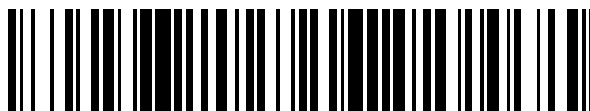


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 315**

51 Int. Cl.:

G10L 19/032 (2013.01)

G10L 19/083 (2013.01)

G10L 21/0232 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.07.2011** **PCT/SE2011/050899**

87 Fecha y número de publicación internacional: **13.09.2012** **WO12121637**

96 Fecha de presentación y número de la solicitud europea: **04.07.2011** **E 11860420 (6)**

97 Fecha y número de publicación de la concesión europea: **21.06.2017** **EP 2681734**

54 Título: **Corrección de ganancia post-cuantificación en codificación de audio**

30 Prioridad:

04.03.2011 US 201161449230 P

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

08.11.2017

73 Titular/es:

TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)
(100.0%)
164 83 Stockholm, SE

72 Inventor/es:

NORVELL, ERIK y
GRANCHAROV, VOLODYA

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 641 315 T3

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

DESCRIPCIÓN

Corrección de ganancia post-cuantificación en codificación de audio

5 Campo técnico

La presente tecnología se refiere a la corrección de ganancia en la codificación de audio basándose en esquemas de cuantificación donde la cuantificación se divide en una representación de ganancia y una representación de forma, denominada codificación de audio en forma de ganancia, y especialmente para corrección de ganancia post-cuantificación.

Antecedentes

Se espera que los servicios de telecomunicaciones modernos manejen muchos tipos diferentes de señales de audio. Si bien el contenido de audio principal son las señales de voz, existe el deseo de manejar señales más generales como música y mezclas de música y voz. Aunque la capacidad en las redes de telecomunicaciones está aumentando continuamente, sigue siendo de gran interés limitar el ancho de banda requerido por canal de comunicación. En las redes móviles, los anchos de banda de transmisión más pequeños para producen un consumo de energía menor tanto en el dispositivo móvil como en la estación base. Esto se traduce en ahorro de energía y de costes para el operador móvil, mientras que el usuario final experimentará una duración prolongada de la batería y un aumento del tiempo de conversación. Además, con menos ancho de banda consumido por usuario, la red móvil puede atender a un mayor número de usuarios en paralelo.

Hoy en día, la tecnología de compresión dominante para servicios de voz móviles es CELP (predicción lineal excitada por código), que logra una buena calidad de audio para la voz en anchos de banda bajos. Es ampliamente utilizado en códecs desplegados como AMR (multivelocidad adaptativa), AMR-WB (ancho de banda de velocidad adaptativa) y GSM-EFR (sistema global para comunicaciones móviles – velocidad completa mejorada). Sin embargo, para señales de audio generales, como la música, la tecnología CELP tiene un rendimiento bajo. Estas señales pueden a menudo ser mejor representadas mediante el uso de codificación basada en la transformada de frecuencia, por ejemplo los códecs ITU-T G.722.1 [1] y G.719 [2]. Sin embargo, los códecs de dominio de transformada operan generalmente a una velocidad de bits mayor que los códecs de voz. Existe una brecha entre los dominios de voz y de audio general en términos de codificación y es deseable aumentar el rendimiento de los códecs de dominio de transformada a velocidades de bits más bajas.

Los códecs de dominio de transformada requieren una representación compacta de los coeficientes de transformada del dominio frecuencia. Estas representaciones a menudo se basan en la cuantificación vectorial (VQ), donde los coeficientes se codifican en grupos. Entre los diversos métodos para la cuantificación vectorial está la VQ en forma de ganancia. Este enfoque aplica la normalización a los vectores antes de codificar los coeficientes individuales. El factor de normalización y los coeficientes normalizados se denominan la ganancia y la forma del vector, que puede codificarse por separado. La estructura en forma de ganancia tiene muchos beneficios. Al dividir la ganancia y la forma, el códec se puede adaptar fácilmente a diferentes niveles de entrada de fuente mediante el diseño del cuantificador de ganancia. También es beneficioso desde una perspectiva perceptual donde la ganancia y la forma pueden tener diferente importancia en diferentes regiones de frecuencia. Por último, la división en forma de ganancia simplifica el diseño del cuantificador y lo hace menos complejo en términos de memoria y recursos computacionales en comparación con un cuantificador de vector sin restricciones. En la figura 1 se puede ver una descripción funcional de un cuantificador en forma de ganancia.

Si se aplica a un espectro de dominio frecuencia, la estructura en forma de ganancia se puede utilizar para formar una envolvente espectral y una representación de estructura fina. La secuencia de valores de ganancia forma la envolvente del espectro mientras que los vectores de forma dan el detalle espectral. Desde una perspectiva perceptual es beneficioso dividir el espectro utilizando una estructura de banda no uniforme que sigue la resolución de frecuencia del sistema auditivo humano. Esto generalmente significa que anchos de banda estrechos se utilizan para frecuencias bajas, mientras que anchos de banda más grandes se utilizan para altas frecuencias. La importancia perceptual de la estructura fina espectral varía con la frecuencia, pero también depende de las características de la propia señal. Los codificadores de transformada emplean a menudo un modelo auditivo para determinar las partes importantes de la estructura fina y asignar los recursos disponibles a las partes más importantes. La envolvente espectral se utiliza a menudo como entrada a este modelo auditivo. El codificador de forma cuantifica los vectores de forma utilizando los bits asignados. Véase la figura 2 para un ejemplo de un sistema de codificación basado en la transformada con un modelo auditivo.

Dependiendo de la precisión del cuantificador de forma, el valor de ganancia utilizado para reconstruir el vector puede ser más o menos apropiado. Especialmente cuando los bits asignados son pocos, el valor de ganancia se aleja del valor óptimo. Una forma de resolver esto es codificar un factor de corrección que explique el desajuste de ganancia después de la cuantificación de forma. Otra solución es codificar primero la forma y luego calcular el factor de ganancia óptima dada la forma cuantificada.

La solución para codificar un factor de corrección de ganancia después de la cuantificación de forma puede consumir una velocidad de bits considerable. Si la velocidad es ya baja, esto significa que se deben tomar más bits en otros lugares y tal vez puede reducir la velocidad de bits disponible para la estructura fina.

- 5 Para codificar la forma antes de codificar la ganancia es una solución mejor, pero si la velocidad de bits para el cuantificador de forma se decide a partir del valor de ganancia cuantificado, entonces la ganancia y la cuantificación de forma dependerían entre sí. Una solución iterativa podría resolver esta codependencia, pero podría resultar demasiado compleja para ejecutarse en tiempo real en un dispositivo móvil.
- 10 El documento US 2011/0002266 A1 (Yang Gao) describe un post-procesamiento de dominio frecuencia basándose en enmascaramiento perceptual, donde un factor de ganancia de modificación adaptativa se aplica a cada coeficiente de frecuencia con el fin de mejorar la calidad percibida de los coeficientes espectrales.

Sumario

- 15 Un objeto es obtener un ajuste de ganancia en la decodificación de audio que se ha codificado con representaciones separadas de ganancia y forma.

Este objeto se consigue de acuerdo con las reivindicaciones adjuntas.

- 20 Un primer aspecto implica un método de ajuste de ganancia que incluye los siguientes pasos:
- Se estima una medida de precisión de la representación de la forma.
 - 25 • Se determina una corrección de ganancia basándose en la medida de precisión estimada.
 - La representación de ganancia se ajusta basándose en la corrección de ganancia determinada.

Un segundo aspecto implica un aparato de ajuste de ganancia que incluye:

- 30 • Un medidor de precisión configurado para estimar una medida de precisión de la representación de forma y para determinar una corrección de ganancia basándose en la medida de precisión estimada.
- 35 • Un ajustador de envolvente configurado para ajustar la representación de ganancia basándose en la corrección de ganancia determinada.

Un tercer aspecto implica un decodificador que incluye un aparato de ajuste de ganancia de acuerdo con el segundo aspecto.

- 40 Un cuarto aspecto implica un nodo de red que incluye un decodificador de acuerdo con el tercer aspecto.

El esquema propuesto para la corrección de ganancia mejora la calidad percibida de un sistema de codificación de audio en forma de ganancia. El esquema tiene una complejidad computacional baja y requiere pocos bits adicionales, si los hay.

Breve descripción de los dibujos

La presente tecnología, junto con otros objetivos y ventajas de la misma, puede entenderse mejor haciendo referencia a la siguiente descripción tomada junto con los dibujos adjuntos, en los que:

- 50 la figura 1 ilustra un esquema de cuantificación de vector en forma de ganancia de ejemplo;
- la figura 2 ilustra un esquema de codificación y decodificación de dominio de transformada de ejemplo;
- 55 la figura 3A-C ilustra la cuantificación del vector en forma de ganancia en un caso simplificado;
- la figura 4 ilustra un decodificador de dominio de transformada de ejemplo que utiliza una medida de precisión para determinar una corrección de envolvente;
- 60 la figura 5A-B ilustra un resultado de ejemplo de escalar la síntesis con factores de ganancia cuando el vector de forma es un vector de impulso escaso;
- la figura 6A-B ilustra cómo la altura de impulso más grande puede indicar la precisión del vector de forma;
- 65 la figura 7 ilustra un ejemplo de una función de atenuación basada en la velocidad para la realización 1;

la figura 8 ilustra un ejemplo de una función de ajuste de ganancia dependiente de la altura máxima del impulso y de la velocidad para la realización 1;

la figura 9 ilustra otro ejemplo de una función de ajuste de ganancia dependiente de la altura de impulso máxima y de la velocidad para la realización 1;

la figura 10 ilustra una realización de la presente tecnología en el contexto de un sistema codificador y decodificador de audio basado en MDCT;

la figura 11 ilustra un ejemplo de una función de mapeo desde la medida de estabilidad al factor de limitación de ajuste de ganancia;

la figura 12 ilustra un ejemplo de un sistema decodificador y codificador ADPCM y con un tamaño de paso adaptativo;

la figura 13 ilustra un ejemplo en el contexto de un sistema codificador y decodificador de audio basado en ADPCM de sub-banda;

la figura 14 ilustra un ejemplo de la presente tecnología en el contexto de un sistema codificador y decodificador de audio basado en ADPCM de sub-banda;

la figura 15 ilustra un codificador de dominio de transformada de ejemplo que incluye un clasificador de señales;

la figura 16 ilustra otro decodificador de dominio de transformada de ejemplo que utiliza una medida de precisión para determinar una corrección de envolvente;

la figura 17 ilustra una realización de un aparato de ajuste de ganancia de acuerdo con la presente tecnología;

la figura 18 ilustra una realización de ajuste de ganancia de acuerdo con la presente tecnología con más detalle;

la figura 19 es un diagrama de flujo que ilustra el método de acuerdo con la presente tecnología;

la figura 20 es un diagrama de flujo que ilustra una realización del método de acuerdo con la presente tecnología; y

la figura 21 ilustra una realización de una red de acuerdo con la tecnología actual.

Descripción detallada

En la siguiente descripción se utilizarán las mismas designaciones de referencia para elementos que realicen la misma función o una función similar.

Antes de describir la presente tecnología en detalle, la codificación en forma de ganancia se ilustrará con referencia a las figuras 1-3.

