descrita en Rec. ITU-T G.718, sec. 6.5.

b. Refinamiento de la parte entera del retardo de tono

La parte entera final del retardo de tono se estima sobre una señal de audio $x[n]$ que corre a la velocidad de

muestreo del codificador central, que en general es superior a aquella de la señal muestreada hacia abajo usada en a. (por ejemplo 12,8 kHz, 16 kHz, 32 kHz...). La señal $x[n]$ puede ser cualquier señal de audio por ejemplo una señal de audio compensada LPC.

- 5 La parte entera del retardo de tono luego es el retardo T_{int} que maximiza la función de auto-correlación

$$C(d) = \sum_{n=0}^L x[n]x[n-d]$$

con d alrededor de un retardo de tono T estimado en el paso 1.a.

$$T - \delta_1 \leq d \leq T + \delta_2$$

10

c. Estimación de la parte fraccional del retardo de tono

La parte fraccional se encuentra interpolando la función de auto correlación $C(d)$ computada en el paso 2.b. y seleccionando el retardo de tono fraccional T_{fr} que maximiza la función de auto correlación interpolada. La interpolación puede ser realizada un filtro de paso bajo FIR como el descrito por ejemplo en Rec. ITU-T G.718, sec. 6.6.7.

15

4. Bit de decisión

- 20 Si la señal de audio de entrada no contiene contenido armónico o si una técnica basada en la predicción introduciría distorsiones en la estructura temporal (por ejemplo, repetición de un transitorio corto), luego no se codifican parámetros en el caudal de bits. Se envía solamente 1 bit de modo tal que el decodificador conozca si debe decodificar los parámetros de filtrado o no. La decisión se toma en base a varios parámetros.

- 25 La correlación normalizada en el retardo de tono entero estimado en el paso 3.b.

$$norm_corr = \frac{\sum_{n=0}^L x[n]x[n-T_{int}]}{\sqrt{\sum_{n=0}^L x[n]x[n]} \sqrt{\sum_{n=0}^L x[n-T_{int}]x[n-T_{int}]}}$$

- 30 La correlación normalizada es 1 si la señal de entrada perfectamente predecible por un retardo de tono entero, y 0 si no es predecible. Un valor alto (cercano a 1) indicaría una señal armónica. Para una decisión más precisa, además de la correlación normalizada para el cuadro actual ($norm_corr(curr)$) la correlación normalizada del cuadro pasado ($norm_corr(prev)$) también puede usarse en la decisión, por ejemplo:

- 35 Si $(norm_corr(curr) * norm_corr(prev)) > 0,25$
 Si $\max(norm_corr(curr), norm_corr(prev)) > 0,5$,

luego el cuadro actual contiene algún contenido armónico (bit=1)

40

a. Características computadas por un detector de transitorio (por ejemplo Medición de planicidad temporal **¡Error! No se encuentra el origen de la referencia.**, Cambio de energía máxima **¡Error! No se encuentra el origen de la referencia.**), para evitar la activación del post filtro en una señal que contiene un transitorio fuerte o cambios temporales grandes. Las características temporales se calculan sobre la señal que contiene el cuadro actual (segmentos N_{new}) y el cuadro pasado hasta el retardo de tono (segmentos N_{past}). Para transitorios que están lentamente decayendo, todas o algunas de las características se calculan solamente hasta la posición del transitorio ($i_{max} - 3$) dado que las distorsiones en la parte no armónica del espectro introducidas por el filtrado LTO serían suprimidas por el enmascaramiento del transitorio prolongado (por ejemplo, platillo de choque).

45

b. Los trenes de pulsos para las señales de bajo tono pueden ser detectados como un transitorio por un detector de transitorio. Para las señales con bajo tono las características del detector de transitorio son ignoradas y hay en cambio un umbral adicional para la correlación normalizada que depende del retardo de tono, por ejemplo:

50

Si $norm_corr \leq 1,2 \cdot T_{int} / L$, luego se fija bit=0 y no se envían parámetros.

- 55 Un ejemplo de decisión se ilustra en la Figura 2 donde b1 es alguna velocidad de bits, por ejemplo 48 kbps, donde TCX_20 indica que el cuadro se codifica usando un bloque largo simple, donde TCX_10 indica que el cuadro se

codifica usando 2,3,4 o más bloques cortos, donde la decisión TCX_20/TCX_10 se basa en la salida del detector de transitorio que se ha descrito. tempFlatness es la Medición de planicidad temporal definida en **¡Error! No se encuentra el origen de la referencia.**, maxEnergyChange es el Cambio de energía máximo definido en **¡Error! No se encuentra el origen de la referencia.**. La condición $\text{norm_corr}(\text{curr}) > 1,2 \cdot T_{int}/L$ también podría ser escrita como $(1,2 \cdot \text{norm_corr}(\text{curr})) \cdot L < T_{int}$.

El principio de la lógica de decisión se ilustra en el diagrama en bloques en la Figura 3. Debe notarse que la Figura 3 es más general que la Figura 2 en el sentido que los umbrales no están limitados. Pueden fijarse de acuerdo con la Figura 2 o de una forma distinta. Por otra parte, la Figura 3 ilustra que el ejemplo de dependencia de la velocidad de bits de la Figura 2 puede dejarse de lado. Naturalmente, la lógica de decisión de la Figura 3 podría modificarse para incluir la dependencia de la velocidad de bits de la Figura 2. Además, la Figura 3 ha sido mantenida inespecífica con respecto al uso de solamente la corriente o además el tono pasado. En tanto, la Figura 3 ilustra que la realización de la Figura 2 puede ser modificada en este aspecto.

El "umbral" en la Figura 3 corresponde a distintos umbrales usados para tempFlatness y maxEnergyChange en la Figura 2. El "umbral_1" en la Figura 3 corresponde a $1,2 \cdot T_{int}/L$ en la Figura 2. El "umbral_2" en la Figura 3 corresponde a $0,44$ o $\max(\text{norm_corr}(\text{curr}), \text{norm_corr}(\text{prev})) > 0,5$ o $(\text{norm_corr}(\text{curr}) * \text{norm_corr}(\text{prev})) > 0,25$ en la Figura 2.

Resulta obvio a partir de los ejemplos anteriores que la detección de un transitorio afecta el mecanismo de decisión para la predicción a largo plazo y qué parte de la señal se usará para las mediciones usadas en la decisión, y no dispara directamente la desactivación de la predicción a largo plazo.

Las mediciones temporales usadas para la decisión de longitud de transformada pueden ser completamente distintas de aquellas empleadas para la decisión LTP o superponerse o ser exactamente las mismas, aunque calculadas en regiones diferentes.

Para señales de bajo tono la detección de transitorios es completamente ignorada si el umbral para la correlación normalizada que depende del retardo de tono es alcanzado.

5. Estimación y cuantización de ganancia

La ganancia en general se estima sobre la señal de audio de entrada a la velocidad de muestreo del codificador central, pero puede ser cualquier señal de audio como la Señal de audio compensada LPC. Esta señal es indicada como $y[n]$ y puede ser igual o distinta de $x[n]$.

La predicción $y_p[n]$ de $y[n]$ primero se encuentra filtrando $y[n]$ con el siguiente filtro

$$P(z) = B(z, T_{fr})z^{-T_{int}}$$

donde T_{int} es la parte entera del retardo de tono (estimado en 0) y $B(z, T_{fr})$ un FIR filtro de paso bajo cuyo coeficiente depende de la parte fraccional del retardo de tono T_{fr} (estimado en 0).

Un ejemplo de $B(z)$ cuando la resolución del retardo de tono es $\frac{1}{4}$:

$$T_{fr} = \frac{0}{4} \quad B(z) = 0,0000z^{-2} + 0,2325z^{-1} + 0,5349z^0 + 0,2325z^1$$

$$T_{fr} = \frac{1}{4} \quad B(z) = 0,0152z^{-2} + 0,3400z^{-1} + 0,5094z^0 + 0,1353z^1$$

$$T_{fr} = \frac{2}{4} \quad B(z) = 0,0609z^{-2} + 0,4391z^{-1} + 0,4391z^0 + 0,0609z^1$$

$$T_{fr} = \frac{3}{4} \quad B(z) = 0,1353z^{-2} + 0,5094z^{-1} + 0,3400z^0 + 0,0152z^1$$

La ganancia g entonces se computa de la siguiente forma:

$$g = \frac{\sum_{n=0}^{L-1} y[n]y_P[n]}{\sum_{n=0}^{L-1} y_P[n]y_P[n]}$$

y se limita entre 0 y 1.

5 Finalmente, la ganancia se cuantiza por ejemplo en 2 bits, usando por ejemplo cuantización uniforme. Si la ganancia se cuantiza en 0, luego no se codifica ningún parámetro en el caudal de bits, solamente el bit de decisión 1 (bit=0).

La descripción que antecede mediante la cual se han indicado las ventajas de las realizaciones de la presente para un control dependiente de la armonicidad de una herramienta de filtro de armónicos, también se aplica a las realizaciones generales que siguen. En algunos casos la descripción anterior ha sido muy específica, aunque el concepto de control dependiente de la armonicidad también puede ser aplicado en el marco de otros códecs de audio y puede modificarse respecto de los detalles antes expuestos. Por esta razón, las realizaciones de la presente solicitud serán descritas nuevamente de una manera más genérica. Sin embargo, la siguiente descripción podrá hacer referencia a la descripción detallada anterior con explicar la implementación de las realizaciones generales. A tal fin, es posible transferir ciertos detalles de implementación de la descripción anterior a los elementos que se describen seguidamente. En consecuencia, cada vez que en la descripción que sigue se haga referencia a la anterior, se entenderá que es independiente de las referencias adicionales a la misma.

Así, una realización más genérica que emerge de la descripción detallada previa se ilustra en la Figura 4. En particular, la Figura 4 ilustra un aparato para ejecutar un control dependiente de la armonicidad de una herramienta de filtro de armónicos, como una herramienta de filtro pre/post armónico o post-filtro, de un códec de audio. El aparato en general se indica mediante la referencia 10. El aparato 10 recibe la señal de audio 12 a procesar por el códec de audio y emite una señal de control 14 para cumplir con la tarea de control del aparato 10. El aparato 10 comprende un estimador de tono 16 configurado para determinar un retardo de tono actual 18 de la señal de audio 12, y un medidor de armonicidad 20 configurado para determinar una medición 22 de la armonicidad de la señal de audio 12 usando un retardo de tono actual 18. En particular, la medición de armonicidad puede ser una ganancia de predicción o puede llevarse a la práctica mediante uno o más coeficientes de filtrado (único o multi-tap) o una correlación normalizada máxima. El bloque de cálculo de la medición de armonicidad de la Figura 1 comprende las tareas del estimador de tono 16 y medidor de armonicidad 20.

El aparato 10 además comprende un analizador de estructura temporal 24 configurado para determinar por lo menos una medición de estructura temporal 26 de acuerdo con el retardo de tono 18, la medición 26 que mide una característica de una estructura temporal de la señal de audio 12. Por ejemplo, la dependencia puede basarse en el posicionamiento de la región temporal dentro de la cual la medición 26 mide la característica de una estructura temporal de la señal de audio 12, según se ha descrito y se describirá con mayor detalle a continuación. En honor a la completitud, sin embargo, adviértase que la dependencia de la determinación de la medición 26 sobre el retardo de tono 18 puede llevarse a la práctica de una manera diferente a la descrita con anterioridad y seguidamente. Por ejemplo, en lugar de posicionar la porción temporal, es decir, la ventana de determinación, de acuerdo con el retardo de tono, la dependencia podría meramente variar temporalmente los pesos a los cuales un respectivo intervalo de tiempo de la señal de audio dentro de una ventana dispuesta independientemente del retardo de tono con relación al cuadro actual, contribuyen a la medición 26. Con relación a la descripción que sigue, esto puede significar que la ventana de determinación 36 podría disponerse sostenidamente para corresponder con la concatenación de los cuadros actual y previo, y que la porción dispuesta con relación al tono meramente funciona como una ventana de un peso mayor al cual la estructura temporal de la señal de audio influencia la medición 26. Sin embargo, por el momento, se asume que la ventana temporal se dispone de acuerdo con el retardo de tono. El analizador de estructura temporal 24 corresponde al bloque de cálculo de la medición de envolvente T/F de la Figura 1.

Finalmente, el aparato de la Figura 4 comprende un controlador 28 configurado para emitir una señal de control 14 de acuerdo con la medición de estructura temporal 26 y la medición 22 de la armonicidad para en consecuencia controlar el pre/post filtro de armónicos o el post-filtro de armónicos. Cuando se compara la Figura 4 con la Figura 1, el bloque de computación de la ganancia de filtro óptima corresponde o representa una implementación posible del controlador 28.

El modo de operación del aparato 10 es el siguiente. En particular, la función del aparato 10 es controlar la herramienta de filtro de armónicos de un códec de audio, y a pesar de que la descripción más detallada anterior con respecto a las Figuras 1 a 3 revela un control o adaptación gradual de esta herramienta en términos de su intensidad de filtro o ganancia de filtro, por ejemplo, el controlador 28 no se limita a esa clase de control gradual. En general, el control mediante el controlador 28 puede gradualmente adaptar la intensidad de filtro o ganancia de la herramienta de filtro de armonicidad entre 0 y un valor máximo valor máximo, inclusive, como en el caso de los ejemplos específicos anteriores con respecto a las Figuras 1 a 3, aunque distintas alternativas son posibles, como un control gradual entre dos valores de ganancia de filtro no-cero, un control progresivo o un control binario como el cambio entre activación (no cero) y desactivación (ganancia cero) para encender y apagar la herramienta de filtro de

armónicos.

Como queda claro a partir de la exposición que antecede, la herramienta de filtro de armónicos que se ilustra en la Figura 4 mediante líneas punteadas 30 tiene por objeto mejorar la calidad subjetiva de un códec de audio como un códec de audio a base de transformada, en especial con respecto a las fases armónicas de la señal de audio. En particular, esta herramienta 30 es especialmente útil en escenarios de baja velocidad de bits donde el ruido de cuantización introducido, sin la herramienta de filtro 30, conducirían tales fases armónicas a artefactos audibles. Es importante, sin embargo, que la herramienta de filtro 30 no afecte negativamente otras fases temporales de la señal de audio que no sean predominantemente armónicas. Además, según se ha indicado, la herramienta de filtro 30 puede ser del enfoque post-filtro o pre-filtro. Los pre y post-filtros pueden operar en el dominio de transformada o dominio de tiempo. Por ejemplo, un post-filtro de la herramienta 30 puede, por ejemplo, cumplir una función de transferencia con una local máxima que se dispone a distancias espectrales que corresponden a, o se fijan de acuerdo con el retardo de tono 18. La implementación de un pre-filtro y/o post-filtro bajo la forma de un filtro de LTP, en la forma de, por ejemplo, un filtro FIR y IIR, respectivamente, también es posible. El pre-filtro puede tener una función de transferencia que es sustancialmente inversa a la función de transferencia del post-filtro. En efecto, el pre-filtro busca ocultar el ruido de cuantización dentro del componente armónico de la señal de audio incrementando el ruido de cuantización dentro del armónico del tono actual de la señal de audio y el post-filtro reformula el espectro transmitido en consecuencia. En el caso del enfoque con post-filtro solamente, el post-filtro modifica la señal de audio transmitida para filtrar el ruido de cuantización que ocurre entre los armónicos del tono de la señal de audio.