La figura 1 ilustra un esquema de cuantificación de vector en forma de ganancia de ejemplo. La parte superior de la figura ilustra el lado del codificador. Un vector x de entrada se envía a una calculadora 10 de normalización, que determina la norma del vector (ganancia) g , típicamente la norma euclidiana. Esta norma exacta se cuantifica en un cuantificador 12 de norma, y la inversa $1/\hat{g}$ de la norma cuantificada $\hat{g}$ se envía a un multiplicador 14 para escalar el vector x de entrada en una forma. La forma se cuantifica en un cuantificador 16 de forma. Las representaciones de la ganancia y la forma cuantificadas se envían a un multiplexor 18 de flujo de bits (mux). Estas representaciones se ilustran con líneas discontinuas para indicar que pueden, por ejemplo, constituir índices en tablas (libros de códigos) en lugar de los valores cuantificados reales.

La parte inferior de la figura 1 ilustra el lado del decodificador. Un demultiplexor 20 de flujo de bits (demux) recibe las representaciones de ganancia y forma. La representación de forma se envía a un descuantificador 22 de forma, y la representación de ganancia se envía a un descuantificador 24 de ganancia. La ganancia obtenida $\hat{g}$ se envía a un multiplicador 26, donde escala la forma obtenida, lo que da el vector reconstruido $\hat{x}$.

La figura 2 ilustra un esquema de codificación y decodificación de dominio de transformada de ejemplo. La parte superior de la figura ilustra el lado del codificador. Una señal de entrada es enviada a un transformador de frecuencia 30, por ejemplo basándose en la Transformada Discreta de Coseno Modificada (MDCT), para producir la transformada X de frecuencia. La transformada X de frecuencia es enviada a una calculadora 32 de envolvente, la cual determina la energía $E(b)$ de cada banda de frecuencia b . Estas energías se cuantifican en energías $\hat{E}(b)$ en un cuantificador 34 de envolvente. Las energías cuantificadas $\hat{E}(b)$ son enviadas a un normalizador 36 de envolvente, que escala los coeficientes de la banda de frecuencias b de la transformada X con la inversa de la energía

cuantificada $\hat{E}(b)$ correspondiente de la envolvente. Las formas escaladas resultantes son enviadas a un cuantificador 38 de estructura fina. Las energías cuantificadas $\hat{E}(b)$ también son enviadas a un asignador 40 de bits, que asigna bits para la cuantificación de estructura fina a cada banda b de frecuencia. Como se ha indicado anteriormente, la asignación $R(b)$ de bits puede basarse en un modelo del sistema auditivo humano. Las representaciones de las ganancias cuantificadas $\hat{E}(b)$ y las correspondientes formas cuantificadas son enviadas al multiplexor 18 de flujo de bits.

La parte inferior de la figura 2 ilustra el lado del decodificador. El demultiplexor 20 de flujo de bits recibe las representaciones de ganancia y forma. Las representaciones de ganancia son enviadas a un descuantificador 42 de envolventes. Las energías de envolvente generadas $\hat{E}(b)$ son enviadas a un asignador 44 de bits, que determina la asignación $R(b)$ de bits de las formas recibidas. Las representaciones de forma son enviadas a un descuantificador 46 de estructura fina, que está controlado por la asignación $R(b)$ de bits. Las formas decodificadas son enviadas a un conformador 48 de envolvente, que las escala con las energías de envolvente correspondientes $\hat{E}(b)$ para formar una transformada de frecuencia reconstruida. Esta transformada es enviada a un transformador 50 de frecuencia inverso, por ejemplo basado en la Transformada de Coseno Discreta Modificada Inversa (IMDCT), que produce una señal de salida que representa audio sintetizado.

La figura 3A-C ilustra la cuantificación de vector en forma de ganancia descrita anteriormente en un caso simplificado en el que la banda b de frecuencias está representada por el vector bidimensional $X(b)$ en la figura 3A. Este caso es suficientemente simple para ser ilustrado en un dibujo, pero también lo suficientemente general como para ilustrar el problema con la cuantificación en forma de ganancia (en la práctica, los vectores tienen típicamente 8 o más dimensiones). El lado derecho de la figura 3A ilustra una representación exacta en forma de ganancia del vector $X(b)$ con una ganancia $E(b)$ y una forma (vector de longitud unitaria) $N'(b)$.

Sin embargo, como se ilustra en la figura 3B, la ganancia exacta $E(b)$ se codifica en una ganancia cuantificada $\hat{E}(b)$ en el lado del codificador. Dado que la inversa de la ganancia cuantificada $\hat{E}(b)$ se utiliza para escalar el vector $X(b)$, el vector escalado resultante $N(b)$ apuntará en la dirección correcta, pero no será necesariamente de longitud unitaria. Durante la cuantificación de forma el vector escalado $N(b)$ se cuantifica en la forma cuantificada $\hat{N}(b)$. En este caso, la cuantificación se basa en un esquema de codificación de impulsos [3], que construye la forma (o dirección) a partir de una suma de impulsos enteros con signo. Los impulsos pueden añadirse uno encima del otro para cada dimensión. Esto significa que las posiciones de cuantificación de forma permitidas están representadas por los puntos grandes en las rejillas rectangulares ilustradas en las figuras 3B-C. El resultado es que la forma cuantificada $\hat{N}(b)$ en general no coincide con la forma (dirección) de $N(b)$ (y $N'(b)$).

La figura 3C ilustra que la precisión de la cuantificación de forma depende de los bits asignados $R(b)$, o de manera equivalente el número total de impulsos disponibles para la cuantificación de forma. En la parte izquierda de la figura 3C la cuantificación de forma se basa en 8 impulsos, mientras que la cuantificación de forma en la parte derecha utiliza sólo 3 impulsos (el ejemplo de la figura 3B utiliza 4 impulsos).

Por lo tanto, se apreciará que, dependiendo de la precisión del cuantificador de forma, el valor de ganancia $\hat{E}(b)$ utilizado para reconstruir el vector (b) en el lado del decodificador puede ser más o menos apropiado. De acuerdo con la presente tecnología, una corrección de ganancia puede basarse en una medida de precisión de la forma cuantificada.

La medida de precisión utilizada para corregir la ganancia puede derivarse de parámetros ya disponibles en el decodificador, pero también puede depender de parámetros adicionales designados para la medida de precisión. Típicamente, los parámetros incluirían el número de bits asignados para el vector de forma y el propio vector de forma, pero también puede incluir el valor de ganancia asociado con el vector de forma y estadísticas prealmacenadas sobre las señales que son típicas para el sistema de codificación y decodificación. En la figura 4 se muestra un resumen de un sistema que incorpora una medida de precisión y una corrección o ajuste de ganancia.

La figura 4 ilustra un ejemplo de decodificador 300 de dominio de transformada que utiliza una medida de precisión para determinar una corrección de envolvente. Con el fin de evitar el desorden del dibujo, sólo se ilustra el lado del decodificador. El lado del codificador puede ser implementado como en la figura 2. La nueva característica es un aparato 60 de ajuste de ganancia. El aparato 60 de ajuste de ganancia incluye un medidor 62 de precisión configurado para estimar una medida $A(b)$ de precisión de la representación $\hat{N}(b)$ de forma, y para determinar una corrección $g_c(b)$ de ganancia basándose en la medida $A(b)$ de precisión estimada. También incluye un ajustador 64 de envolvente configurado para ajustar la representación $\hat{E}(b)$ de ganancia basándose en la corrección de ganancia determinada.

Como se ha indicado anteriormente, la corrección de ganancia puede, en algunas realizaciones, ser formada sin gastar bits adicionales. Esto se realiza estimando la corrección de ganancia a partir de parámetros ya disponibles en el decodificador. Este proceso se puede describir como una estimación de la precisión de la forma codificada. Típicamente, esta estimación incluye derivar la medida $A(b)$ de precisión de las características de cuantificación de forma que indican la resolución de la cuantificación de forma.

Realización 1

En una realización, la presente tecnología se utiliza en un sistema codificador/decodificador de audio. El sistema se basa en la transformada y la transformada utilizada es la Transformada Discreta de Coseno Modificada (MDCT) que utiliza ventanas sinusoidales con superposición del 50%. Sin embargo, se entiende que cualquier transformada adecuada para la codificación de transformada puede utilizarse junto con una segmentación y ventana apropiadas.

Codificador de la realización 1

El audio de entrada se extrae en tramas utilizando 50% de superposición y ventanas con una ventana sinusoidal simétrica. Cada trama con ventanas se transforma a continuación en un espectro MDCT X . El espectro se divide en sub-bandas para procesamiento, donde las anchuras de la sub-banda no son uniformes. Los coeficientes espectrales de la trama m que pertenecen a la banda b se indican $X(b, m)$ y tienen el ancho de banda $BW(b)$. Puesto que la mayoría de los pasos del codificador y del decodificador se pueden describir dentro de una trama, omitimos el índice de trama y apenas utilizamos la notación $X(b)$. Los anchos de banda deberían aumentar preferentemente con frecuencia creciente para cumplir con la resolución de frecuencia del sistema auditivo humano. El valor cuadrático medio (RMS) de cada banda se utiliza como factor de normalización y se denomina $E(b)$:

$$E(b) = \sqrt{\frac{X(b)^T X(b)}{BW(b)}} \quad (1)$$

donde $X(b)^T$ indica la transposición de $X(b)$.

El valor RMS puede ser visto como el valor de energía por coeficiente. La secuencia de factores de normalización $E(b)$ para $b = 1, 2, \dots, N_{\text{bandas}}$ forma la envolvente del espectro MDCT, donde N_{bandas} indica el número de bandas. A continuación, la secuencia se cuantifica para ser transmitida al decodificador. Para asegurar que la normalización puede invertirse en el decodificador, se obtiene la envolvente cuantificada $\hat{E}(b)$. En esta realización de ejemplo, los coeficientes de envolvente son cuantificados escalares en el dominio de registro utilizando un tamaño de paso de 3 dB y los índices de cuantificador se codifican diferencialmente utilizando la codificación de Huffman. La envolvente cuantificada se utiliza para la normalización de las bandas espectrales, es decir:

$$N(b) = \frac{1}{\hat{E}(b)} X(b) \quad (2)$$

Téngase en cuenta que si la envolvente no cuantificada $E(b)$ se utiliza para la normalización, la forma tendría RMS = 1, es decir:

$$N'(b) = \frac{1}{E(b)} X(b) \Rightarrow \sqrt{\frac{N'(b)^T N'(b)}{BW(b)}} = 1 \quad (3)$$

Mediante el uso de la envolvente cuantificada $\hat{E}(b)$, el vector de forma tendrá un valor RMS cercano a 1. Esta característica se utilizará en el decodificador para crear una aproximación del valor de ganancia.

La unión de los vectores de forma normalizados $N(b)$ forma la estructura fina del espectro MDCT. La envolvente cuantificada se utiliza para producir una asignación $R(b)$ de bits para la codificación de los vectores $N(b)$ de forma normalizados. El algoritmo de asignación de bits utiliza preferentemente un modelo auditivo para distribuir los bits a las partes perceptualmente más relevantes. Puede utilizarse cualquier esquema de cuantificación para codificar el vector de forma. Común para todos es que pueden ser diseñados bajo el supuesto de que la entrada es normalizada, lo que simplifica el diseño de cuantificador. En esta realización, la cuantificación de forma se realiza utilizando un esquema de codificación de impulsos que construye la forma de síntesis a partir de una suma de impulsos de enteros señalizados [3]. Los impulsos pueden añadirse uno encima del otro para formar impulsos de diferente altura. En esta realización, la asignación $R(b)$ de bits indica el número de impulsos asignados a la banda b .

Los índices de cuantificador a partir de la cuantificación de envolvente y la cuantificación de forma son multiplexados en un flujo de bits para ser almacenados o transmitidos a un decodificador.