Debe advertirse que la Figura 4 está, en algún sentido, concebida de una forma simple. Por ejemplo, a pesar de que la Figura 4 sugiere que el estimador de tono 16, el medidor de armonicidad 20 y el analizador de estructura temporal 24 operan, es decir, cumplen sus funciones, sobre la señal de audio 12 en forma directa, ello no es necesario. En cambio, el estimador de tono 16, el analizador de estructura temporal 24 y el medidor de armonicidad 20 pueden operar sobre diferentes versiones de la señal de audio 12 como aquellas distintas de la original y alguna versión pre-modificada de la misma, en tanto estas versiones pueden variar entre los elementos 16, 20 y 24 internamente y también con relación al códec de audio también, pudiendo además operar sobre alguna versión modificada de la señal de audio original. Por ejemplo, el analizador de estructura temporal 24 puede operar sobre la señal de audio 12 a la velocidad de muestreo de entrada de la misma, es decir la velocidad de muestreo original de la señal de audio 12, o puede operar sobre una versión internamente codificada/decodificada de la misma. El códec de audio, a su vez, puede operar a una velocidad de muestreo central interna que es usualmente inferior a la velocidad de muestreo interna. El estimador de tono 16, a su vez, puede realizar su estimación de tono sobre una versión pre-modificada de la señal de audio, como, por ejemplo, una versión psicoacústicamente compensada de la señal de audio 12 con el fin de mejorar la estimación de tono con respecto a los componentes espectrales que son, en términos de perceptibilidad, más significativos que otros componentes espectrales. Por ejemplo, según se ha descrito, el estimador de tono 16 puede estar configurado para determinar el retardo de tono 18 en etapas que comprenden una primera etapa y una segunda etapa, la primera etapa da como resultado una estimación preliminar del retardo de tono que luego es refinada en la segunda etapa. Por ejemplo, según se ha descrito, el estimador de tono 16 puede determinar una estimación preliminar del retardo de tono en un dominio muestreado hacia abajo que corresponde a una primera velocidad de muestreo, y luego refinando la estimación preliminar del retardo de tono a una segunda velocidad de muestreo que es mayor que la primera velocidad de muestreo.

En lo relativo al medidor de armonicidad 20, queda claro a partir de la exposición anterior con respecto a las Figuras 1 a 3 que puede determinar la medición 22 de la armonicidad computando una correlación normalizada de la señal de audio o una versión pre-modificada de la misma en el retardo de tono 18. Debe advertirse que el medidor de armonicidad 20 aún puede configurarse para computar la correlación normalizada aún a varias distancias de correlación además del retardo de tono 18 como en un intervalo de retardo temporal que incluye y circunda el retardo de tono 18. Esto puede ser favorable, por ejemplo, en el caso de una herramienta de filtro 30 que use una LTP multi-tap o LTP posible con tono fraccional. En ese caso, el medidor de armonicidad 20 puede analizar o evaluar la correlación aún a índices de retardo cercanos al retardo de tono actual 18, como el retardo de tono entero en el ejemplo concreto anterior que se expone en las Figuras 1 a 3.

Para detalles y posibles implementaciones adicionales del estimador de tono 16, refiérase a la sección "estimación de tono" anterior. Las posibles implementaciones del medidor de armonicidad 20 fueron expuestas anteriormente con respecto a la ecuación norm.corr. Sin embargo, también según lo descrito, la expresión "medición de armonicidad" no solamente incluirá una correlación normalizada sino que además sugiere la ganancia de predicción del filtro de armónicos, en tanto el filtro de armónicos puede ser igual o no al pre-filtro del filtro 230 en caso de usar en enfoque pre/post-filtro independientemente del códec de audio que use este filtro de armónicos o si este filtro de armónicos es meramente usado por el medidor de armónicos 20 para establecer la medición 22.

Como en el caso de las Figuras 1 a 3 antes descritas, el analizador de estructura temporal 24 puede estar configurado para determinar la por lo menos única medición de estructura temporal 26 dentro de una región temporal temporalmente colocada de acuerdo con el retardo de tono 18. Para mayor ilustración, véase la Figura 5. La Figura 5 ilustra un espectrograma 32 de la señal de audio, es decir su descomposición espectral hasta su

frecuencia más alta f_H de acuerdo con, por ejemplo, la velocidad de muestreo de la versión de la señal de audio internamente usada por el analizador de estructura temporal 24, temporalmente muestreada a una velocidad de bloque de transformada que puede o no coincidir con la velocidad de bloque de transformada del códec de audio, en su caso. Con fines ilustrativos, la Figura 5 muestra el espectrograma 32 como temporalmente subdividido en cuadros en unidades con las cuales el controlador puede realizar el control de la herramienta de filtro 30, la subdivisión puede, por ejemplo, además coincidir con la subdivisión de cuadros usada por el códec de audio que comprende o usa la herramienta de filtro 30.

Por el momento, a los fines de la ilustración se asume que el cuadro actual para el cual el controlador 28 ejecuta el control, es el cuadro 34a. Según se ha descrito y conforme lo ilustrado en la Figura 5, la región temporal 36, dentro de la cual el determinador de analizador de estructura temporal determina la por lo menos única medición de estructura temporal 26, no necesariamente coincide con los cuadros actuales 34a. En cambio, tanto el extremo de encabezado temporalmente pasado 38 como el extremo de encabezado temporalmente futuro 40 de la región temporal 36 pueden apartarse de los extremos de encabezado temporalmente pasado y futuro 42 y 44 del cuadro actual 34a. Según se ha descrito, el analizador de estructura temporal 24 dispone el extremo de encabezado temporalmente pasado 38 de la región temporal 36 de acuerdo con el retardo de tono 18 determinado por el estimador de tono 16 que establece el retardo de tono 18 para cada cuadro 34, por cuadro actual 34a. Tal como se advertirá de la descripción anterior, el analizador de estructura temporal 24 dispone el extremo de encabezado temporalmente pasado 38 de la región temporal de modo tal que el extremo de encabezado temporalmente pasado 38 se desplace hacia el pasado con relación al extremo de encabezado temporalmente pasado 42 del cuadro actual 34a, por ejemplo, en un grado 46 que monótonicamente se incrementa con el aumento del retardo de tono 18. En otras palabras, cuanto mayor es el retardo de tono 18, mayor será 46. Según se advertirá a partir de la exposición anterior con respecto a las Figuras 1 a 3, el grado puede fijarse de acuerdo con la ecuación 8, donde N_{past} es una medición para el desplazamiento temporal 46.

El extremo de encabezado temporalmente futuro 40 de región temporal 36, a su vez, puede ser fijado por el analizador de estructura temporal 24 de acuerdo con la estructura temporal de la señal de audio dentro de una región candidato temporal 48 que se extiende desde el extremo de encabezado temporalmente pasado 38 de la región temporal 36 al extremo de encabezado temporalmente futuro del cuadro actual, 44. En particular, según se ha expuesto, el analizador de estructura temporal 24 puede evaluar una medición de disparidad de muestras de energía de la señal de audio dentro de la región candidato temporal 48 para decidir sobre la posición del extremo de encabezado temporalmente futuro 40 de la región temporal 36. En los detalles específicos anteriores presentados con respecto a las Figuras 1 a 3, una medición de la diferencia entre las muestras de energía máxima y mínima dentro de la región candidato temporal 48 fue usada como medición de disparidad, como una relación de amplitud entre ellas. En particular, en el ejemplo concreto anterior, la variable N_{new} mide el extremo de encabezado temporalmente futuro 40 del temporalmente futuro 36 con respecto al extremo de encabezado temporalmente pasado 42 del cuadro actual 34a según lo indicado en 50 en la Figura 5.

Tal como se advertirá de la exposición que antecede, la disposición de la región temporal 36 dependiente del retardo de tono 18 es conveniente por la habilidad del aparato 10 para identificar correctamente situaciones donde la herramienta de filtro de armónicos 30 puede ser convenientemente usada. En particular, la detección correcta de tales situaciones es más confiable, es decir tales situaciones se detectan con mayor probabilidad sin incrementar sustancialmente la detección de falsos positivos.

Según se ha descrito con respecto a las Figuras 1 a 3, el analizador de estructura temporal 24 puede determinar la por lo menos única medición de estructura temporal dentro de la región temporal 36 sobre la base de un muestreo temporal de la energía de la señal de audio dentro de esa región temporal 36. Esto se ilustra en la Figura 6, donde las muestras de energía son indicadas por puntos trazados sobre un plano de tiempo/energía abarcado por ejes de tiempo y energía arbitrarios. Según se ha explicado, las muestras de energía 52 pueden haber sido obtenidas muestreando la energía de la señal de audio a una velocidad de muestreo superior a la velocidad de cuadro de los cuadros 34. En la determinación de la por lo menos única medición de estructura temporal 26, el analizador 24 puede, según se ha expuesto, computar por ejemplo un conjunto de valores de cambio de energía durante un cambio entre los pares de muestras de energía inmediatamente consecutivos 52 dentro de la región temporal 36. En la descripción anterior, la ecuación 5 fue usada a tal fin. Mediante esta medición, un valor de cambio de energía puede ser obtenido de cada par de muestras de energía inmediatamente consecutivas 52. El analizador 24 puede entonces someter el conjunto de valores de cambio de energía obtenido de las muestras de energía 52 dentro de la región temporal 36 a una función escalar para obtener la por lo menos única medición de energía estructural 26. En el ejemplo concreto que antecede, la medición de planicidad temporal, por ejemplo, ha sido determinada sobre la base de una suma de sumandos, que dependen respectivamente de exactamente un conjunto de valores de cambio de energía. El cambio de energía máximo, a su vez, fue determinado de acuerdo con la ecuación 7 usando un operador máximo aplicado sobre los valores de cambio de energía.

Según se ha indicado, las muestras de energía 52 no necesariamente miden la energía de la señal de audio 12 en su versión original sin modificar. En cambio, el muestreo de energía 52 puede medir la energía de la señal de audio

en algún dominio modificado. En el ejemplo concreto anterior, por ejemplo, las muestras de energía miden la energía de la señal de audio obtenida luego del filtrado de paso alto de la misma. En consecuencia, la energía de la señal de audio en una región espectralmente inferior influencia las muestras de energía 52 menos que los componentes espectralmente superiores de la señal de audio. También hay otras posibilidades, sin embargo. En particular, debe notarse que el ejemplo donde el analizador de estructura temporal 24 meramente usa un valor de la por lo menos única medición de estructura temporal 26 por instante de acuerdo con los ejemplos hasta aquí presentados, ello es simplemente una realización y existen alternativas de acuerdo con las cuales el analizador de estructura temporal determina la medición de estructura temporal de una forma espectralmente discriminante con el fin de obtener un valor de la por lo menos única medición de estructura temporal por banda espectral de una pluralidad de bandas espectrales. En consecuencia, el analizador de estructura temporal 24 entonces proveería al controlador 28 más de un valor de la por lo menos única medición de estructura temporal 26 para el cuadro actual 34a determinada dentro de la región temporal 36, a saber, una por dicha banda espectral, en tanto las bandas espectrales se dividen, por ejemplo, el intervalo espectral general del espectrograma 32.

La Figura 7 ilustra el aparato 10 y su uso en un códec de audio que soporta la herramienta de filtro de armónicos 30 de acuerdo con el enfoque de filtro pre/post armónico. La Figura 7 ilustra un codificador basado en transformada 70 así como un decodificador basado en transformada 72 con el codificador 70 que codifica la señal de audio 12 en un caudal de datos 74 y el decodificador 72 recibe el caudal de datos 74 para reconstruir la señal de audio en el dominio espectral según lo ilustrado en 76 u, opcionalmente, en el dominio de tiempo ilustrado en 78. Debe quedar claro que el codificador y decodificador 70 y 72 son entidades distintas/separadas y se ilustran conjuntamente en la Figura 7 solamente con fines ilustrativos.

El codificador a base de transformada 70 comprende un transformador 80 que somete la señal de audio 12 a una transformada. El transformador 80 puede usar una transformada “superpuesta” como una transformada superpuesta críticamente muestreada, por ejemplo, MDCT. En el ejemplo de la Figura 7, el codificador de audio a base de transformada 70 además comprende un conformador espectral 82 que conforma espectralmente la señal del espectro de audio como salida del transformador 80. El conformador espectral 82 puede conformar espectralmente el espectro de la señal de audio de acuerdo con una función de transferencia que es sustancialmente inversa a una función perceptual espectral. La función perceptual espectral puede ser derivada por predicción lineal y así, la información relativa a la función perceptual espectral puede ser transferida al decodificador 72 dentro del caudal de datos 74 en la forma de, por ejemplo, coeficientes de predicción lineal en la forma de, por ejemplo, un par espectral de líneas cuantizadas de valores de frecuencia espectrales lineales. Alternativamente, es posible usar un modelo perceptual para determinar la función perceptual espectral en la forma de factores de escala, un factor de escala por banda de factores de escala, las bandas pueden, por ejemplo, coincidir con las bandas Bark. El codificador 70 además comprende un cuantificador 84 que cuantiza el espectro con, por ejemplo, una función de cuantización que es igual para todas las líneas espectrales. El espectro espectralmente formado y cuantizado es transportado dentro del caudal de datos 74 al decodificador 72.

Por motivo de completitud solamente, debe notarse que el orden entre el transformador 80 y conformador espectral 82 ha sido seleccionado en la Figura 7 para fines ilustrativos solamente. Teóricamente, el conformador espectral 82 podría generar la formación espectral dentro del dominio de tiempo, es decir, por arriba del transformador 80. Además, con el fin de determinar la función perceptual espectral, el conformador espectral 82 podría acceder a la señal de audio 12 en el dominio de tiempo, aunque no se indica específicamente en la Figura 7. En el lado del decodificador, el decodificador 72 se ilustra en la Figura 7 como comprendiendo el conformador espectral 86 configurado para formar el espectro interno espectralmente formado y cuantizado obtenido del caudal de datos 74 con la inversa de la función de transferencia del conformador espectral 82, es decir sustancialmente con la función perceptual espectral, seguido de un transformador inverso opcional 88. El transformador inverso 88 realiza la transformación inversa con relación al transformador 80 y puede, por ejemplo, a tal fin realizar una transformación inversa en base a bloques seguida de un proceso de superposición-suma con el fin de realizar la cancelación de la distorsión del dominio de tiempo, por medio de lo cual se reconstruye la señal de audio en el dominio de tiempo.

De acuerdo con lo ilustrado en la Figura 7, un pre-filtro de armónicos puede estar comprendido por el codificador 70 en una posición por arriba o debajo del transformador 80. Por ejemplo, un pre-filtro de armónicos 90 por arriba del transformador 80 puede someter la señal de audio 12 dentro del dominio de tiempo a un filtrado para efectivamente atenuar el espectro de la señal de audio en los armónicos además de la función de transferencia o conformador espectral 82. Alternativamente, el pre-filtro de armónicos puede disponerse por debajo del transformador 80 en tanto este pre-filtro 92 realiza o provoca la misma atenuación en el dominio espectral. Según lo ilustrado en la Figura 7, los correspondientes post-filtros 94 y 96 están dispuestos dentro del decodificador 72: el pre-filtro 92, dentro del post-filtro del dominio espectral 94 dispuesto por arriba del transformador inverso 88 forma inversamente el espectro de la señal de audio, inversa a la función de transferencia de pre-filtro 92, y en el caso de usarse el pre-filtro 90, el post filtro 96 realiza un filtrado de la señal de audio reconstruida en el dominio de tiempo, por debajo del transformador inverso 88, con una función de transferencia inversa a la función de transferencia del pre-filtro 90.

En el caso de la Figura 7, el aparato 10 controla la herramienta de filtro de armónicos del códec de audio

implementada por el par 90 y 96 o 92 y 94 enviando explícitamente las señales de control 98 mediante el códec de audio del caudal de datos 74 al lado de decodificación para controlar el respectivo post-filtro y, en línea con el control del post-filtro en el lado de decodificación, controlando el pre-filtro en el lado del codificador.

5 En honor a la completitud, la Figura 8 ilustra el uso del aparato 10 que usa un códec de audio basado en transformada además los elementos 80, 82, 84, 86 y 88, sin embargo, aquí se ilustra el caso en el cual el códec de audio soporta solamente el post-filtro de armónicos. En este caso, la herramienta de filtro de armónicos 30 puede ser llevada a la práctica mediante un post-filtro 100 dispuesto por arriba del transformador inverso 88 dentro del decodificador 72, con el fin de ejecutar el filtrado post-armónico en el dominio espectral, o mediante el uso de un
10 post-filtro 102 dispuesto por debajo del transformador inverso 88 para ejecutar el post-filtrado armónico dentro del decodificador 72 dentro del dominio de tiempo. El modo de operación de los post-filtros 100 y 102 es sustancialmente similar al de los post-filtros 94 y 96: el objeto de estos post-filtros es atenuar el ruido de cuantización entre los armónicos. El aparato 10 controla estos post-filtros mediante una señalización explícita dentro del caudal de datos 74, la señalización explícita es indicada en la Figura 8 usando la referencia 104.

15 Según se ha expuesto, la señal de control 98 o 104 es enviada, por ejemplo, regularmente, como por cuadro 34. En cuanto a los cuadros, se advierte que los mismos no son necesariamente de la misma longitud. La longitud de los cuadros 34 también puede variar.

20 La descripción anterior, en especial aquella de la Figura 2 y 3, revela posibilidades acerca de cómo el controlador 28 controla la herramienta de filtro de armónicos. Tal como surge de esta exposición, puede ser que la por lo menos única medición de estructura temporal mida una variación de energía promedio o máxima de la señal de audio dentro de la región temporal 36. Además, el controlador 28 puede incluir, dentro de sus opciones de control, la desactivación de la herramienta de filtro de armónicos 30. Esto se ilustra en la Figura 9. La Figura 9 ilustra el
25 controlador 28 que comprende una lógica 120 configurada para determinar si una condición predeterminada es satisfecha por la por lo menos única medición de estructura temporal y la medición de armonicidad, para obtener un resultado del chequeo 122, que es de naturaleza primaria e indica si la condición predeterminada se cumple o no. El controlador 28 comprende un interruptor 124 configurado para activar y desactivar la herramienta de filtro de armónicos de acuerdo con el resultado de la prueba 122. Si el resultado del chequeo 122 indica que la condición predeterminada ha sido aprobada por la lógica 120, el interruptor 124 indica directamente la situación mediante la
30 señal de control 14, o el interruptor 124 indica la situación junto con un grado de ganancia de filtro para la herramienta de filtro de armónicos 30. Es decir, en el último caso, el interruptor 124 no alternaría entre desactivar la herramienta de filtro de armónicos 30 por completo y encender la herramienta de filtro de armónicos 30 por completo, solamente, sino que pondría la herramienta de filtro de armónicos 30 en algún estado intermedio que varíe la intensidad del filtrado o ganancia de filtro, respectivamente. En ese caso, si el interruptor 124 además
35 adapta/controla la herramienta de filtro de armónicos 30 en algún punto entre la activación/desactivación completa de la herramienta 30, el interruptor 124 puede basarse en la medición de estructura temporal 26 y la medición de armonicidad 22 con el fin de establecer los estados intermedios de la señal de control 14, es decir con el fin de adaptar la herramienta 30. En otras palabras, el interruptor 124 podría determinar el factor de ganancia o factor de adaptación para controlar la herramienta de filtro de armónicos 30 sobre la base de las mediciones 26 y 22. Alternativamente, el interruptor 124 usa para todos los estados de la señal de control 14 que no indican el estado off de la herramienta de filtro de armónicos 30, la señal de audio 12 directamente. Si el resultado del chequeo 122 indica que una condición predeterminada no se ha satisfecho, luego la señal de control 14 indica la desactivación de la herramienta de filtro de armónicos 30.

45 Tal como surge de la descripción anterior de las Figuras 2 y 3, la condición predeterminada puede cumplirse si la por lo menos única medición de estructura temporal es menor que un primer umbral predeterminado y la medición de armonicidad es, para un cuadro actual y/o a cuadro previo, superior a un segundo umbral. Puede haber otra alternativa: la condición predeterminada además puede ser satisfecha si la medición de armonicidad es, para un
50 cuadro actual, superior a un tercer umbral y la medición de armonicidad es, para un cuadro actual y/o un cuadro previo, superior a un cuarto umbral que disminuye con un aumento del retardo de tono.

En particular, en el ejemplo de las Figuras 2 y 3, hay virtualmente tres alternativas para las cuales se cumple con la condición predeterminada, las alternativas dependen de la por lo menos única medición de estructura temporal:

- 55
1. Una medición de estructura temporal < umbral y armonicidad combinada para cuadro actual y previo > segundo umbral;
 2. Una medición de estructura temporal < tercer umbral y (armonicidad para cuadro actual y previo) > cuarto umbral;
 - 60 3. (Una medición de estructura temporal < quinto umbral o todas las mediciones temp. < umbrales) y armonicidad para el cuadro actual > sexto umbral.

Así, la Figura 2 y la Figura 3, revelan la implementación de ejemplos posibles para la lógica 124.

De acuerdo con lo ilustrado anteriormente con respecto a las Figuras 1 a 3, es posible que el aparato 10 no se use solamente para controlar una herramienta de filtro de armónicos de un códec de audio. En cambio, el aparato 10 puede formar, junto con una detección transitoria, un sistema capaz de realizar el control de la herramienta de filtro de armónicos, así como la detección de transitorios. La Figura 10 ilustra esta posibilidad. La Figura 10 ilustra un sistema 150 compuesto por el aparato 10 y un detector de transitorio 152, y a pesar que el aparato 10 emite la señal de control 14 según se ha expuesto, el detector de transitorio 152 está configurado para detectar transitorios en la señal de audio 12. A tal fin, sin embargo, el detector de transitorio 152 explota un resultado intermedio que ocurre dentro del aparato 10: el detector de transitorio 152 usa para su detección las muestras de energía 52 temporalmente o, alternativamente, muestreando espectro-temporalmente la energía de la señal de audio, sin embargo, evaluado opcionalmente las muestras de energía dentro de una región temporal además de la región temporal 36 como dentro del cuadro actual 34a, por ejemplo. Sobre la base de las muestras de energía, el detector de transitorio 152 ejecuta la detección de transitorios y señala los transitorios detectados mediante una señal de detección 154. En el caso del ejemplo previo, la señal de detección de transitorios sustancialmente indica las posiciones donde se cumple con la condición de la ecuación 4, es decir cuando el cambio de energía de las muestras de energía temporalmente consecutivos excede cierto umbral.

Tal como queda claro a partir de la exposición anterior, un codificador a base de transformada como el ilustrado en la Figura 8 o un codificador de excitación codificado por transformada puede comprender o usar el sistema de la Figura 10 para conmutar un bloque de transformada y/o longitud de superposición de acuerdo con la señal de detección de transitorio 154. Además, o alternativamente, un codificador de audio que comprende o usa el sistema de la Figura 10 puede ser de tipo conmutación. Por ejemplo, USAC y EVS usan conmutación entre modos. Así, este codificador podría estar configurado para soportar la conmutación entre un modo de excitación codificado por transformada y un modo de predicción lineal excitado por código y el codificador podría estar configurado para realizar la conmutación de acuerdo con la señal de detección de transitorio 154 del sistema de la Figura 10. En cuanto al modo de excitación codificado por transformada, la conmutación del bloque de transformada y/o longitud de superposición podrían, nuevamente, depender de la señal de detección de transitorio 154.

Ejemplos de los beneficios de las realizaciones previas

30 **Ejemplo 1:**

El tamaño de la región donde las mediciones temporales para la decisión LTP se calculan depende del tono (véase la ecuación **¡Error! No se encuentra el origen de la referencia.**) y esta región es distinta de la región donde las mediciones temporales para la longitud de transformada se calculan (usualmente cuadro actual y posterior).

35 En el ejemplo de la Figura 11 el transitorio está dentro de la región donde las mediciones temporales se calculan y así se influencia la decisión LTP. La motivación, según se ha dicho, que una LTP para el cuadro actual, que usa las muestras pasadas del segmento indicado "retardo de tono", alcanzaría una porción del transitorio.

40 En el ejemplo en la Figura 12 el transitorio está fuera de la región donde las mediciones temporales se calculan y por ello no influencia la decisión LTP. Esto es razonable dado que, a diferencia de la figura previa, una LTP para el cuadro actual no alcanzaría el transitorio.

45 En ambos ejemplos (Figura 11 y Figura 12) la configuración de longitud de transformada se decide sobre las mediciones temporales solamente dentro del cuadro actual, es decir la región marcada con "longitud de cuadro". Esto significa que, en ambos ejemplos, no se detectaría transitorio en el cuadro actual y preferentemente una sola transformada larga (en lugar de muchas cortas) se emplearía.

50 **Ejemplo 2:**

En este caso describimos el comportamiento de la LTP para los transitorios de impulso y paso dentro de la señal armónica, según se ilustra en el espectrograma de la Figura 13.

55 Cuando se codifica la señal incluye la LTP para la señal completa (dado que la decisión LTP se basa solamente en la ganancia de tono), el espectrograma de la salida luce como se presenta en la Figura 14.

60 La forma de onda de la señal, cuyo espectrograma está en la Figura 14, se presenta en la Figura 15. La Figura 15 además incluye la misma señal filtrada por paso bajo (LP) paso alto (HP). En la señal filtrada por LP la estructura armónica se torna más clara y en la señal filtrada por HO la ubicación del impulso como transitorio y su rastro es más evidente. El nivel de la señal completa, la señal LP y HP se modifica en la figura en honor a la presentación.

Para transitorios de impulso corto (como el primer transitorio en la Figura 13), la predicción de largo plazo produce repeticiones del transitorio tal como puede observarse en la Figura 14 y la Figura 15. El uso de la predicción de largo plazo durante los transitorios largos tipo paso (como el segundo transitorio en la Figura 13) no introduce distorsiones

adicionales dado que el transitorio es lo suficientemente fuerte durante un período prolongado y así enmascara (simultáneamente y post-enmascaramiento) las porciones de la señal construida usando esta predicción. El mecanismo de decisión permite la LTP para transitorios tipo paso (para explotar el beneficio de la predicción) y desactiva la LTP para el transitorio de impulso corto (para prevenir artefactos).

5 En la Figura 16 y la Figura 17, se ilustran las energías de los segmentos computados en el detector de transitorio. La Figura 16 ilustra el transitorio de impulso, la Figura 17 ilustra el transitorio tipo paso. Para el transitorio de impulso en la Figura 16 las características temporales se calculan sobre la señal que contiene el cuadro actual (N_{new}

segmentos) y el cuadro pasado hasta el retardo de tono (segmentos N_{past}), dado que la relación $\frac{E_{TD}(i_{max})}{E_{TD}(i_{min})}$ está

10 por sobre el umbral (1/0,375). Para el transitorio tipo paso en la Figura 17, la relación $\frac{E_{TD}(i_{max})}{E_{TD}(i_{min})}$ se encuentra por

debajo del umbral (1/0,375) y así solamente las energías de los segmentos -8, -7 y -6 se usan en el cálculo de las mediciones temporales. Estas elecciones distintas de los segmentos donde se calculan las mediciones temporales conducen a la determinación de fluctuaciones de energía mucho más altas para los transitorios de impulsos y en consecuencia para desactivar la LTP para los transitorios de impulsos y activar la LTP para los transitorios tipo paso.

15 **Ejemplo 3:**

Sin embargo, en algunos casos el uso de las mediciones temporales puede no ser conveniente. El espectrograma en la Figura 18 y la forma de onda en la Figura 19 exhiben un extracto de aproximadamente 35 milisegundos desde el inicio de "Kalifornia" por Fatboy Slim.

20 La decisión LTP que depende de la Medición de planicidad temporal y el Cambio de energía máximo desactiva la LTP para este tipo de señal dado que detecta las grandes fluctuaciones de energía temporales.

25 Esta muestra es un ejemplo de ambigüedad entre los transitorios y el tren de pulsos que forman la señal de bajo tono.

Tal como puede observarse en la Figura 20, donde el extracto de 600 milisegundos de la misma señal se presenta, la señal contiene un transitorio de impulso muy corto repetido (el espectrograma se produce usando un FFT de longitud corta).

30 Tal como puede observarse en el mismo extracto de 600 milisegundos en la Figura 21 la señal se ve como si contuviera una señal muy armónica con tono bajo y cambiante (el espectrograma se produce usando un FFT de longitud larga).

35 Esta clase de señales se beneficia de la LTP dado que hay una clara estructura repetitiva (equivalente a la estructura armónica clara). Dada la clara fluctuación de energía (que puede observarse en la Figura 18, Figura 19 y Figura 20), la LTP sería desactivado por exceder el umbral para la Medición de planicidad temporal o para el Cambio de energía máximo. Sin embargo, en nuestra propuesta, la LTP se activa debido a que la correlación normalizada excede el umbral dependiente del retardo de tono ($norm_corr(curr) \leq 1,2 - T_{int}/L$).

40 Así, las realizaciones previas, revelan, por ejemplo, un concepto para una mejor decisión de filtro de armónicos para la codificación de audio. Debe advertirse que es posible introducir ciertos cambios a partir de lo revelado. En particular, según se ha indicado, la señal de audio 12 puede ser una señal de voz o música y puede ser reemplazada por una versión pre-procesada de la señal 12 a los fines de la estimación del tono, medición de la armonicidad, o análisis o medición de la estructura temporal. Además, la estimación de tono puede no limitarse a las mediciones de retardo de tonos, aunque, según advertirán los expertos en el arte, podrán realizarse mediante mediciones de una frecuencia fundamental, en el dominio de tiempo o espectral, que pueden fácilmente convertirse en un retardo de tono equivalente mediante una ecuación como "retardo de tono = frecuencia de muestreo / frecuencia de tono". En general, el estimador de tono 16 estima el tono de la señal de audio lo cual, a su vez, se manifiesta en el retardo de tono y frecuencia de tono.

55 A pesar de haberse descrito algunos aspectos en el contexto de un aparato, queda claro que estos aspectos representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa del método o una característica de una etapa del método. Análogamente, los aspectos descritos en el contexto de un paso del método además representan una descripción de un bloque, elemento o característica correspondiente de un aparato dado. Parte o todas las etapas del método pueden ejecutarse (u usarse) mediante un aparato como por ejemplo un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguno, uno o más de los pasos más importantes del método pueden ser ejecutados por este aparato.

60 La señal de audio codificada de la invención puede ser almacenada en un medio de almacenamiento digital o puede

ser transmitida por un medio de transmisión como un medio de transmisión inalámbrico o un medio de transmisión cableado como Internet.

5 De acuerdo con ciertos requerimientos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación puede ser ejecutada usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, con señales de control electrónicamente legibles almacenadas, que cooperan (o son capaces de cooperar) con un sistema informático programable de modo que se ejecute el método respectivo. En consecuencia, el medio de almacenamiento digital puede ser legible por ordenador.

10 Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos con señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, con el fin de ejecutar uno de los métodos descritos en el presente documento.

15 En general, las realizaciones de la presente invención pueden ser implementadas como un producto informático con un código de programa, el código de programa puede ejecutar uno de los métodos cuando se ejecuta en un ordenador. El código de programa puede por ejemplo estar almacenado en un soporte legible por máquina.