Decodificador de la realización 1

El decodificador desmultiplexa los índices del flujo de bits y envía los índices relevantes a cada módulo de decodificación. En primer lugar, se obtiene la envolvente cuantificada $\hat{E}(b)$. A continuación, la asignación de bits de

estructura fina se deriva de la envolvente cuantificada utilizando una asignación de bits idéntica a la utilizada en el codificador. Los vectores $\hat{N}(b)$ de forma de la estructura fina se decodifican utilizando los índices y la asignación $R(b)$ de bits obtenida.

- 5 Ahora, antes de escalar la estructura fina decodificada con la envolvente, se determinan factores de corrección de ganancia adicionales. En primer lugar, la ganancia de correspondencia RMS se obtiene como:

$$g_{RMS}(b) = \sqrt{\frac{BW(b)}{\hat{N}(b)^T \hat{N}(b)}} \quad (4)$$

- 10 El factor $g_{RMS}(b)$ es un factor de escala que normaliza el valor RMS a 1, es decir:

$$\sqrt{\frac{(g_{RMS}(b)\hat{N}(b))^T (g_{RMS}(b)\hat{N}(b))}{BW(b)}} = 1 \quad (5)$$

En esta realización buscamos minimizar el error cuadrático medio (MSE) de la síntesis:

15

$$g_{MSE}(b) = \arg \min_g |N(b) - g \cdot \hat{N}(b)| \quad (6)$$

con la solución

$$g_{MSE}(b) = \frac{\hat{N}(b)^T N(b)}{N(b)^T N(b)} \quad (7)$$

20

Dado que $g_{MSE}(b)$ depende de la forma de entrada $N(b)$, no se conoce en el decodificador. En esta realización, el impacto se estima utilizando una medida de precisión. La relación de estas ganancias se define como un factor de corrección de ganancia $g_c(b)$:

25

$$g_c(b) = \frac{g_{MSE}(b)}{g_{RMS}(b)} \quad (8)$$

Cuando la precisión de la cuantificación de la forma es buena, el factor de corrección es cercano a 1, es decir:

30

$$\hat{N}(b) \rightarrow N(b) \Rightarrow g_c(b) \rightarrow 1 \quad (9)$$

Sin embargo, cuando la precisión de $\hat{N}(b)$ es baja, $g_{MSE}(b)$ y $g_{RMS}(b)$ divergirán. En esta realización, en la que la forma se codifica utilizando un esquema de codificación de impulsos, una velocidad baja hará que el vector de forma sea escaso y $g_{RMS}(b)$ dará una sobreestimación de la ganancia apropiada en términos de MSE. Para este caso $g_c(b)$ debe ser inferior a 1 para compensar el rebasamiento. Véanse las figuras 5A-B para una ilustración de ejemplo del caso de forma de impulso de baja velocidad. La figura 5A-B ilustra un ejemplo de escala de la síntesis con factores de ganancia g_{MSE} (figura 5B) y g_{RMS} (figura 5A) cuando el vector de forma es un vector de impulso escaso. La escala de g_{RMS} da impulsos que son demasiado altos en un sentido MSE.

35

40 Por otra parte, una señal de destino pico o escasa puede estar bien representada con una forma de impulso. Aunque la dispersión de la señal de entrada puede no ser conocida en la etapa de síntesis, la escasez de la forma de síntesis puede servir como un indicador de la precisión del vector de forma sintetizado. Una forma de medir la dispersión de la forma de síntesis es la altura del pico máximo en la forma. El razonamiento detrás de esto es que una señal de entrada escasa es más probable que genere altos picos en la forma de síntesis. Véase la figura 7A-B para una ilustración de cómo la altura del pico puede indicar la precisión de dos vectores de impulso de velocidad igual. En la figura 7A hay 5 impulsos disponibles ($R(b) = 5$) para representar la forma discontinua. Puesto que la forma es bastante constante, la codificación generó 5 impulsos distribuidos de igual altura 1, es decir $p_{max} = 1$. En la figura 7B también hay 5 impulsos disponibles para representar la forma discontinua. Sin embargo, en este caso la forma es pico o escasa, y el pico más grande está representado por 3 impulsos uno encima del otro, es decir $p_{max} = 3$. Esto indica que la corrección $g_c(b)$ de ganancia depende de una dispersión estimada p_{max} de la forma cuantificada.

45

50

Como se ha indicado anteriormente, la forma $N(b)$ de entrada no es conocida por el decodificador. Puesto que $g_{MSE}(b)$ depende de la forma $N(b)$ de entrada, esto significa que la corrección o compensación $g_c(b)$ de ganancia no puede en la práctica basarse en la ecuación ideal (8). En esta realización se decide en cambio la corrección de ganancia $g_c(b)$ basándose en la velocidad binaria en términos del número de impulsos $R(b)$, la altura del impulso más grande en el vector de forma $p_{max}(b)$ y la banda b de frecuencias, es decir:

$$g_c(b) = f(R(b), p_{max}(b), b) \quad (10)$$

Se ha observado que las velocidades más bajas requieren generalmente una atenuación de la ganancia para minimizar el MSE. La dependencia de la velocidad puede ser implementada como una tabla $t(R(b))$ de consulta que es entrenada en datos de señal de audio relevantes. En la figura 7 se puede ver una tabla de consulta de ejemplo. Dado que los vectores de forma en esta realización tienen anchuras diferentes, la velocidad puede expresarse preferentemente como número de impulsos por muestra. De esta manera se puede utilizar la misma atenuación dependiente de la velocidad para todos los anchos de banda. Una solución alternativa, que se utiliza en esta realización, es utilizar un tamaño T de paso en la tabla dependiendo de la anchura de la banda. Aquí, utilizamos 4 anchos de banda diferentes en 4 grupos diferentes y por lo tanto requieren 4 tamaños de paso. Un ejemplo de tamaños de paso se encuentra en la tabla 1. Utilizando el tamaño de paso, el valor de búsqueda se obtiene utilizando una operación de redondeo $t(\lfloor R(b) \cdot T \rfloor)$, donde $\lfloor \cdot \rfloor$ representa el redondeo al entero más cercano.

Tabla 1

Grupo de banda	Ancho de banda	Tamaño T de paso
1	8	4
2	16	4/3
3	24	2
4	34	1

Otro cuadro de consulta de ejemplo se da en la tabla 2.

Tabla 2

Grupo de banda	Ancho de banda	Tamaño T de paso
1	8	4
2	16	4/3
3	24	2
4	32	1

La dispersión estimada se puede implementar como otra tabla $u(R(b), p_{max}(b))$ de consulta basándose tanto en el número de impulsos $R(b)$ como en la altura del impulso máximo $p_{max}(b)$. En la figura 8 se muestra una tabla de consulta de ejemplo. La tabla u de consulta sirve como medida $A(b)$ de precisión para la banda b , es decir:

$$A(b) = u(R(b), p_{max}(b)) \quad (11)$$

Se observó que la aproximación de g_{MSE} era más adecuada para el rango de frecuencias más bajas desde una perspectiva perceptual. Para las frecuencias más altas, la estructura fina se vuelve menos importante perceptualmente y la coincidencia del valor de energía o RMS se hace vital. Por esta razón, la atenuación de ganancia se puede aplicar sólo por debajo de cierto número de banda b_{THR} . En este caso, la corrección de ganancia $g_c(b)$ tendrá una dependencia explícita de la banda b de frecuencias. La función de corrección de ganancia resultante puede definirse en este caso como:

$$g_c(b) = \begin{cases} t(R(b)) \cdot A(b), & b < b_{THR} \\ 1, & \text{de otro modo} \end{cases} \quad (12)$$

La descripción hasta este punto también se puede utilizar para describir las características esenciales del ejemplo de realización de la figura 4. Así, en la realización de la figura 4, la síntesis final $X(b)$ se calcula como:

$$\hat{X}(b) = \underbrace{g_c(b) g_{RMS}(b) \hat{E}(n)}_{\hat{E}(n)} \hat{N}(b) \quad (13)$$

Como alternativa, la función $u(R(b), p_{max}(b))$ se puede implementar como una función lineal de la altura máxima de impulso p_{max} y la velocidad $R(b)$ de bits asignada, por ejemplo como:

$$u(R(b), p_{\max}(b)) = k \cdot (p_{\max}(b) - R(b)) + 1 \quad (14)$$

donde la inclinación k se determina por:

$$\begin{aligned} k &= \frac{1 - (a_{\min} + R(b) \cdot \Delta a)}{R(b) - 1} \\ \Delta a &= (a_{\max} - a_{\min}) / R(b) \\ a_{\max} &= 1 - \frac{1 - a_{\min}}{R(b) - 1} \end{aligned} \quad (15)$$

La función depende del parámetro $a_{\min}$ de sintonía que da el factor de atenuación inicial para $R(b) = 1$ y $p_{\max}(b) = 1$. La función se ilustra en la figura 9, con el parámetro $a_{\min} = 0,41$ de ajuste. Típicamente $u_{\max} \in [0,7, 1,4]$ y $u_{\min} \in [0, u_{\max}]$. En la ecuación (14) u es lineal en la diferencia entre $p_{\max}(b)$ y $R(b)$. Otra posibilidad es tener diferentes factores de inclinación para $p_{\max}(b)$ y $R(b)$.

La velocidad de bits para una banda dada puede cambiar drásticamente para una banda dada entre tramas adyacentes. Esto puede conducir a variaciones rápidas de la corrección de ganancia. Tales variaciones son especialmente críticas cuando la envolvente es bastante estable, es decir, los cambios totales entre las tramas son bastante pequeños. Esto sucede a menudo para señales de música que normalmente tienen envolventes de energía más estables. Para evitar que la atenuación de ganancia introduzca inestabilidad, se puede añadir una adaptación adicional. En la figura 10 se da una visión general de tal realización, en la que se ha añadido un medidor 66 de estabilidad al aparato 60 de ajuste de ganancia en el decodificador 300.

La adaptación puede por ejemplo basarse en una medida de estabilidad de la envolvente $E(b)$. Un ejemplo de una medida de este tipo es calcular la distancia euclidiana al cuadrado entre vectores de envolvente $\log_2$ adyacentes:

$$\Delta E(m) = \frac{1}{N_{\text{bands}}} \sum_{b=0}^{N_{\text{bands}}-1} (\log_2 \hat{E}(b, m) - \log_2 \hat{E}(b, m-1))^2 \quad (16)$$

Aquí, $\Delta E(m)$ indica la distancia euclidiana al cuadrado entre los vectores de envolvente para la trama m y la trama $m - 1$. La medida de estabilidad también puede ser filtrada en paso bajo para tener una adaptación más suave:

$$\Delta \tilde{E}(m) = \alpha \Delta E(m) + (1 - \alpha) \Delta \tilde{E}(m-1) \quad (17)$$

Un valor adecuado para el factor α de olvido puede ser 0,1. La medida de estabilidad se puede utilizar para crear una limitación de la atenuación, por ejemplo, una función sigmoide, tal como:

$$g_{\min} = \frac{1}{1 + e^{C_1(\Delta \tilde{E}(m) - C_2) - C_3}}, \quad (18)$$

donde los parámetros se pueden establecer en $C_1 = 6$, $C_2 = 2$ y $C_3 = 1,9$. Debe tenerse en cuenta que estos parámetros han de considerarse como ejemplos, mientras que los valores reales pueden elegirse con más libertad. Por ejemplo:

$$C_1 \in [1, 10]$$

$$C_2 \in [1, 4]$$

$$C_3 \in [-5, 10]$$

La figura 11 ilustra un ejemplo de una función de mapeo desde la medida de estabilidad $\Delta \tilde{E}(m)$ al factor $g_{\min}$ de limitación de ajuste de ganancia. La expresión anterior para $g_{\min}$ se implementa preferentemente como una tabla de búsqueda o con una función de paso simple, tal como:

$$g_{\min} = \begin{cases} 1, & \Delta\tilde{E}(m) < C_3 / C_1 + C_2 \\ 0, & \Delta\tilde{E}(m) \geq C_3 / C_1 + C_2 \end{cases} \quad (19)$$

La variable $g_{\min} \in [0,1]$ de limitación de atenuación se puede utilizar para crear una modificación $\tilde{g}_c(b)$ de ganancia adaptada a la estabilidad como:

$$\tilde{g}_c(b) = \max(g_c(b), g_{\min}) \quad (20)$$

Después de la estimación de la ganancia, la síntesis final $X(b)$ se calcula como:

$$\hat{X}(b) = \underbrace{\tilde{g}_c(b)g_{RMS}(b)\hat{E}(n)}_{\tilde{E}(n)}\hat{N}(b) \quad (21)$$

En las variantes descritas de la realización 1, la unión de los vectores sintetizados $X(b)$ forma el espectro sintetizado $\hat{X}$, que se procesa adicionalmente utilizando la transformada MDCT inversa, se enciende con la ventana sinusoidal simétrica y se añade a la síntesis de salida utilizando la estrategia de superposición y suma.