20 Otras realizaciones comprenden el programa informático para ejecutar uno de los métodos descritos, almacenado en un soporte legible por máquina.

En otras palabras, una realización de un método de la invención es, en consecuencia, un programa informático que posee un código para ejecutar uno de los métodos descritos, cuando el programa se ejecuta en un ordenador.

25 Una realización adicional de los métodos de la invención es, en consecuencia, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para ejecutar uno de los métodos descritos. El soporte de datos, el medio de almacenamiento digital o el medio grabado típicamente son tangibles y/o no transitorios.

30 Una realización adicional del método de la invención es, en consecuencia, un caudal de datos o una secuencia de señales que representan el programa informático para ejecutar uno de los métodos descritos en la presente. El caudal de datos o la secuencia de señales pueden, por ejemplo, configurarse para ser transferidos por una conexión de comunicación de datos, por ejemplo, mediante Internet.

35 Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los métodos descritos en el presente documento.

Una realización adicional comprende un ordenador que tiene instalado el programa para ejecutar uno de los métodos descritos en el presente documento.

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

45 En algunas realizaciones, un dispositivo lógico programable (por ejemplo, un conjunto de compuertas programables por campo) puede usarse para ejecutar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, un conjunto de compuertas programables por campo puede cooperar con un microprocesador con el fin de ejecutar uno de los métodos descritos en la presente. En general, los métodos preferentemente se ejecutan mediante cualquier aparato.

50 Las realizaciones antes descritas son meramente ilustrativas de los principios de la presente invención. Queda entendido que los expertos en el arte advertirán modificaciones y variantes a lo revelado. En consecuencia, el invento estará solamente limitado por las reivindicaciones adjuntas y no por detalles específicos expuestos con fines descriptivos e ilustrativos de las realizaciones del mismo.

REIVINDICACIONES

1. Aparato (10) para ejecutar un control dependiente de la armonicidad de una herramienta de filtro de armónicos de un códec de audio, que comprende:
- 5 un estimador de tono (16) configurado para determinar un tono (18) de una señal de audio (12) a ser procesada por el códec de audio;
- un medidor de armonicidad (20) configurado para determinar una medición (22) de armonicidad de la señal de audio (12) usando el tono (18);
- 10 un analizador de estructura temporal (24) configurado para determinar, de acuerdo con el tono (18), por lo menos una medición de estructura temporal (26) que mide una característica de una estructura temporal de la señal de audio (12);
- un controlador (28) configurado para controlar la herramienta de filtro de armónicos (30) de acuerdo con la medición de estructura temporal (26) y la medición (22) de armonicidad,
- 15 en el que el analizador de estructura temporal (24) está configurado para determinar la por lo menos una medición de estructura temporal (26) dentro de una región temporal temporalmente colocada de acuerdo con el tono (18), y en el que el analizador de estructura temporal (24) está configurado para disponer un extremo de encabezado temporalmente pasado (38) de la región temporal, o de una región de mayor influencia sobre la determinación de la medición de estructura temporal (26), de acuerdo con el tono (18).
- 20 2. Aparato de acuerdo con la reivindicación 1, en el que el medidor de armonicidad (20) está configurado para determinar la medición (22) de armonicidad computando una correlación normalizada de la señal de audio (12) o una versión pre-modificada de la misma en o alrededor de un retardo de tono del tono (18).
3. Aparato de acuerdo con la reivindicación 1 o 2, en el que el estimador de tono (16) está configurado para
- 25 determinar el tono (18) en etapas que comprenden una primera etapa y una segunda etapa.
4. Aparato de acuerdo con la reivindicación 3, en el que el estimador de tono (16) está configurado para, en la primera etapa, determinar una estimación preliminar del tono en un dominio muestreado hacia abajo de una primera
- 30 velocidad de muestreo y, dentro de la segunda etapa, refinar la estimación preliminar del tono a una segunda velocidad de muestreo, mayor que la primera velocidad de muestreo.
5. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el estimador de tono (16) está configurado para determinar el tono (18) usando auto correlación.
- 35 6. Aparato de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el que el analizador de estructura temporal (24) está configurado para disponer el extremo de encabezado temporalmente pasado (38) de la región temporal o, de la región de mayor influencia sobre la determinación de la medición de la estructura temporal, de modo tal que el extremo de encabezado temporalmente pasado (38) de la región temporal o, de la región de mayor influencia sobre
- 40 la determinación de la medición de la estructura temporal, se desplace en una dirección pasada en una cantidad temporal que aumenta monotónicamente con una disminución del tono (18).
7. Aparato de acuerdo con cualquiera de las reivindicación 1 a 6, en el que el analizador de estructura temporal (24) está configurado para disponer un extremo de encabezado temporalmente futuro (40) de la región temporal (36) o,
- 45 de la región de mayor influencia sobre la determinación de la medición de estructura temporal (26), de acuerdo con la estructura temporal de la señal de audio (12) dentro de una región candidato temporal que se extiende desde el extremo de encabezado temporalmente pasado (38) de la región temporal, o de la región de mayor influencia sobre la determinación de la medición de la estructura temporal, hasta un extremo de encabezado temporalmente futuro (44) de un cuadro actual (34a).
- 50 8. Aparato de acuerdo con la reivindicación 7, en el que el analizador de estructura temporal (24) está configurado para usar una amplitud o relación entre las muestras de energía máxima y mínima dentro de la región candidato temporal con el fin de disponer el extremo de encabezado temporalmente futuro (40) de la región temporal (36) o, de la región de mayor influencia sobre la determinación de la medición de estructura temporal (26).
- 55 9. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el controlador (28) comprende una lógica (120) configurada para determinar si una condición predeterminada es satisfecha por la por lo menos única medición de estructura temporal (26) y la medición (22) de la armonicidad para obtener un resultado del chequeo; y
- 60 un interruptor (124) configurado para activar y desactivar la herramienta de filtro de armónicos (30) de acuerdo con el resultado de la prueba.
10. Aparato de acuerdo con la reivindicación 9, en el que la por lo menos única medición de estructura temporal (26) mide una variación de energía promedio o máxima de la señal de audio dentro de la región temporal y la lógica está configurada de modo tal que la condición predeterminada sea satisfecha si

ambas la por lo menos única medición de estructura temporal (26) es menor que un primer umbral predeterminado y la medición (22) de la armonicidad es, para un cuadro actual y/o un cuadro previo, por sobre un segundo umbral.

11. Aparato de acuerdo con la reivindicación 10, en el que la lógica (120) está configurada de modo tal que la condición predeterminada también sea satisfecha si

5 la medición (22) de la armonicidad es, para un cuadro actual, por sobre un tercer umbral, y la medición de armonicidad es, para un cuadro actual y/o un cuadro previo, por sobre un cuarto umbral que disminuye con el aumento del retardo de tono del tono (18).

12. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el controlador (28) está configurado para controlar la herramienta de filtro de armónicos (30)

10 enviando explícitamente una señal de control mediante un caudal de datos de códec de audio a un lado de decodificación; o

15 enviando explícitamente una señal de control mediante un caudal de datos de códec de audio a un lado de decodificación para controlar un post-filtro en el lado de decodificación y, en línea con el control del post-filtro en el lado de decodificación, controlar un pre-filtro en un lado de codificador.

13. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el analizador de estructura temporal (24) está configurado para determinar la por lo menos única medición de estructura temporal (26) de una forma espectralmente discriminante con el fin de obtener un valor de la por lo menos única medición de estructura temporal (26) por banda espectral de una pluralidad de bandas espectrales.

14. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el controlador (28) está configurado para controlar la herramienta de filtro de armónicos (30) en unidades de cuadros, y el analizador de estructura temporal (24) está configurado para muestrear la energía de la señal de audio (12) a una velocidad de muestreo superior a la velocidad de muestreo de los cuadros con el fin de obtener muestras de energía de la señal de audio y determinar la por lo menos única medición de estructura temporal (26) sobre la base de las muestras de energía.

15. Aparato de acuerdo con la reivindicación 14, en el que el analizador de estructura temporal (24) está configurado para determinar la por lo menos única medición de estructura temporal (26) dentro de una región temporal temporalmente colocada de acuerdo con el tono (18) y el analizador de estructura temporal (24) está configurado para determinar la por lo menos única medición de estructura temporal (26) sobre la base de las muestras de energía computando un conjunto de valores de cambio de energía que miden un cambio entre los pares de muestras de energía inmediatamente consecutivos de las muestras de energía dentro de la región temporal y someter el conjunto de valores de cambio de energía a una función escalar que incluye un operador máximo o una suma de sumandos que dependen respectivamente de exactamente un conjunto de valores de cambio de energía.

16. Aparato de acuerdo con cualquiera de las reivindicaciones 14 y 15, en el que el analizador de espectro temporal (24) está configurado para realizar el muestreo de la energía de la señal de audio (12) dentro de un dominio de filtro de paso alto.

17. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el estimador de tono (16), el medidor de armonicidad (20) y el analizador de estructura temporal (24) ejecutan su determinación en base a distintas versiones de la señal de audio (12) que incluye la señal de audio original y alguna versión pre-modificada de la misma.

18. Aparato de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el controlador (28) está configurado para, al controlar la herramienta de filtro de armónicos (30), de acuerdo con la medición de estructura temporal (26) y la medición (22) de la armonicidad

50 alternar entre activar y desactivar un pre-filtro y/o un post-filtro de la herramienta de filtro de armónicos (30), o gradualmente adaptar una intensidad de filtro del pre-filtro y/o el post-filtro de la herramienta de filtro de armónicos (30),

55 en el que la herramienta de filtro de armónicos (30) es de un enfoque pre-filtro más post-filtro y el pre-filtro de la herramienta de filtro de armónicos (30) está configurado para incrementar el ruido de cuantización dentro de un armónico del tono de la señal de audio y el post-filtro de la herramienta de filtro de armónicos (30) está configurado para reformular un espectro transmitido en consecuencia, o la herramienta de filtro de armónicos (30) es de un enfoque post-filtro solamente y el post-filtro de la herramienta de filtro de armónicos (30) está configurado para filtrar el ruido de cuantización que ocurre entre el armónico del tono de la señal de audio.

60 19. Codificador de audio o decodificador de audio, que comprende una herramienta de filtro de armónicos (30) y el aparato para ejecutar un control dependiente de la armonicidad de la herramienta de filtro de armónicos de acuerdo con cualquiera de las reivindicaciones anteriores.

20. Sistema que comprende

- un aparato (10) para ejecutar un control dependiente de la armonicidad de una herramienta de filtro de armónicos de acuerdo con cualquiera de las reivindicaciones 14 a 16, y
un detector de transitorio configurado para detectar transitorios en una señal de audio a procesar por el códec de audio sobre la base de las muestras de energía.
- 5 21. Codificador basado en transformada que comprende el sistema de la reivindicación 20, configurado para alternar un bloque de transformada y/o una longitud de superposición de acuerdo con los transitorios detectados.
- 10 22. Codificador de audio que comprende el sistema de la reivindicación 20, configurado para soportar el cambio entre un modo de excitación codificado por transformada y un modo de predicción lineal excitado por código de acuerdo con los transitorios detectados.
- 15 23. Codificador de audio de acuerdo con la reivindicación 22, configurado para cambiar entre un bloque de transformada y/o longitud de superposición en el modo de excitación codificado por transformada de acuerdo con los transitorios detectados.
- 20 24. Método (10) para ejecutar un control dependiente de la armonicidad de una herramienta de filtro de armónicos de un códec de audio, que comprende
determinar un tono (18) de una señal de audio (12) a ser procesada por el códec de audio; determinar una medición (22) de la armonicidad de la señal de audio (12) usando el tono (18);
determinar, de acuerdo con el tono (18), por lo menos una medición de estructura temporal (26) que mide una característica de una estructura temporal de la señal de audio;
controlar la herramienta de filtro de armónicos (30) de acuerdo con la medición de estructura temporal (26) y la medición (22) de armonicidad,
25 en el que la por lo menos una medición de estructura temporal (26) se determina dentro de una región temporal temporalmente colocada de acuerdo con el tono (18), y
en el que un extremo de encabezado temporalmente pasado (38) de la región temporal, o de una región de mayor influencia sobre la determinación de la medición de estructura temporal (26), está colocado de acuerdo con el tono (18).
- 30 25. Programa informático que posee un código para ejecutar, al ejecutarse en un ordenador, un método de acuerdo con la reivindicación 24.

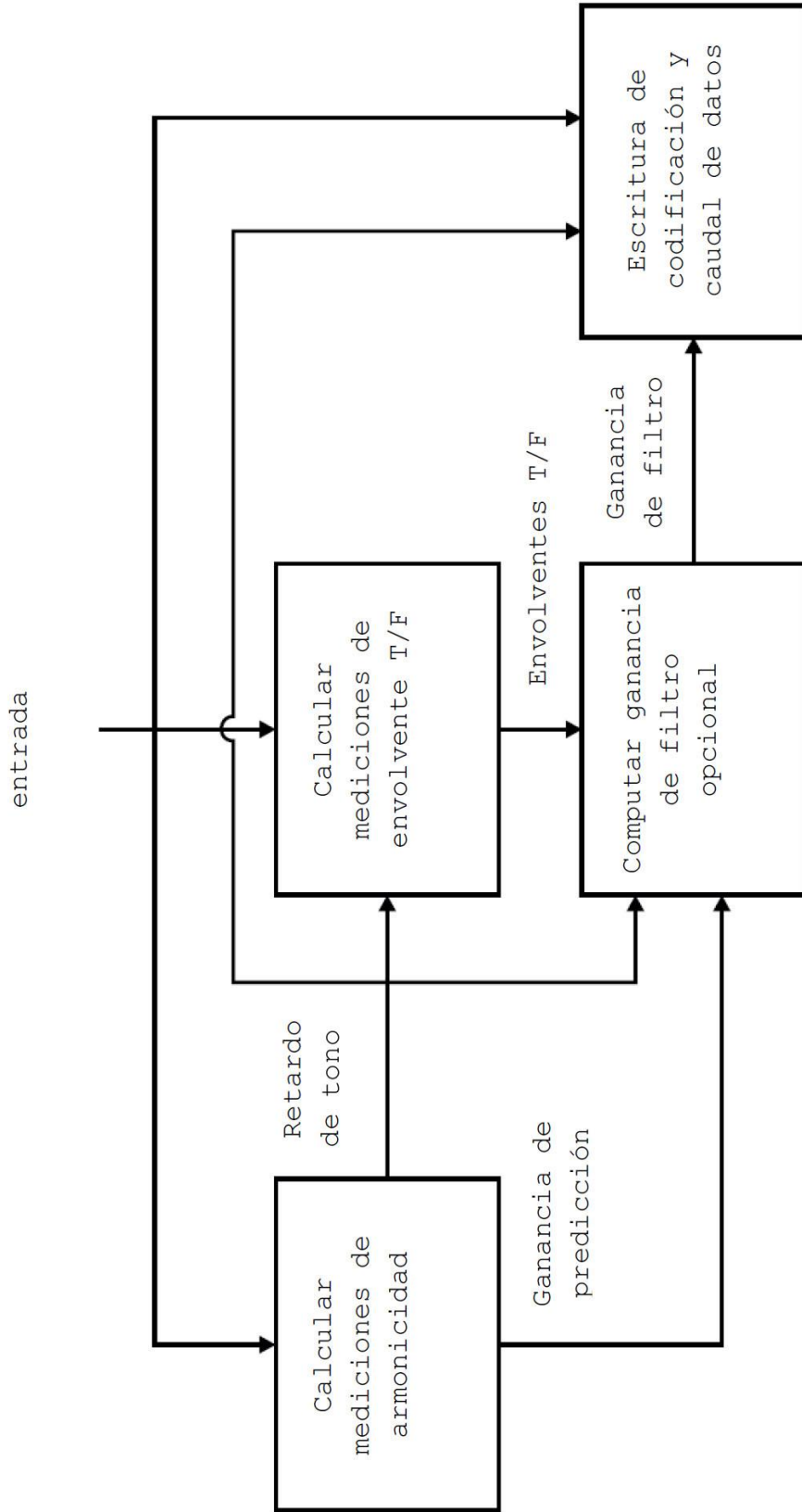


FIG 1

```

enableLTP =
(bitrate < b1 && tcxMode == TCX_20 && (norm_corr(curr) * norm_corr(prev)) > 0,25 && tempFlatness < 3,5) ||
(bitrate >= b1 && tcxMode == TCX_10 && max(norm_corr(curr), norm_corr(prev)) > 0,5 && maxEnergyChange < 3,5) ||
(bitrate >= b1 && norm_corr(curr) > 0,44 && norm_corr(curr) > 1,2 * Tint/L) ||
(bitrate >= b1 && tcxMode == TCX_20 && (norm_corr(curr) > 0,44 &&
(tempFlatness < 6,0 || (tempFlatness < 7,0 && maxEnergyChange < 22,0)));