Ejemplo 2

En otro ejemplo, la forma se cuantifica utilizando un banco de filtros QMF (filtro espejo en cuadratura) y un esquema ADPCM (modulación por código de impulso diferencial adaptativo) para la cuantificación de la forma. Un ejemplo de un esquema ADPCM de sub-banda es el ITU-T G.722 [4]. La señal de audio de entrada se procesa preferentemente en segmentos. Un ejemplo de esquema ADPCM se muestra en la figura 12, con un tamaño S de paso adaptativo. Aquí, el tamaño de paso adaptativo del cuantificador de forma sirve como una medida de precisión que ya está presente en el decodificador y no requiere señalización adicional. Sin embargo, el tamaño del paso de cuantificación debe extraerse de los parámetros utilizados por el proceso de decodificación y no de la propia forma sintetizada. Una vista general de esta realización se muestra en la figura 14. Sin embargo, antes de que esta realización se describa en detalle, se describirá un esquema ADPCM de ejemplo basándose en un banco de filtros QMF con referencia a las figuras 12 y 13.

La figura 12 ilustra un ejemplo de un sistema de codificador y decodificador ADPCM con un tamaño de paso de adaptativo. Un cuantificador ADPCM 70 incluye un sumador 72, que recibe una señal de entrada y resta una estimación de la señal de entrada anterior para formar una señal e de error. La señal de error se cuantifica en un cuantificador 74, cuya salida es enviada al multiplexor 18 de flujo de bits, y también a una calculadora de tamaño de paso 76 y un descuantificador 78. La calculadora 76 de tamaño de paso adapta el tamaño S de paso de cuantificación para obtener un error aceptable. El tamaño S de paso de cuantificación es enviado al multiplexor 18 de flujo de bits, y también controla el cuantificador 74 y el descuantificador 78. El descuantificador 78 emite una estimación $\hat{e}$ de error a un sumador 80. La otra entrada del sumador 80 recibe una estimación de la señal de entrada que ha sido retardada por un elemento 82 de retardo. Esto forma una estimación actual de la señal de entrada, que es enviada al elemento 82 de retardo. La señal retardada también es enviada a la calculadora 76 de tamaño de paso y (con un cambio de signo) al sumador 72 para formar la señal e de error.

Un descuantificador ADPCM 90 incluye un decodificador 92 de tamaño de paso, que decodifica el tamaño S de paso recibido y lo envía a un descuantificador 94. El descuantificador 94 decodifica la estimación $\hat{e}$ de error, que es enviada a un sumador 98, cuya otra entrada recibe la señal de salida del sumador retardada por un elemento 96 de retardo.

La figura 13 ilustra un ejemplo en el contexto de un sistema codificador y decodificador de audio basado en ADPCM de sub-banda. El lado del codificador es similar al lado del codificador de la realización de la figura 2. Las diferencias esenciales son que el transformador de frecuencia 30 ha sido reemplazado por un banco 100 de filtros de análisis QMF (filtro espejo en cuadratura), y que el cuantificador 38 de estructura fina ha sido reemplazado por un cuantificador ADPCM, tal como el cuantificador 70 en la figura 12. El lado del decodificador es similar al lado del decodificador de la realización de la figura 2. Las diferencias esenciales son que el transformador 50 de frecuencia inversa ha sido reemplazado por un banco 102 de filtro de síntesis QMF, y que el descuantificador 46 de estructura fina ha sido reemplazado por un descuantificador ADPCM tal como el descuantificador 90 en la figura 12.

La figura 14 ilustra una realización de la presente tecnología en el contexto de un sistema codificador y decodificador de audio basado en ADPCM de sub-banda. Con el fin de evitar el desorden del dibujo, sólo se ilustra el lado del decodificador 300. El lado del codificador puede ser implementado como en la figura 13.

Codificador del ejemplo 2

El codificador aplica el banco de filtros QMF para obtener las señales de sub-banda. Los valores de RMS de cada señal de sub-banda se calculan y las señales de sub-bandas se normalizan. La envolvente $E(b)$, la asignación de bits de sub-banda $R(b)$ y los vectores $N(b)$ de forma normalizados se obtienen como en la realización 1. Cada sub-banda normalizada se alimenta en el cuantificador ADPCM. En esta realización, el ADPCM funciona de una manera adaptativa hacia delante y determina un paso $S(b)$ de escalado que se utilizará para la sub-banda b . El paso de escalado se elige para minimizar el MSE a través de la trama de sub-banda. En esta realización, el paso se elige intentando todos los pasos posibles y seleccionando el que da el mínimo MSE:

$$S(b) = \min_s \frac{1}{BW(b)} (N(b) - Q(N(b), s))^T (N(b) - Q(N(b), s)) \quad (22)$$

donde $Q(x, s)$ es la función de ADPCM de la variable x que utiliza un tamaño de paso de s . El tamaño de paso seleccionado puede utilizarse para generar la forma cuantificada:

$$\hat{N}(b) = Q(N(b), S(b)) \quad (23)$$

Los índices de cuantificador a partir de la cuantificación de envolvente y la cuantificación de forma son multiplexados en un flujo de bits para ser almacenados o transmitidos a un decodificador.

Decodificador del ejemplo 2

El decodificador desmultiplexa los índices del flujo de bits y envía los índices relevantes a cada módulo de decodificación. La envolvente cuantificada $E(b)$ y la asignación $R(b)$ de bits se obtienen como en la realización 1. Los vectores $N(b)$ de forma sintetizados se obtienen a partir del decodificador ADPCM o descuantificador junto con los tamaños $S(b)$ de pasos de adaptación. Los tamaños de paso indican una precisión del vector de forma cuantificado, donde un tamaño de paso más pequeño corresponde a una mayor precisión y viceversa. Una posible implementación es hacer que la precisión $A(b)$ sea inversamente proporcional al tamaño del paso utilizando un factor de proporcionalidad γ :

$$A(b) = \gamma \frac{1}{S(b)} \quad (24)$$

donde γ debe establecerse para lograr la relación deseada. Una elección posible es $\gamma = S_{\min}$ donde $s_{\min}$ es el tamaño de paso mínimo, que da la precisión 1 para $S(b) = S_{\min}$.

El factor de corrección de ganancia g_c puede obtenerse utilizando una función de mapeo:

$$g_c(b) = h(R(b), b) \cdot A(b) \quad (1)$$

La función h de mapeo puede implementarse como una tabla de consulta basada en la velocidad $R(b)$ y la banda b de frecuencia. Esta tabla puede definirse agrupando los valores g_{MSE}/g_{RMS} de corrección de ganancia óptimos por estos parámetros y calculando la entrada de tabla promediando los valores de corrección de ganancia óptimos para cada grupo.

Después de la estimación de la corrección de ganancia, la síntesis $X(b)$ de sub-banda se calcula como:

$$\hat{X}(b) = \underbrace{g_c(b) g_{RMS}(b) \hat{E}(n)}_{\hat{E}(n)} \hat{N}(b) \quad (2)$$

La trama de audio de salida se obtiene aplicando el banco de filtros de síntesis QMF a las sub-bandas.

En el ejemplo ilustrado en la figura 14, el medidor 62 de precisión en el aparato 60 de ajuste de ganancia recibe el tamaño $S(b)$ de paso de cuantificación aún no decodificado directamente del flujo de bits recibido. Una alternativa, como se ha indicado anteriormente, es decodificarlo en el descuantificador 90 de ADPCM y enviarlo en forma decodificada al medidor 62 de precisión.

Hay que señalar que el ejemplo 2 descrito anteriormente no forma parte de la invención reivindicada pero da un ejemplo útil de una implementación alternativa para entender la invención.

Otras alternativas

La medida de precisión se podría complementar con un parámetro de clase de señal derivado en el codificador. Esto puede ser, por ejemplo, un discriminador de voz/música o un estimador de nivel de ruido de fondo. En la figura 15-16 se muestra un resumen de un sistema que incorpora un clasificador de señales. El lado del codificador en la figura 15 es similar al lado del codificador en la figura 2, pero ha sido provisto de un clasificador 104 de señales. El lado 300 del decodificador en la figura 16 es similar al lado del decodificador en la figura 4, pero ha sido provisto de una entrada de clase de señal adicional en el medidor 62 de precisión.

La clase de señal podría incorporarse en la corrección de ganancia, por ejemplo, teniendo una adaptación dependiente de la clase. Si asumimos que las clases de señal son la voz o la música correspondientes a los valores $C = 1$ y $C = 0$, respectivamente, podemos limitar el ajuste de ganancia a ser efectivo sólo durante la voz, es decir:

$$g_c(b) = \begin{cases} t(R(b)) \cdot A(b), & b < b_{THR} \wedge C = 1 \\ 1, & \text{de otro modo} \end{cases} \quad (27)$$

En otra realización alternativa el sistema puede actuar como un predictor junto con una corrección o compensación de ganancia parcialmente codificada. En esta realización, la medida de precisión se utiliza para mejorar la predicción de la corrección o compensación de ganancia de tal manera que el error de ganancia restante pueda codificarse con menos bits.

Cuando se crea la corrección de ganancia o el factor g_c de compensación, es posible que se quiera realizar una compensación entre el valor RMS o la energía y minimizar el MSE. En algunos casos, hacer coincidir la energía es más importante que una forma de onda precisa. Esto es por ejemplo verdadero para las frecuencias más altas. Para acomodar esto, la corrección de ganancia final puede, en una realización adicional, ser formada utilizando una suma ponderada de los diferentes valores de ganancia:

$$g'_c = \frac{\beta g_{RMS} + (1 - \beta) g_{MSE}}{g_{RMS}} = \beta + (1 - \beta) \frac{g_{MSE}}{g_{RMS}} = \beta + (1 - \beta) g_c \quad (28)$$

donde g_c es la corrección de ganancia obtenida de acuerdo con uno de los enfoques descritos anteriormente. El factor β de ponderación puede adaptarse por ejemplo a la frecuencia, velocidad de bits o el tipo de señal.

Los pasos, funciones, procedimientos y / o bloques descritos en el presente documento pueden implementarse en el equipo físico utilizando cualquier tecnología convencional, tal como tecnología de circuito discreto o de circuito integrado, incluyendo tanto circuitería electrónica de uso general como circuitería específica de aplicación.

Alternativamente, al menos algunos de los pasos, funciones, procedimientos y/o bloques descritos en el presente documento pueden implementarse en el equipo lógico para su ejecución por un dispositivo de procesamiento adecuado, tal como un microprocesador, procesador de señal digital (DSP) y/o cualquier lógica programable adecuada, como un dispositivo de matriz de puertas programables (FPGA).

También debe entenderse que puede ser posible reutilizar las capacidades generales de procesamiento del decodificador. Esto puede hacerse, por ejemplo, mediante la reprogramación del equipo lógico existente o mediante la adición de nuevos componentes de equipo lógico.

La figura 17 ilustra una realización de un aparato 60 de ajuste de ganancia de acuerdo con la presente tecnología. Esta realización se basa en un procesador 110, por ejemplo un microprocesador, que ejecuta un componente 120 de equipo lógico para estimar la medida de precisión, un componente 130 de equipo lógico para determinar la ganancia de la corrección y un componente 140 de equipo lógico para ajustar la representación de ganancia. Estos componentes de equipo lógico se almacenan en la memoria 150. El procesador 110 se comunica con la memoria a través de un bus de sistema. Los parámetros $\hat{N}(b)$, $R(b)$, $\hat{E}(b)$ son recibidos por un controlador 160 de entrada/salida (I/O) que controla un bus I/O, al que están conectados el procesador 110 y la memoria 150. En esta realización, los parámetros recibidos por el controlador 160 de I/O se almacenan en la memoria 150, donde son procesados por los componentes de equipo lógico. Los componentes 120, 130 de equipo lógico pueden implementar la funcionalidad del bloque 62 en las realizaciones descritas anteriormente. El componente 140 de equipo lógico puede implementar la funcionalidad del bloque 64 en las realizaciones descritas anteriormente. La representación $\tilde{E}(b)$ de ganancia ajustada obtenida a partir del componente 140 de equipo lógico es emitida desde la memoria (150) por el controlador 160 de I/O a través del bus de I/O.

La figura 18 ilustra una realización de ajuste de ganancia de acuerdo con la presente tecnología con más detalle. Un estimador 200 de atenuación está configurado para utilizar la asignación $R(b)$ de bits recibida para determinar una atenuación de ganancia $t(R(b))$. El estimador 200 de atenuación puede, por ejemplo, ser implementado como una tabla de consulta o en un equipo lógico basándose en una ecuación lineal tal como la ecuación (14) anterior. La

asignación $R(b)$ de bits también es enviada a un estimador 202 de precisión de forma, que también recibe una dispersión estimada $p_{\max}(b)$ de la forma cuantificada, por ejemplo representada por la altura del impulso más alto en la representación $\hat{N}(b)$ de forma. El estimador 202 de precisión de forma puede, por ejemplo, ser implementado como una tabla de consulta. La atenuación estimada $t(R(b))$ y la precisión $A(b)$ de forma estimada se multiplican en un multiplicador 204. En una realización, este producto $t(R(b)) \cdot A(b)$ forma directamente la corrección $g_c(b)$ de ganancia. En otra realización, la corrección de ganancia $g_c(b)$ se forma de acuerdo con la ecuación (12) anterior. Esto requiere un conmutador 206 controlado por un comparador 208, que determina si la banda b de frecuencia es menor que un límite b_{THR} de frecuencia. Si este es el caso, entonces $g_c(b)$ es igual a $t(R(b)) \cdot A(b)$. De lo contrario $g_c(b)$ se establece en 1. La corrección $g_c(b)$ de ganancia es enviada a otro multiplicador 210, cuya otra entrada recibe la ganancia $g_{\text{RMA}}(b)$ de correspondencia RMS. La ganancia $g_{\text{RMA}}(b)$ de correspondencia RMS es determinada por una calculadora 212 de ganancia de correspondencia RMS basándose en la representación $N(b)$ de forma recibida y ancho de banda $BW(b)$ correspondiente, véase la ecuación (4) anterior. El producto resultante es enviado a otro multiplicador 214, que también recibe la representación $N(b)$ de forma y la representación $\hat{E}(b)$ de ganancia, y forma la síntesis $\hat{X}(b)$.

La detección de estabilidad descrita con referencia a la figura 10 puede incorporarse en la realización 2 así como las otras realizaciones descritas anteriormente. La figura 19 es un diagrama de flujo que ilustra el método de acuerdo con la presente tecnología. El paso S1 estima una medida $A(b)$ de precisión de la representación $\hat{N}(b)$ de forma. La medida de precisión puede derivarse, por ejemplo, de características de cuantificación de forma, tales como $R(b)$, $S(b)$, que indican la resolución de la cuantificación de forma. El paso S2 determina una corrección de ganancia, tal como $g_c(b)$, $g'_c(b)$, basándose en la medida de precisión estimada. El paso S3 ajusta la representación $\hat{E}(b)$ de ganancia basándose en la corrección de ganancia determinada.

La figura 20 es un diagrama de flujo que ilustra una realización del método de acuerdo con la tecnología actual, en la que la forma ha sido codificada utilizando un esquema de codificación de impulsos y la corrección de ganancia depende de una dispersión estimada $p_{\max}(b)$ de la forma cuantificada. Se supone que una medida de precisión ya se ha determinado en un paso S1 (figura 19). El paso S4 estima una atenuación de ganancia que depende de la velocidad de bits asignada. El paso S5 determina una corrección de ganancia basándose en la medida de precisión estimada y la atenuación de ganancia estimada. A continuación, el procedimiento procede al paso S3 (figura 19) para ajustar la representación de ganancia.

La figura 21 ilustra una realización de una red de acuerdo con la tecnología actual. Incluye un decodificador 300 provisto de un aparato de ajuste de ganancia de acuerdo con la presente tecnología. Esta realización ilustra un terminal de radio, pero también son factibles otros nodos de red. Por ejemplo, si se utiliza voz sobre IP (Protocolo de Internet) en la red, los nodos pueden comprender ordenadores.

En el nodo de red de la figura 21, una antena 302 recibe una señal de audio codificada. Una unidad 304 de radio transforma esta señal en parámetros de audio, los cuales son enviados al decodificador 300 para generar una señal de audio digital, como se describe con referencia a las diversas realizaciones anteriores. La señal de audio digital es entonces convertida y amplificada D/A en una unidad 306 y finalmente transmitida a un altavoz 308.

Aunque la descripción anterior se centra en la codificación de audio basada en transformada, los mismos principios también pueden aplicarse a la codificación de audio en el dominio tiempo con representaciones separadas de ganancia y forma, por ejemplo codificación CELP.

Los expertos en la técnica comprenderán que se pueden hacer diversas modificaciones y cambios en la presente tecnología sin apartarse del alcance de la misma, que se define por las reivindicaciones adjuntas.

Abreviaturas

ADPCM Modulación por código de impulso diferencial adaptativo

AMR Multivelocidad adaptativa

AMR-WB Ancho de banda de multivelocidad adaptativa

CELP Predicción lineal excitada por código

GSM-EFR Sistema global para comunicaciones móviles – Velocidad completa mejorada

DSP Procesador de señal digital

FPGA Matriz de puertas programable

IP Protocolo de Internet

MDCT Transformada discreta de coseno modificada

MSE Error cuadrático medio

5

QMF Filtro espejo en cuadratura

RMS Media cuadrática

10 VQ Cuantificación de vector

Referencias

15 [1] "ITU-T G.722.1 ANNEX C: A NEW LOW-COMPLEXITY 14 KHZ AUDIO CODING STANDARD", ICASSP 2006

[2] "ITU-T G.719: A NEW LOW-COMPLEXITY FULL-BAND (20 KHZ) AUDIO CODING STANDARD FOR HIGHQUALITY CONVERSATIONAL APPLICATIONS", WASPA 2009

20 [3] U. Mittal, J. Ashley, E. Cruz-Zeno, "Low Complexity Factorial Pulse Coding of MDCT Coefficients using Approximation of Combinatorial Functions", ICASSP 2007

[4] "7 kHz Audio Coding Within 64 kbit/s", [G.722], IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, 1988

REIVINDICACIONES

1.- Un método de ajuste de ganancia en la decodificación de audio que ha sido codificado con representaciones separadas de ganancia y forma, incluyendo dicho método los pasos de:

estimar (S1) una medida (A(b)) de precisión de la representación de forma ($\hat{N}(b)$) para una banda (b) de frecuencia, comprendiendo la banda (b) de frecuencia pluralidad de coeficientes, en el que la forma ha sido codificada utilizando un esquema de codificación de vector de impulso en el que los impulsos pueden ser añadidos uno encima de otro para formar impulsos de diferentes alturas, y la medida (A(b)) de precisión se basa en un número de impulsos (R(b)) y una altura de un impulso máximo ($p_{\max}(b)$);

determinar (S2) basándose en la medida (A(b)) de precisión estimada una corrección ($g_c(b)$) de ganancia;

ajustar (S3) la representación ($\hat{E}(b)$) de ganancia basándose en la corrección de ganancia determinada.

2.- El método de la reivindicación 1, en el que la corrección ($g_c(b)$) de ganancia también depende de la banda (b) de frecuencia.

3.- El método de cualquiera de las reivindicaciones precedentes, incluidos los pasos de:

estimar (S4) una atenuación ($t(R(b))$) de ganancia que depende de la velocidad (R(b)) de bits asignada;

determinar (S5) la corrección ($g_c(b)$) de ganancia basándose en la medida (A(b)) de precisión estimada y la atenuación ($t(R(b))$) de ganancia estimada.

4.- El método de la reivindicación 3, en el que la atenuación ($t(R(b))$) de ganancia es estimada a partir de una tabla (200) de búsqueda.

5.- El método de la reivindicación 3 ó 4, que incluye el paso de estimar (S5) la medida (A(b)) de precisión a partir de una tabla (200) de búsqueda.

6.- El método de la reivindicación 3 ó 4, que incluye el paso de estimar la medida (A(b)) de precisión a partir de una función lineal de la altura ($p_{\max}$) de altura de impulso máxima y la velocidad (R(b)) de bits asignada

7.- El método de cualquiera de las reivindicaciones precedentes, que incluye el paso de adaptar la corrección ($g_c(b)$) de ganancia a una clase de señal de audio determinada.

8.- Un aparato (60) de ajuste de ganancia para su uso en decodificar audio que ha sido codificado con representaciones de forma y ganancia separadas, incluyendo dichos aparatos:

un medidor (62) de precisión configurado para estimar una medida (A(b)) de precisión de la representación ($\hat{N}(b)$) de forma para una banda (b) de frecuencia, la banda (b) de frecuencia comprendiendo pluralidad de coeficientes, en el que la forma ha sido codificada utilizando un esquema de codificación de vector de impulso en el que los impulsos pueden ser añadidos uno encima de otro para formar impulsos de diferentes alturas, la medida (A(b)) de precisión se basa en un número de impulsos (R(b)) y una altura de un impulso máximo ($p_{\max}(b)$), y para determinar una corrección ($g_c(b)$) de ganancia, en el que la corrección ($g_c(b)$) de ganancia se determina basándose en la medición (A(b)) de precisión estimada;

un ajustador (64) de envolvente configurado para ajustar la representación ($\hat{E}(b)$) de ganancia basándose en la corrección de ganancia determinada.

9.- El aparato de la reivindicación 8, en el que la corrección ($g_c(b)$) de ganancia también depende de la banda (b) de frecuencia.

10.- El aparato de la reivindicación 8 ó 9, en el que el medidor de precisión incluye

un estimador (200) de atenuación configurado para estimar una atenuación ($t(R(b))$) de ganancia que depende de la velocidad (R(b)) de bits asignada;

un estimador (202) de precisión de forma configurado para estimar la medida (A(b)) de precisión;

un corrector (204, 206, 208) de ganancia configurado para determinar una corrección ($g_c(b)$) de ganancia basándose en la medida (A(b)) de precisión estimada y la atenuación ($t(R(b))$) de ganancia estimada.