```

FIG 2

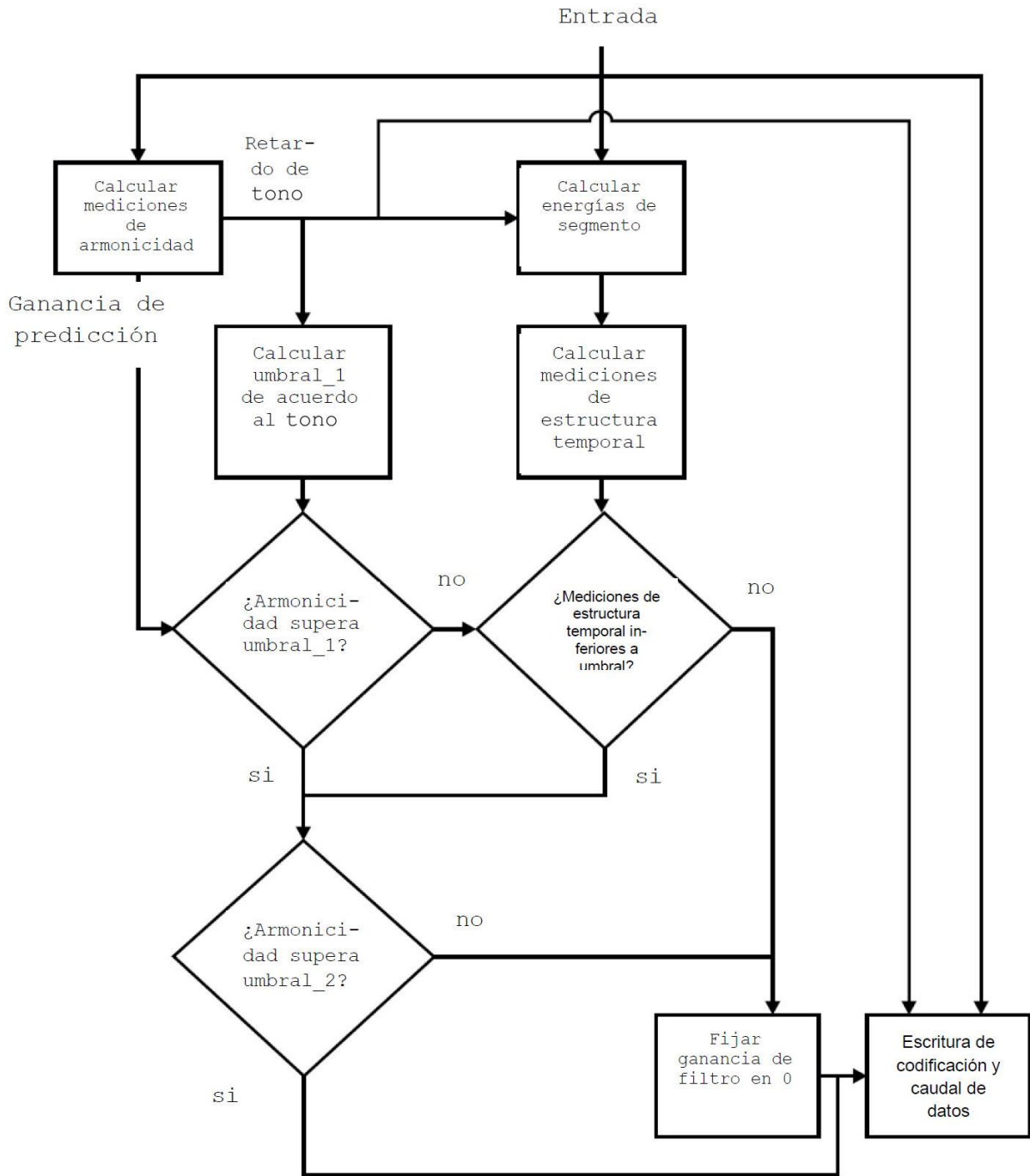


FIG 3

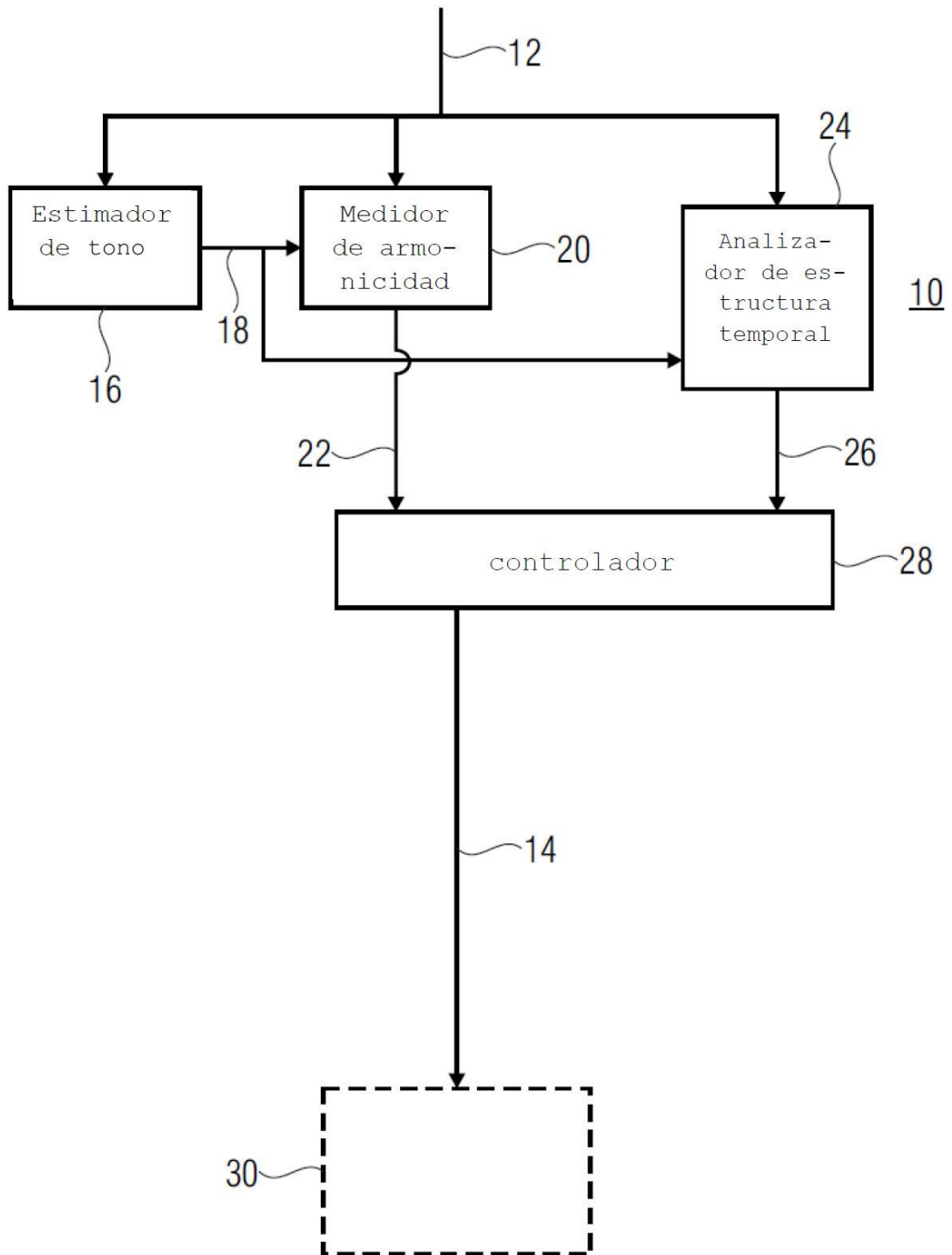


FIG 4

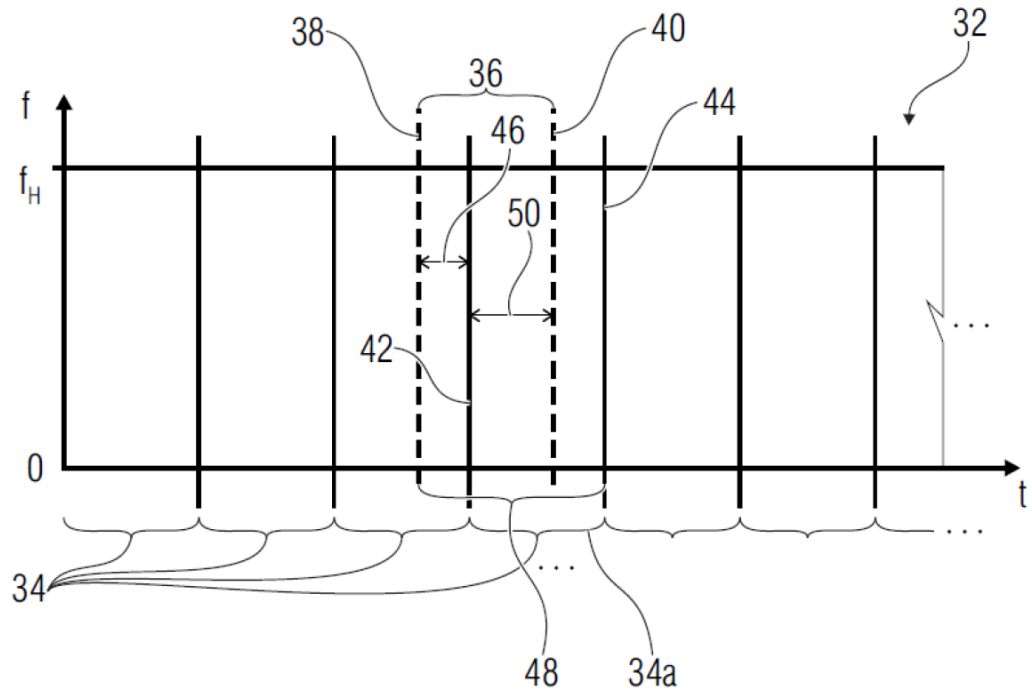


FIG 5

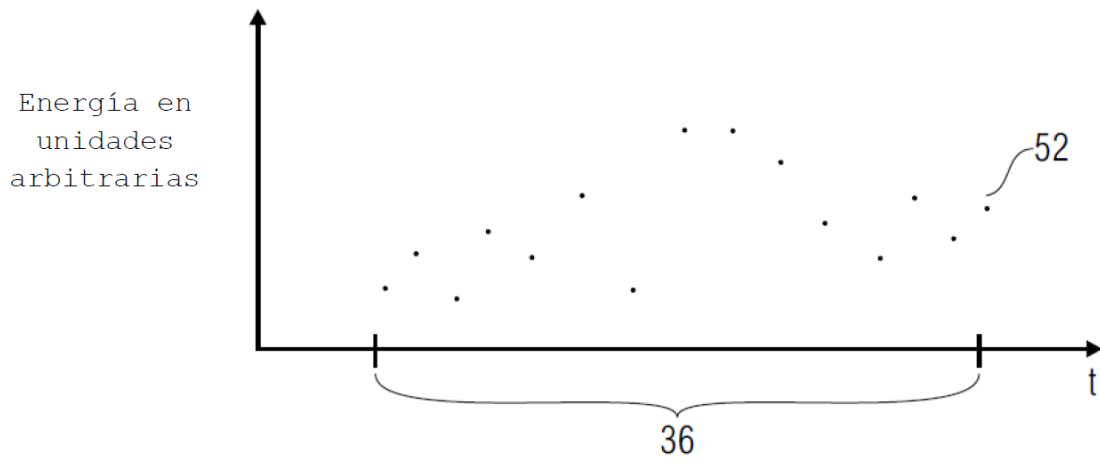