11.- El aparato de la reivindicación 10, en el que el estimador (200) de atenuación se implementa como una tabla de

búsqueda.

12.- El aparato de la reivindicación 10 u 11, en el que el estimador (202) de precisión de forma es una tabla de búsqueda.

5 13.- El aparato de la reivindicación 10 u 11, en el que el estimador (202) de precisión de forma es configurado para estimar la medida ($A(b)$) de precisión a partir de una función lineal de la altura ($p_{\max}$) de impulso máxima y la velocidad ($R(b)$) de bits asignada.

10 14.- El aparato de cualquiera de las reivindicaciones 8-11, en el que el medidor (62) de precisión es configurado para adaptar la corrección ($g_c(b)$) de ganancia a una clase de señal de audio determinada.

15.- Un decodificador que incluye un aparato (60) de ajuste de ganancia de acuerdo con cualquiera de las reivindicaciones 8-14.

15 16.- Un nodo de red que incluye un decodificador de acuerdo con la reivindicación 15.

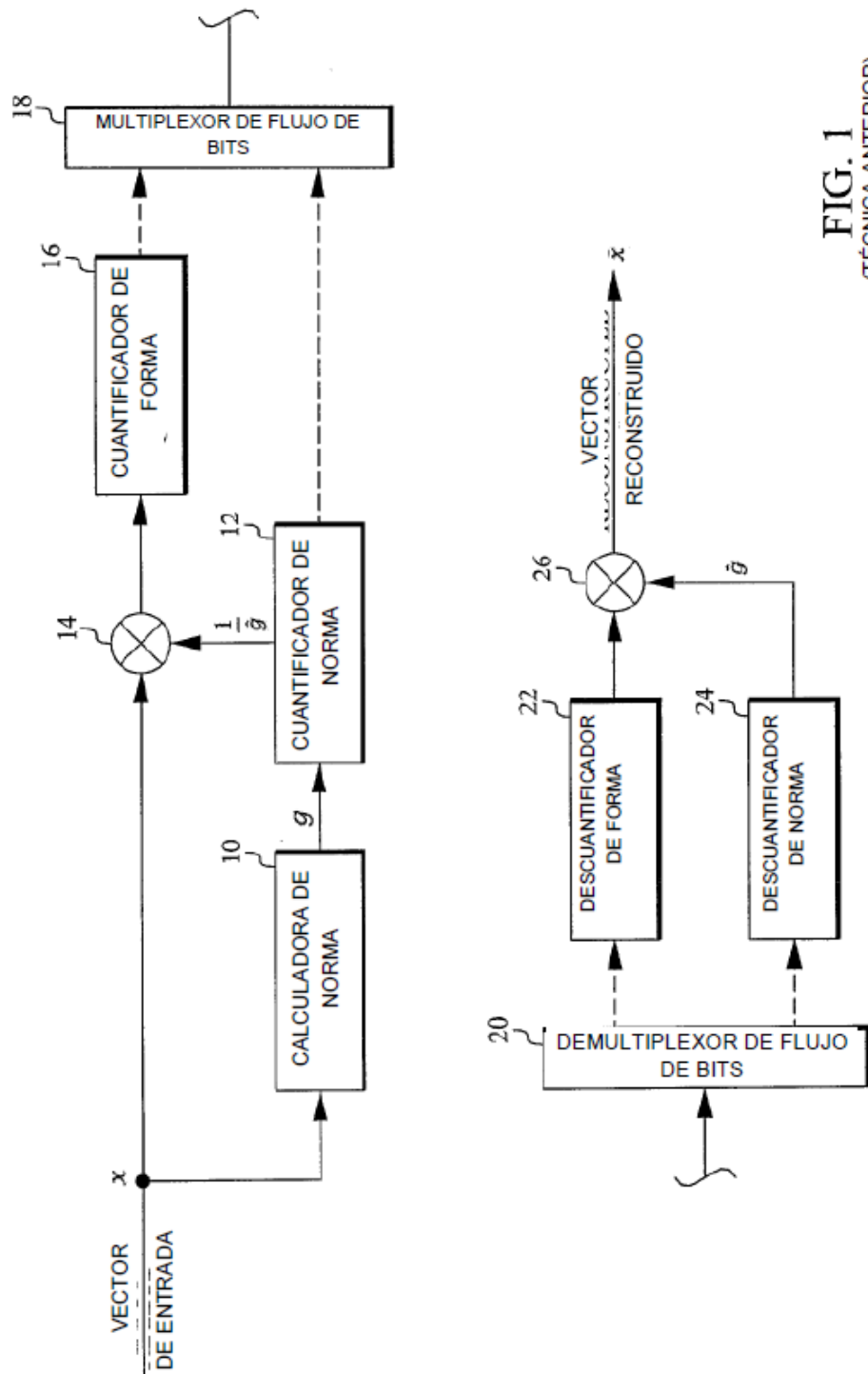


FIG. 1
(TÉCNICA ANTERIOR)

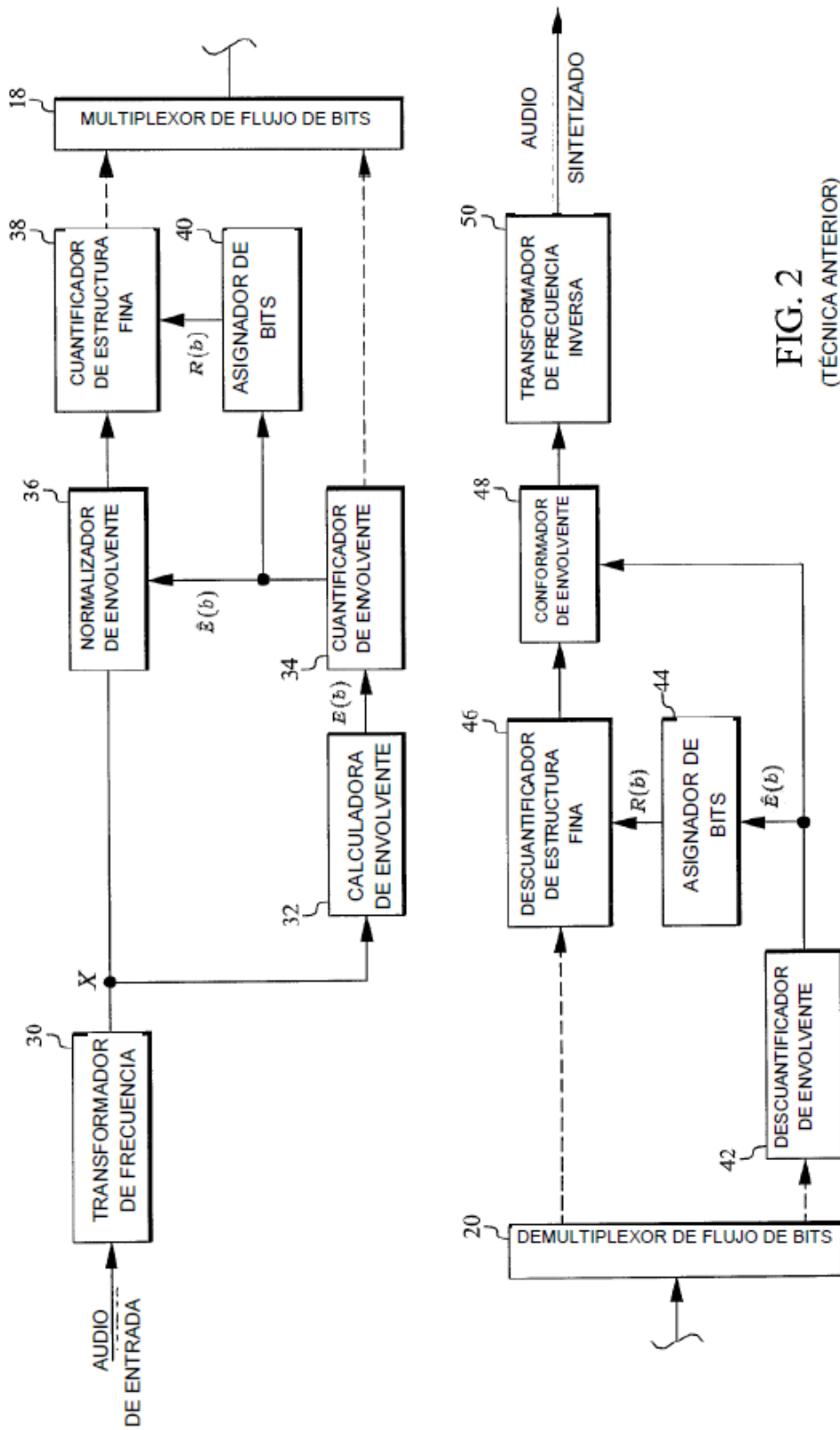
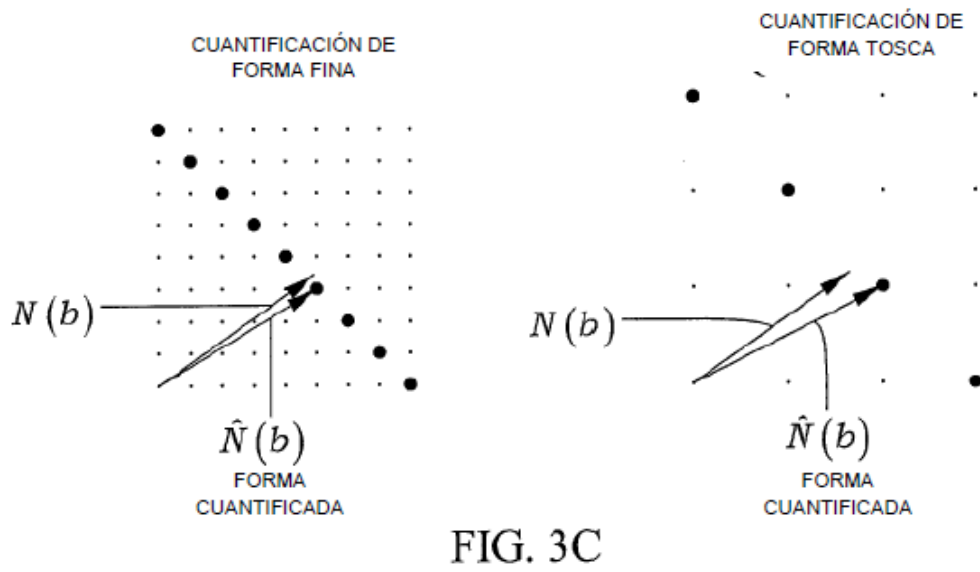
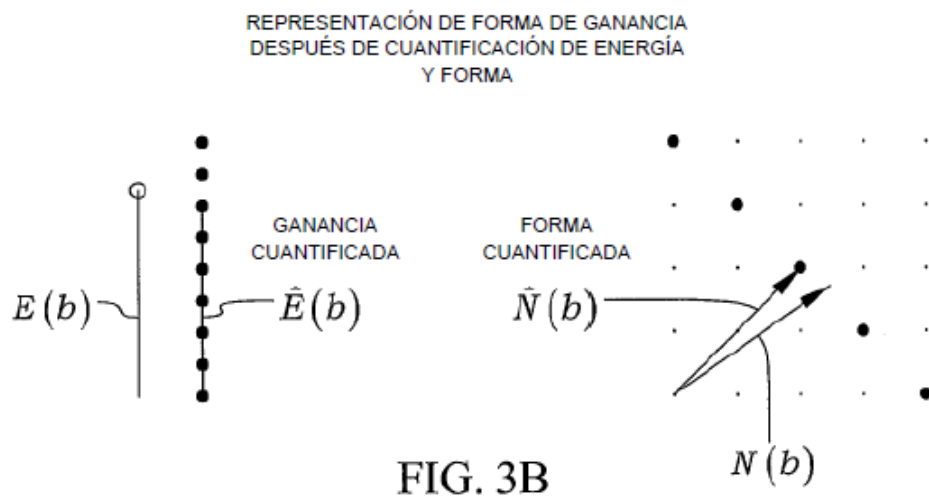
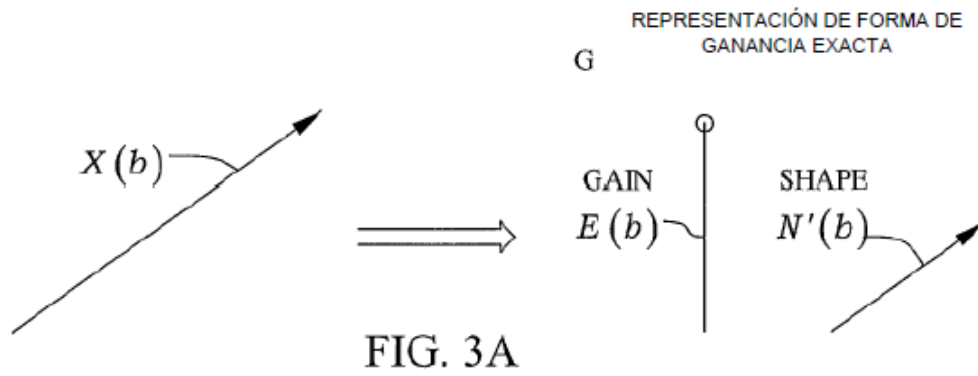


FIG. 2
(TÉCNICA ANTERIOR)



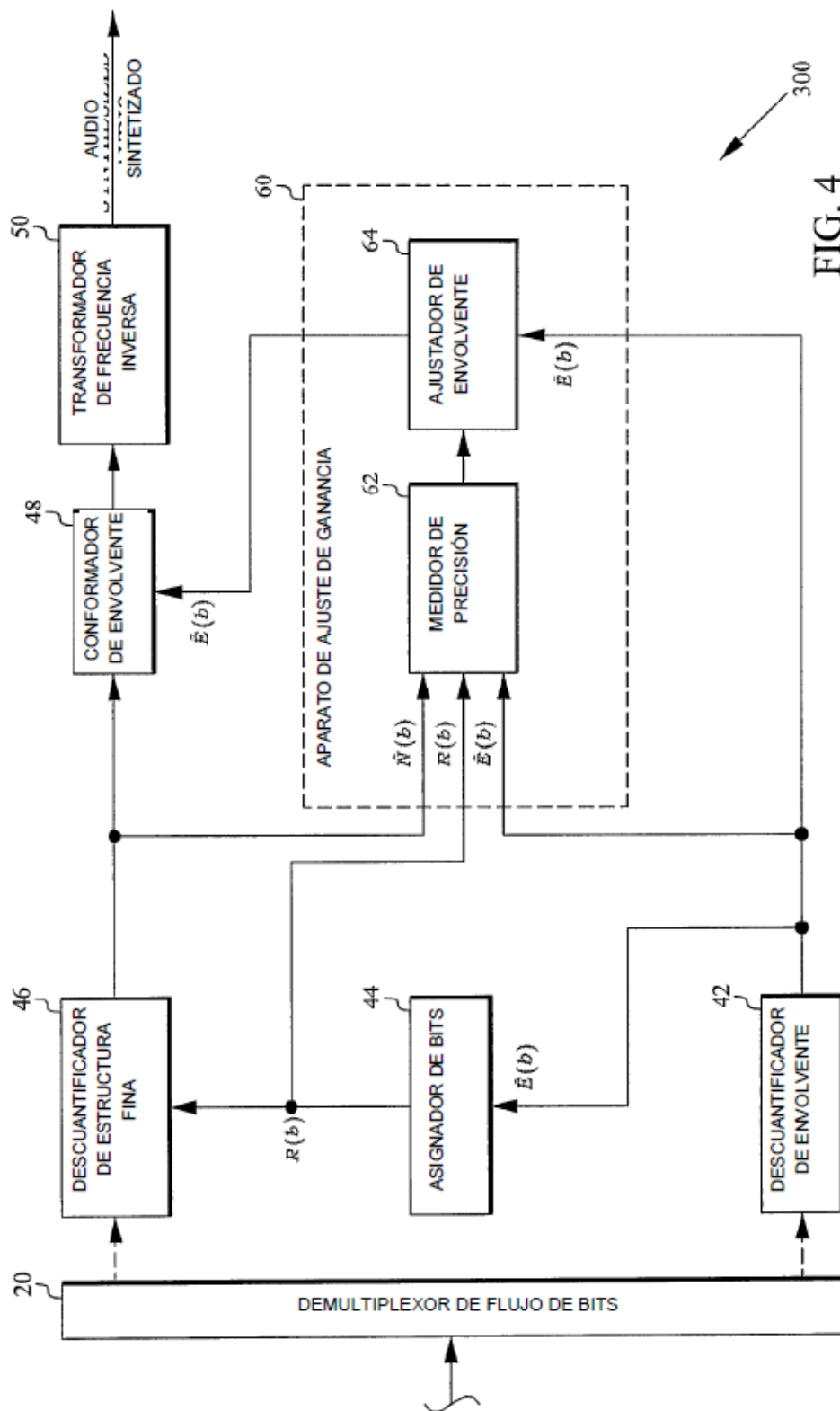
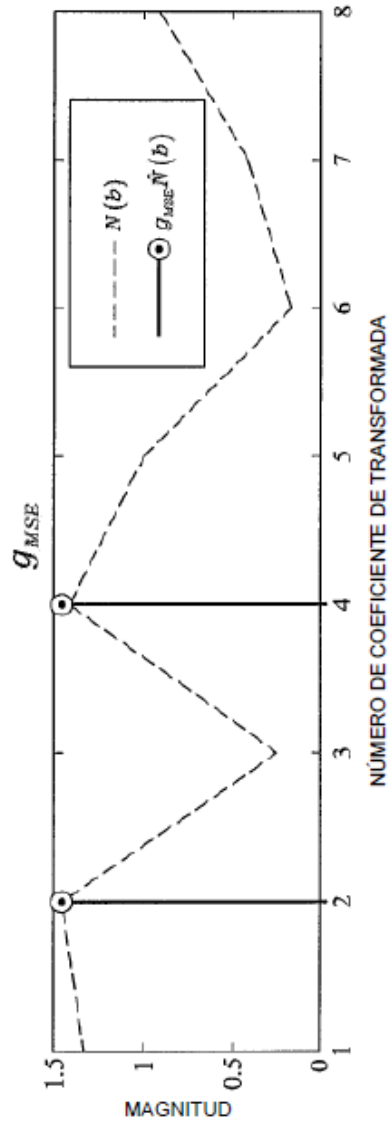
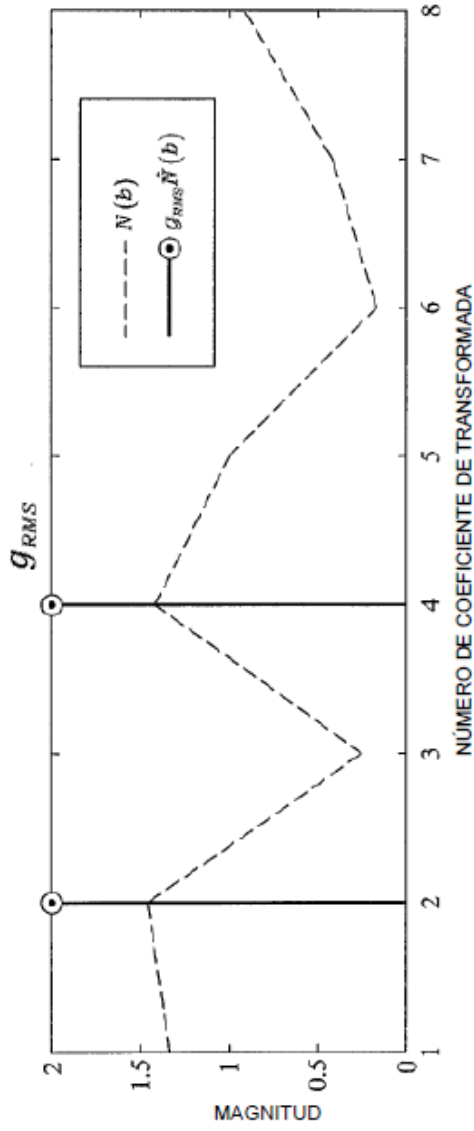