FIG 6

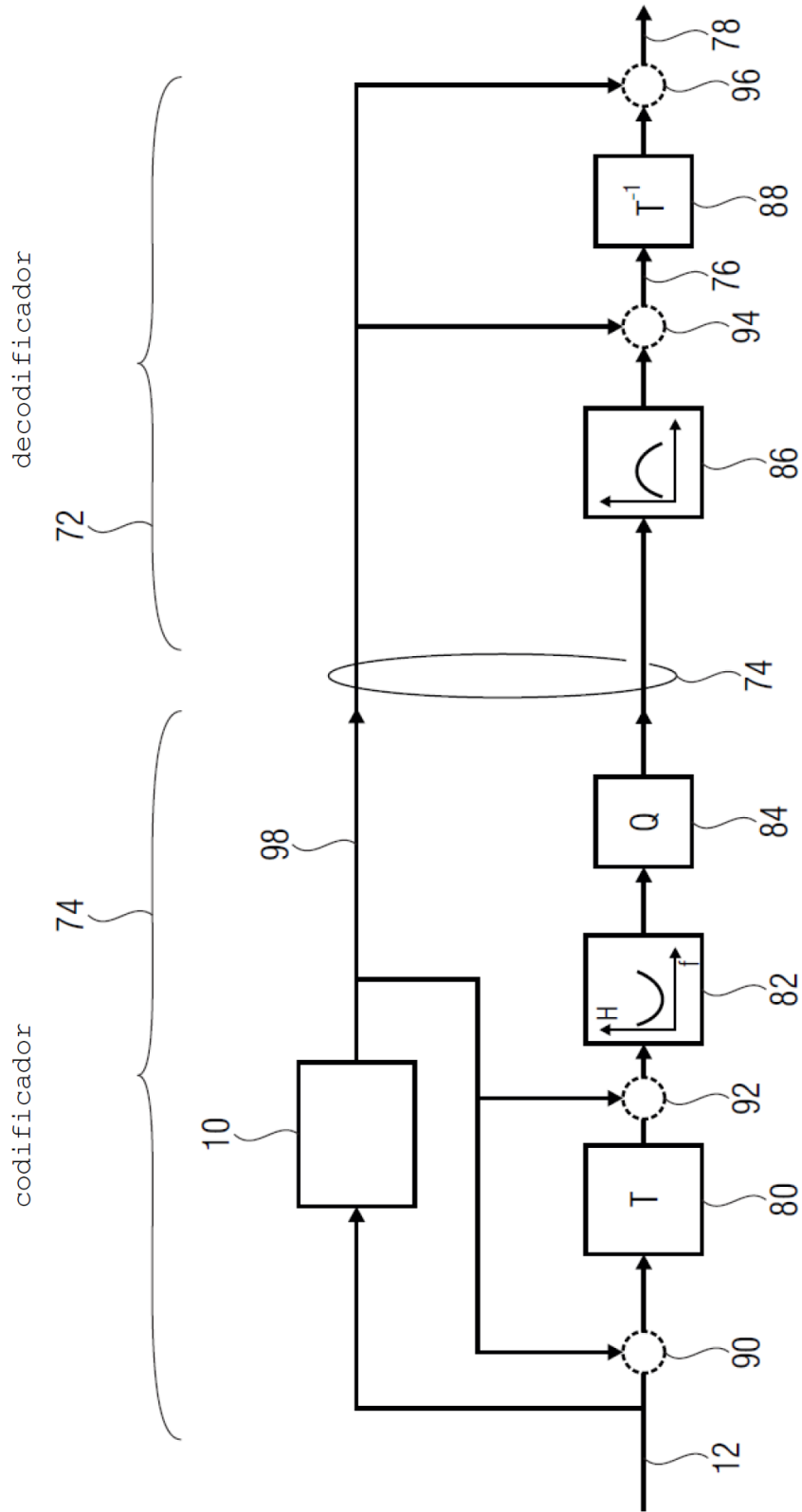


FIG 7

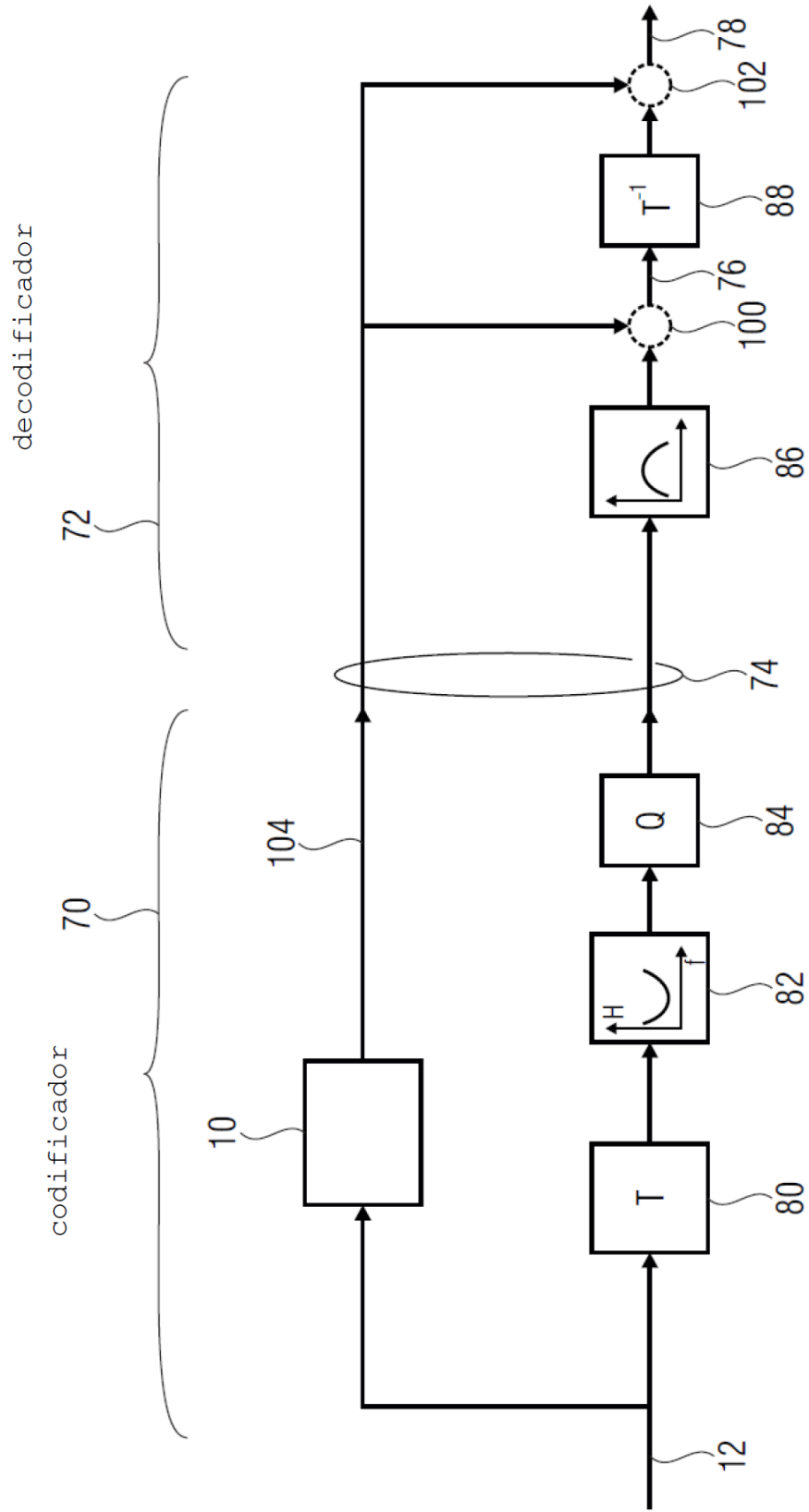


FIG 8

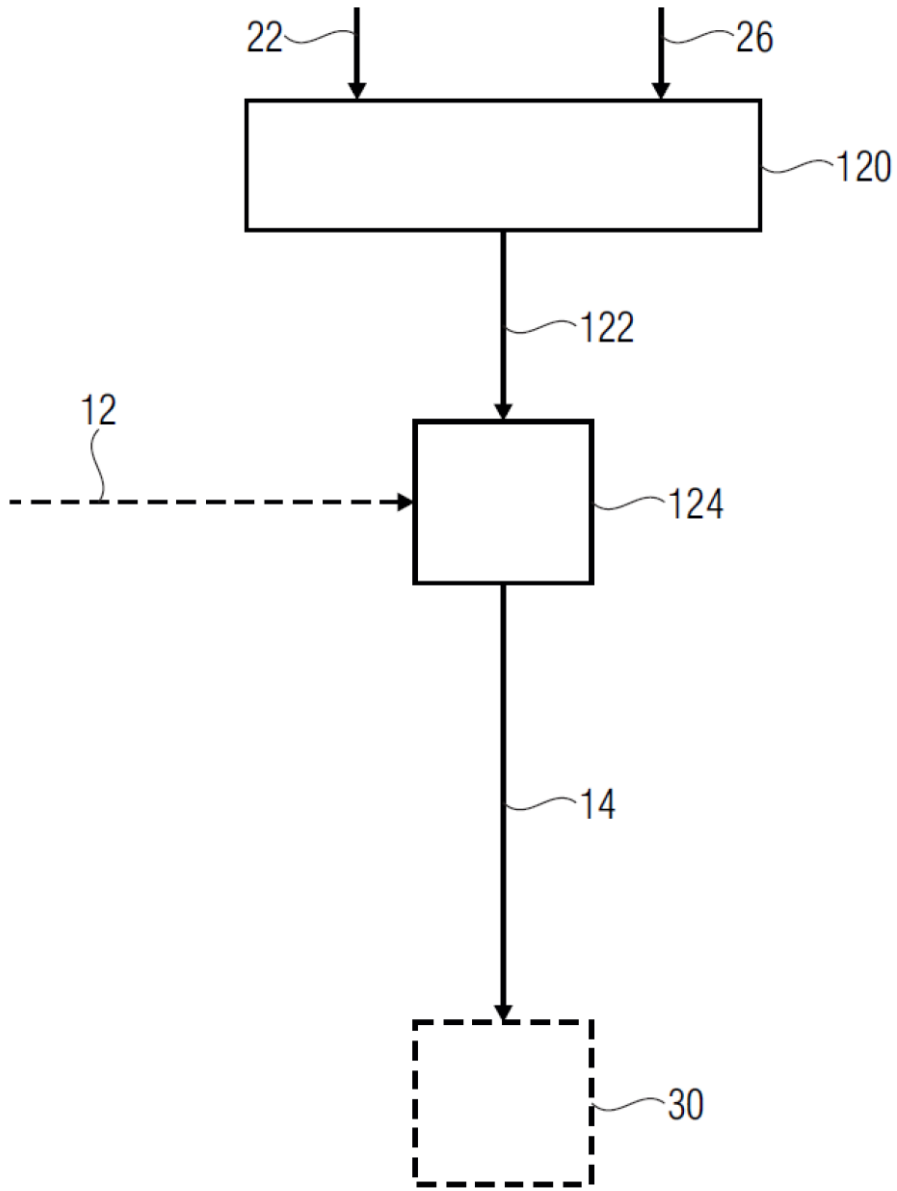
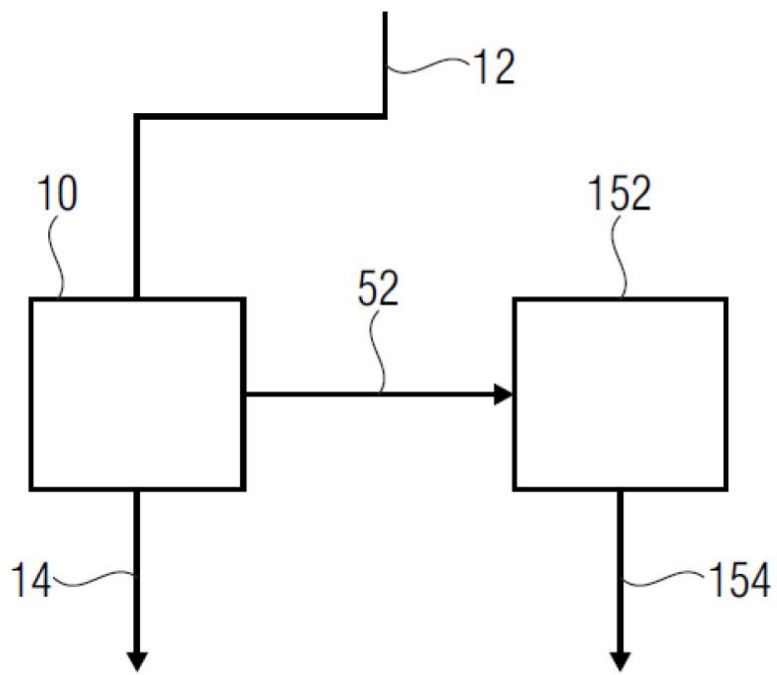


FIG 9



150

FIG 10

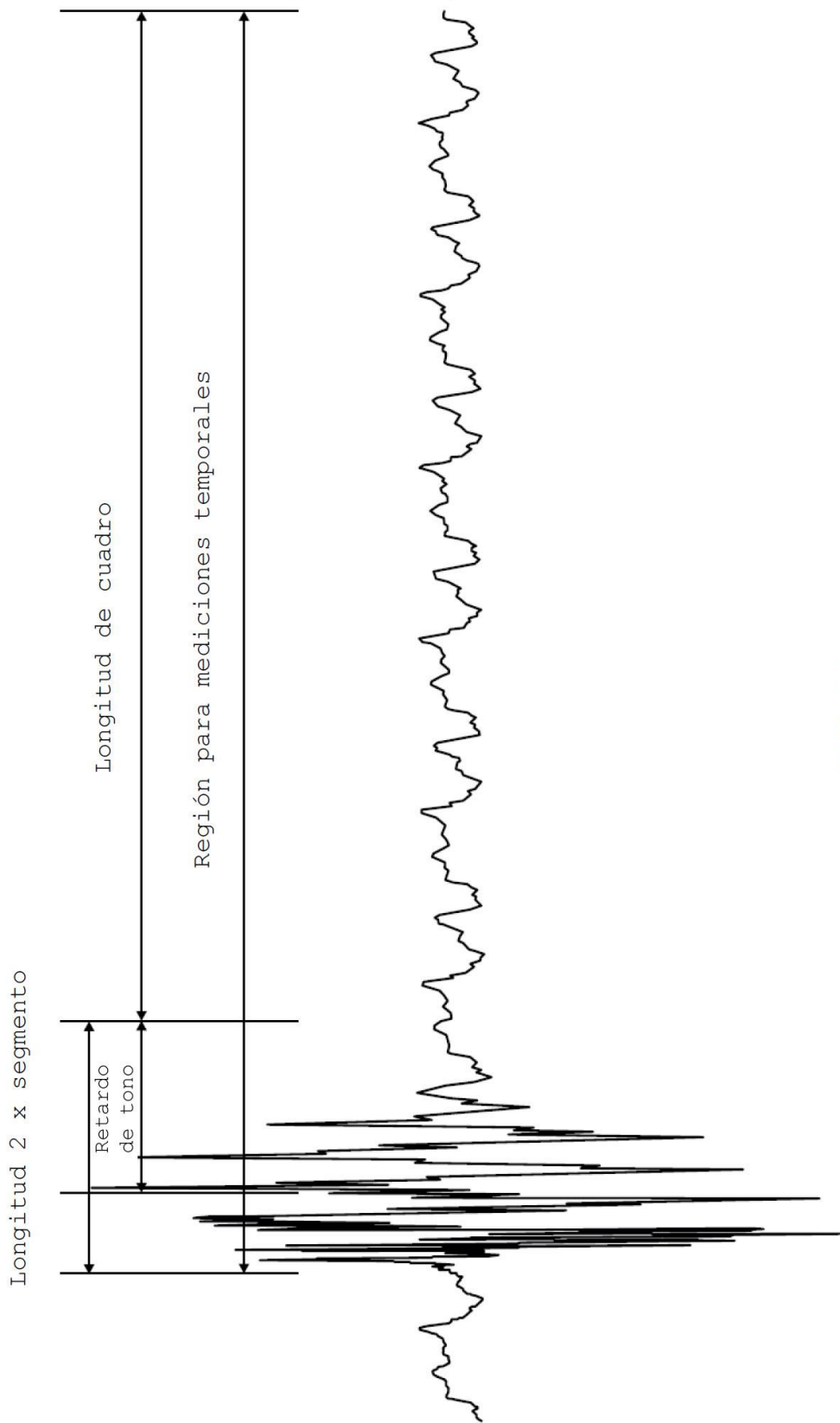


FIG 11

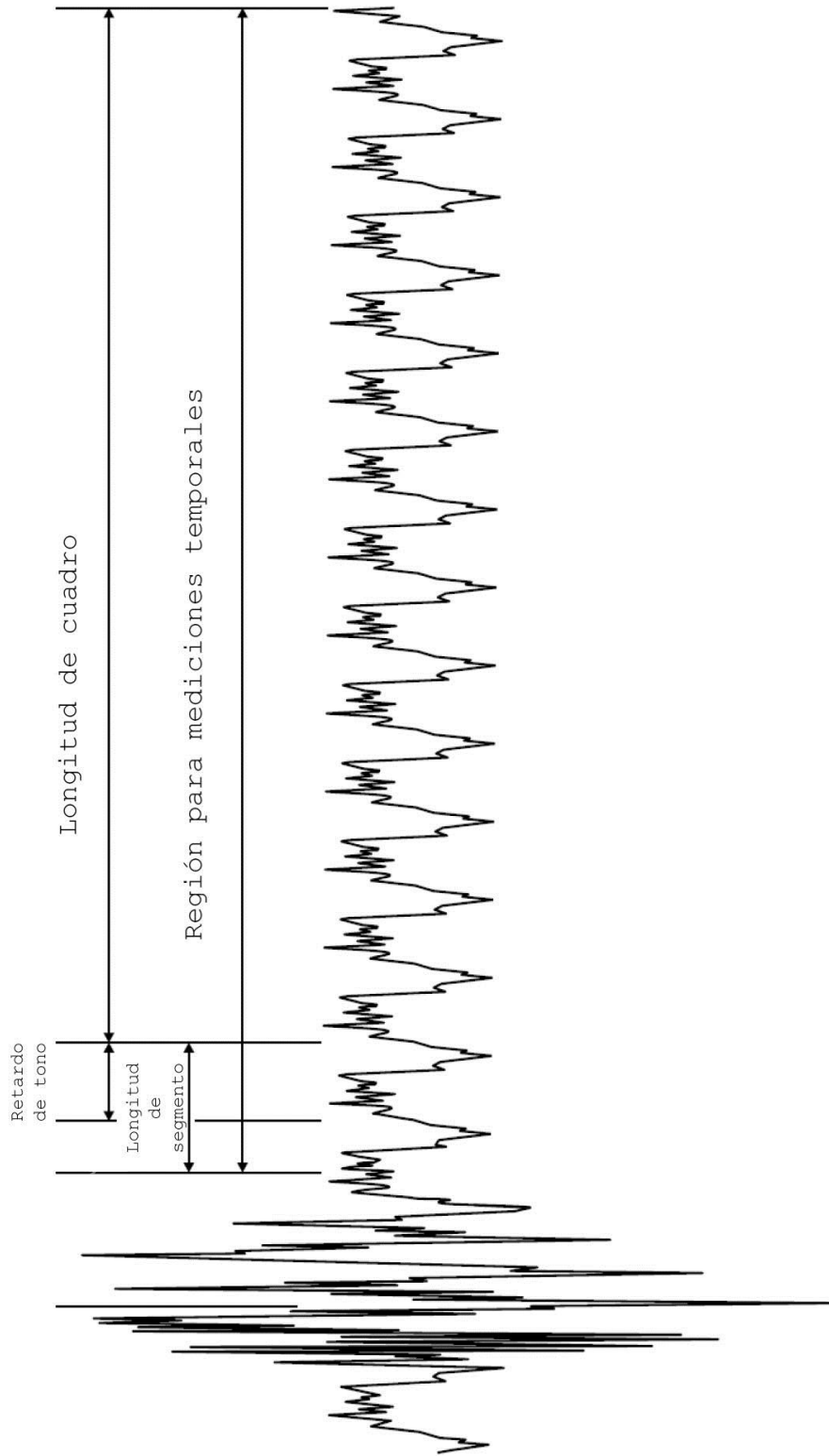


FIG 12

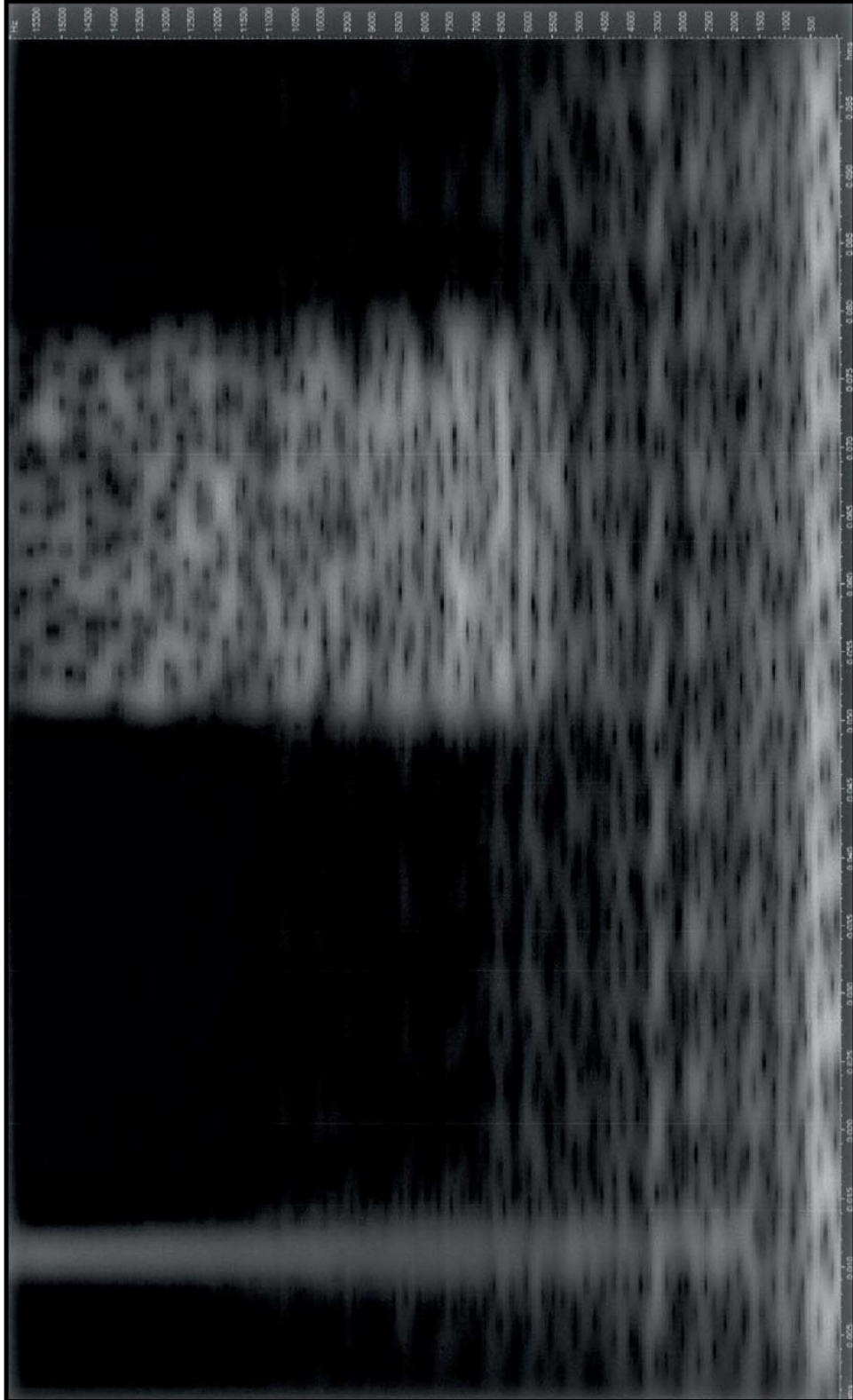


FIG 13

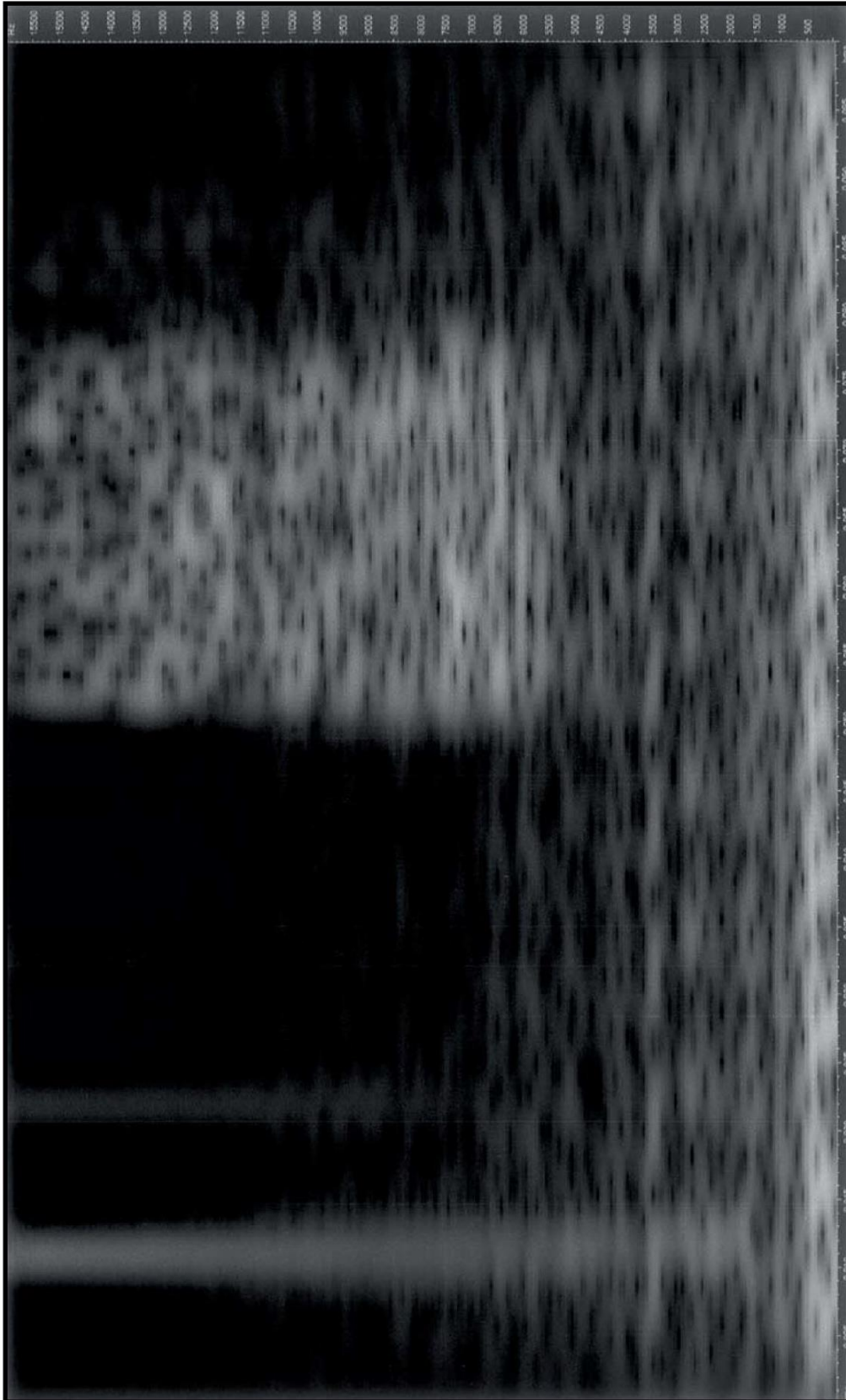


FIG 14

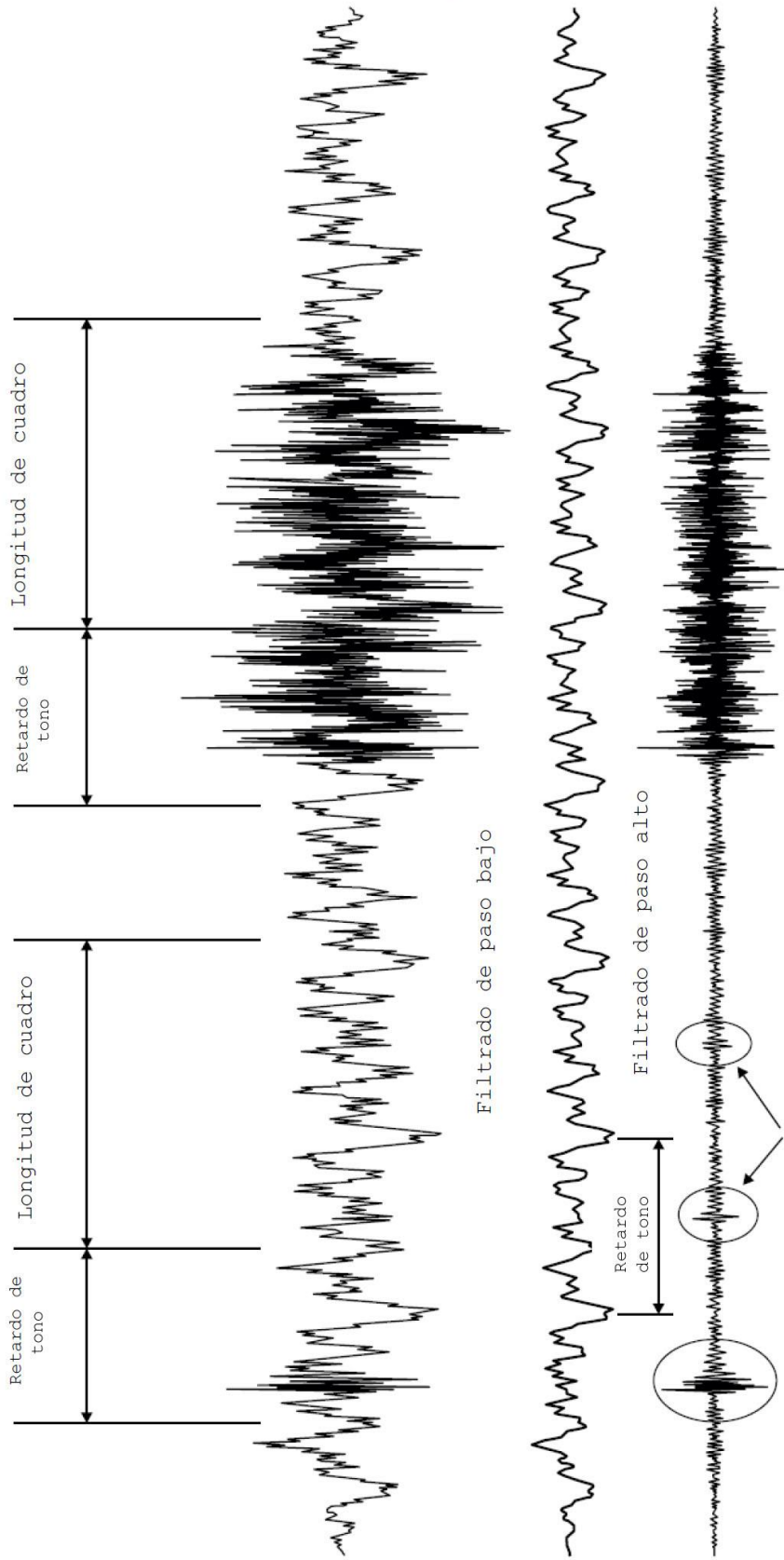


FIG 15

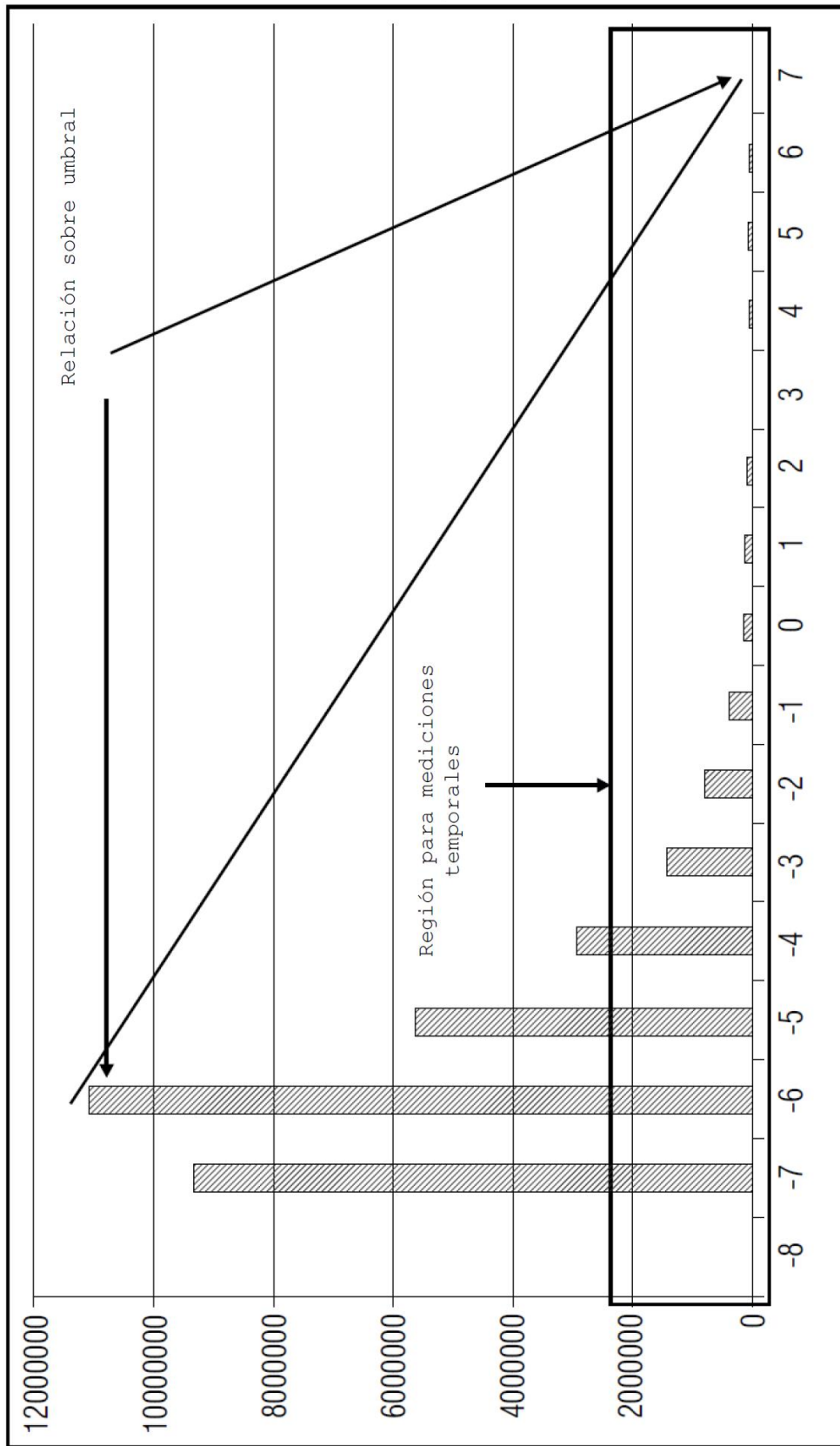


FIG 16

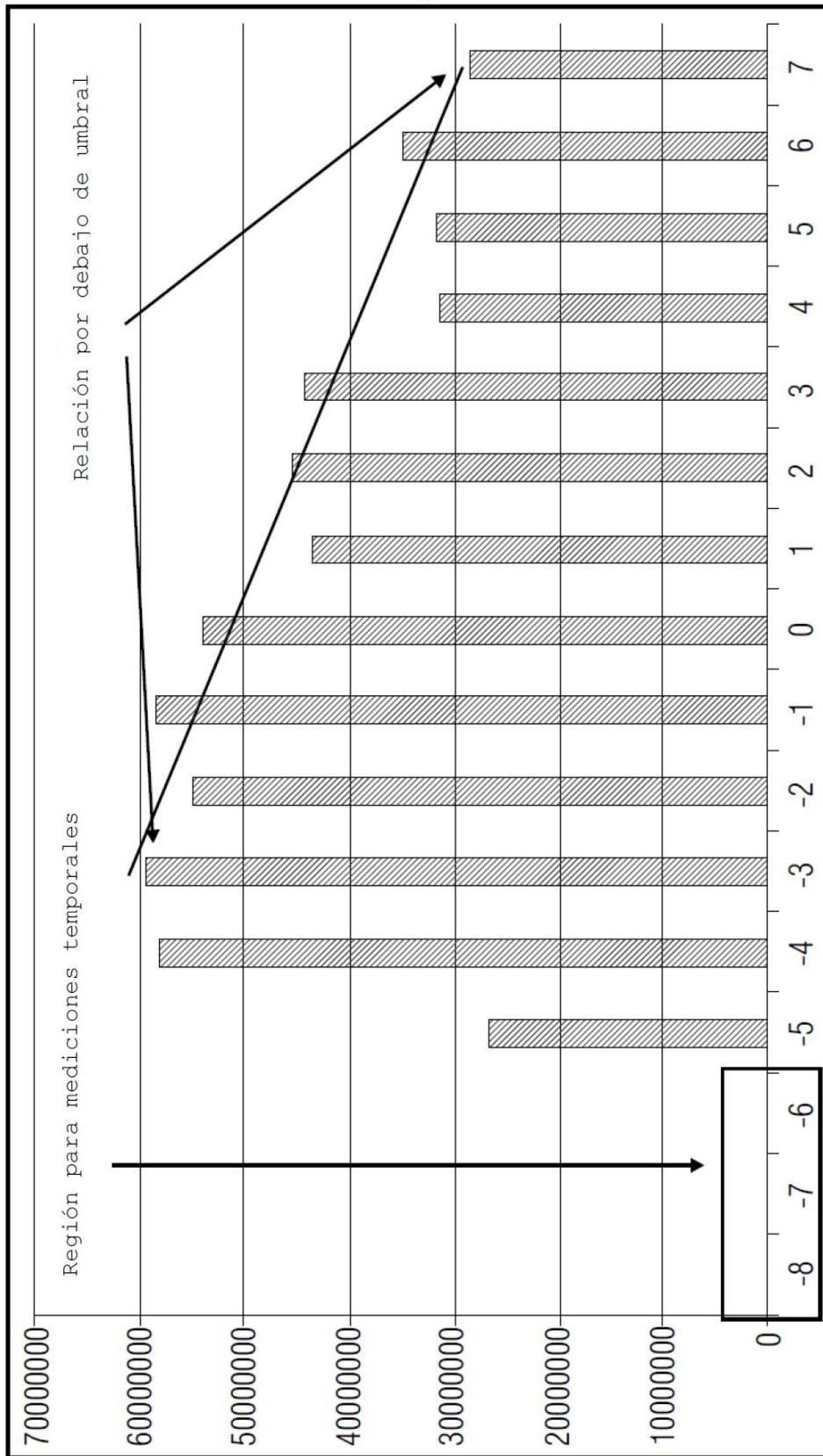


FIG 17

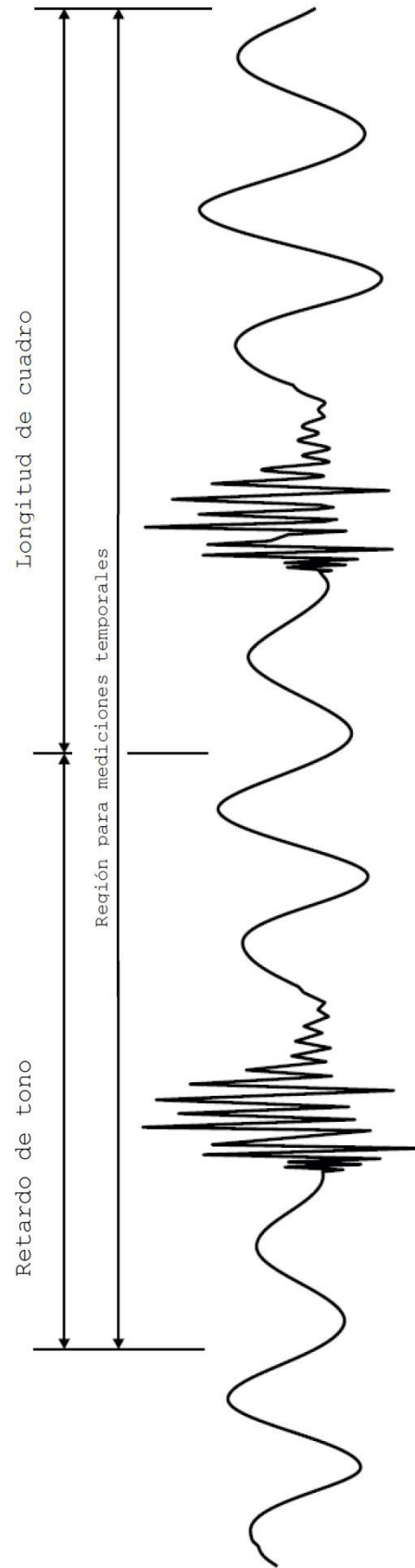


FIG 19

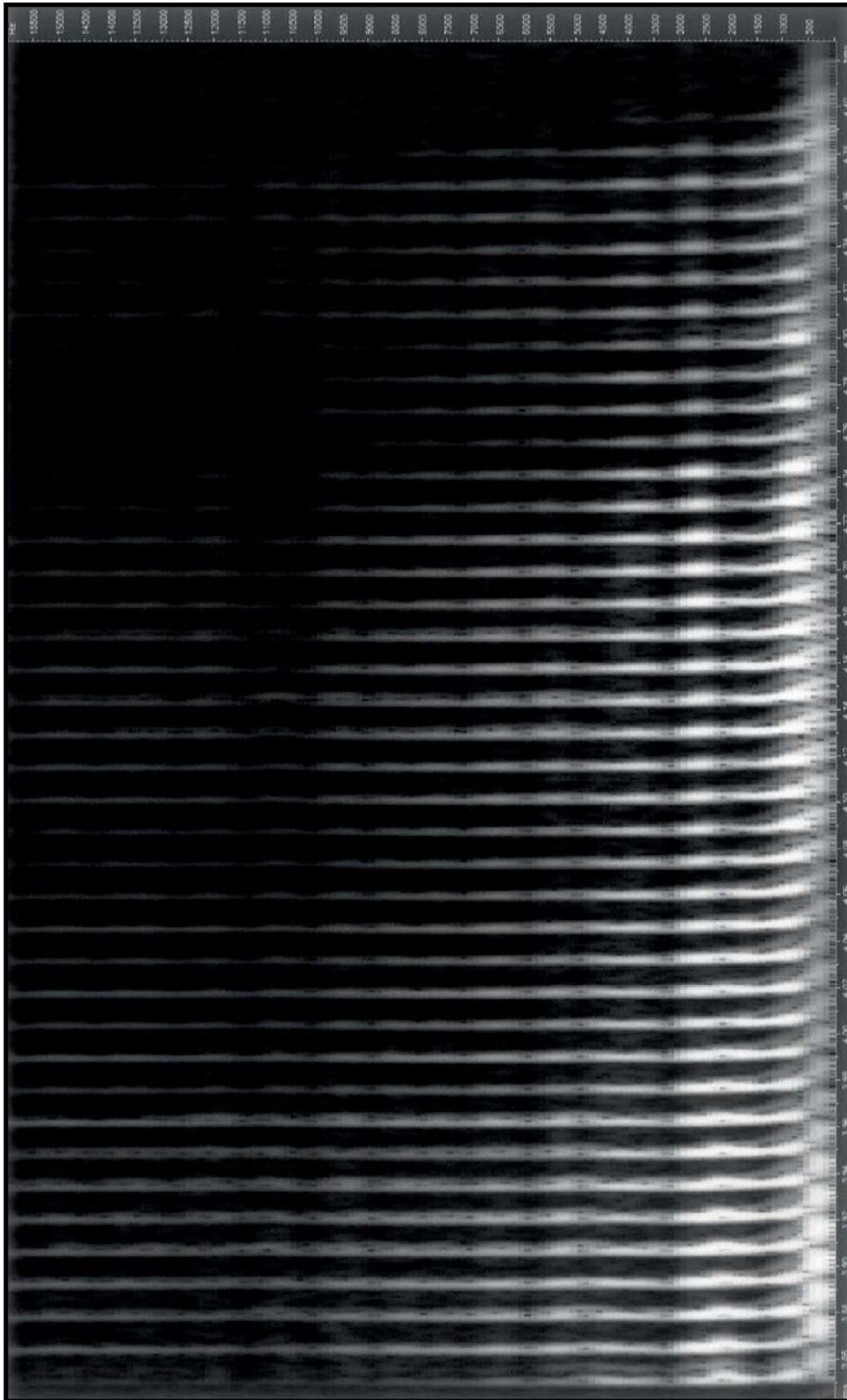


FIG 20

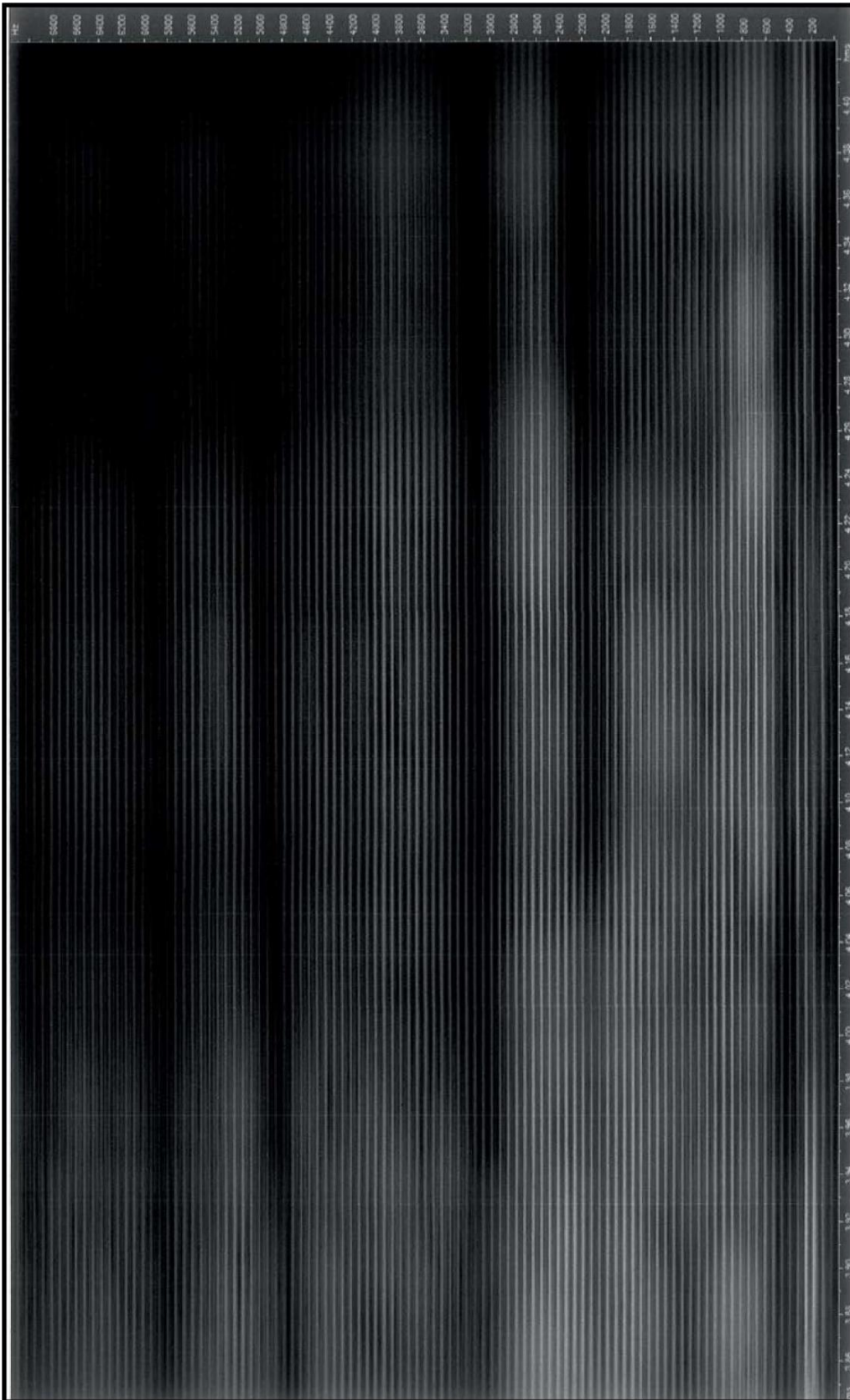


FIG 21