FIG. 4



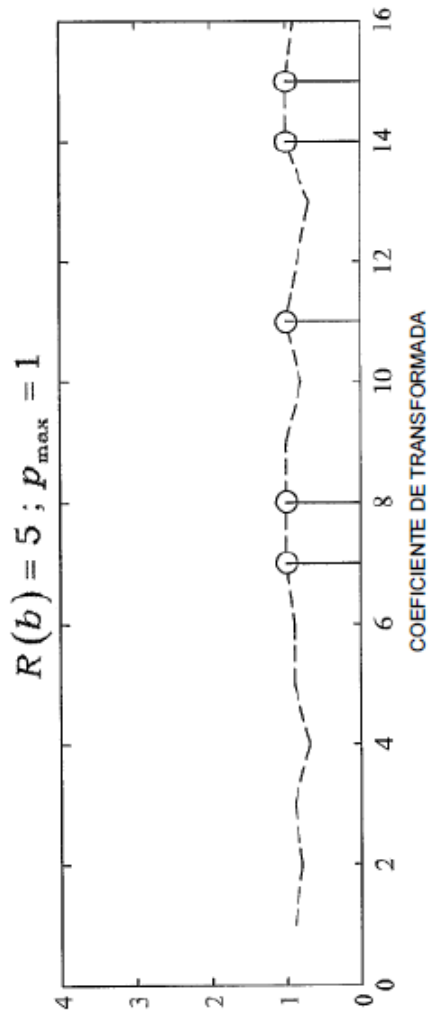


FIG. 6A

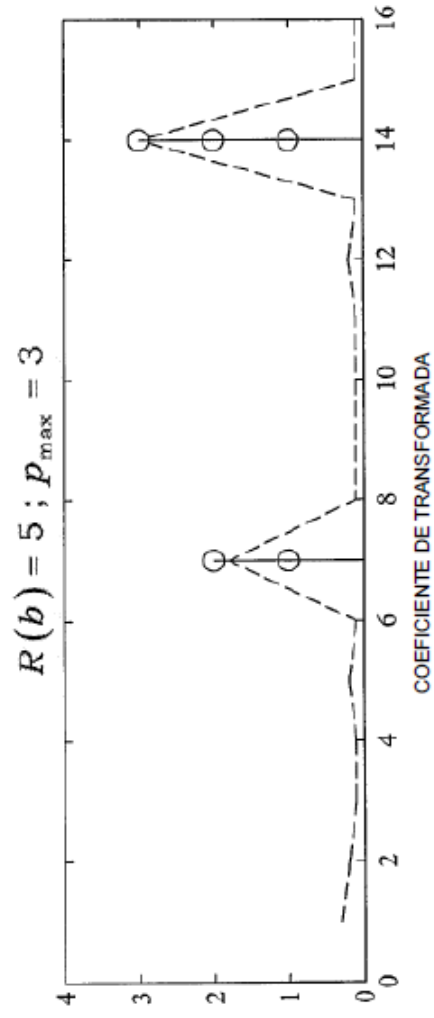


FIG. 6B

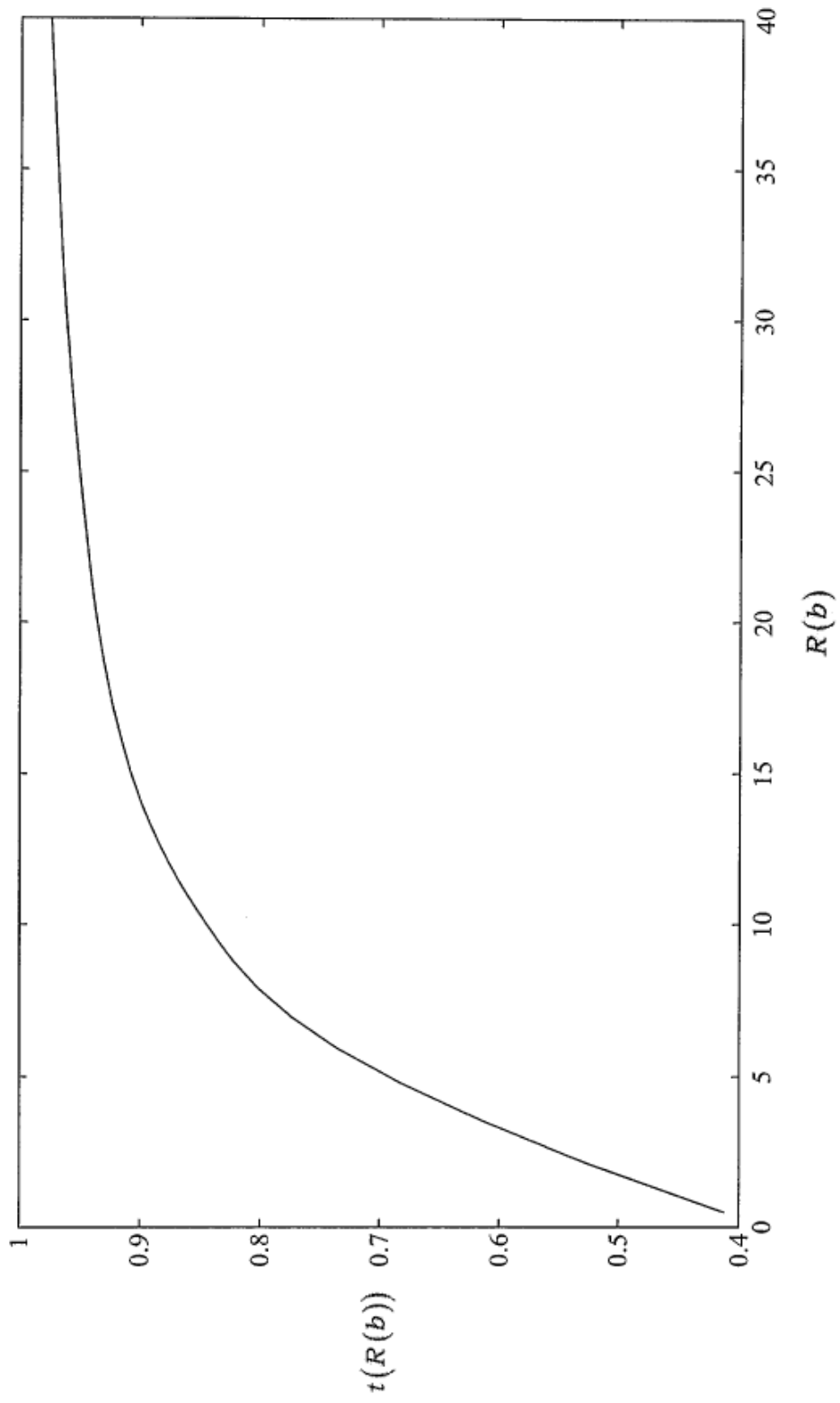


FIG. 7

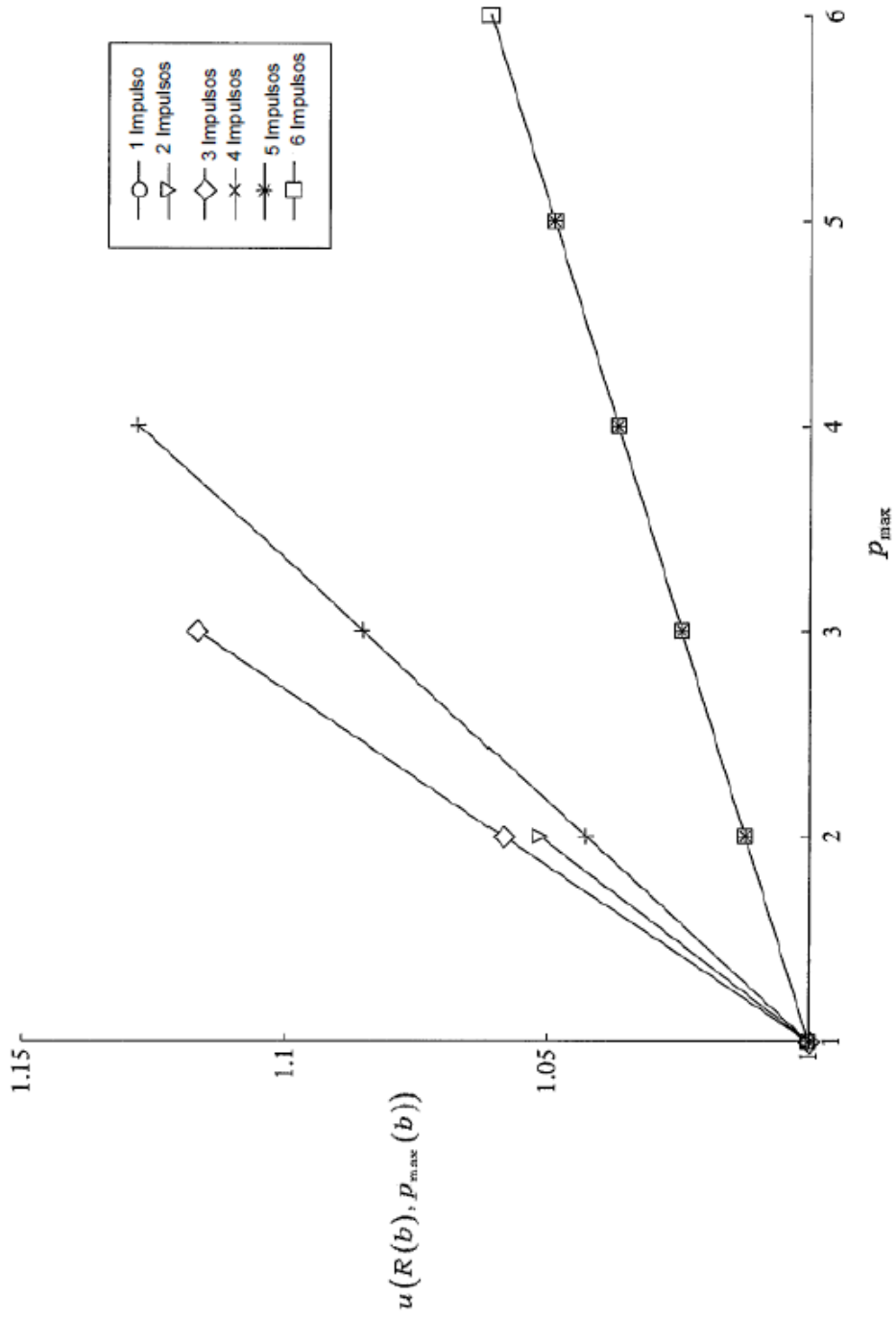


FIG. 8

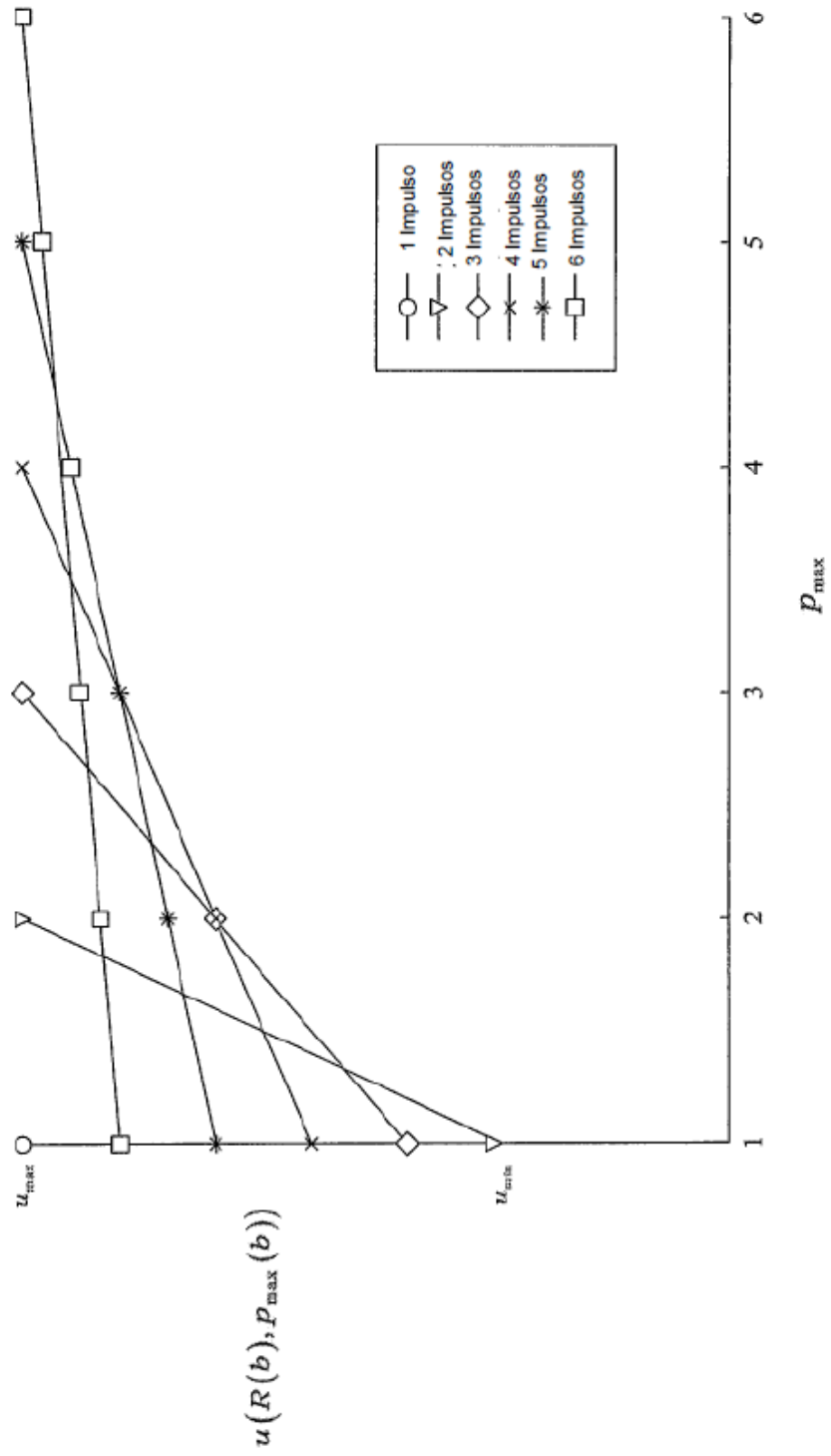
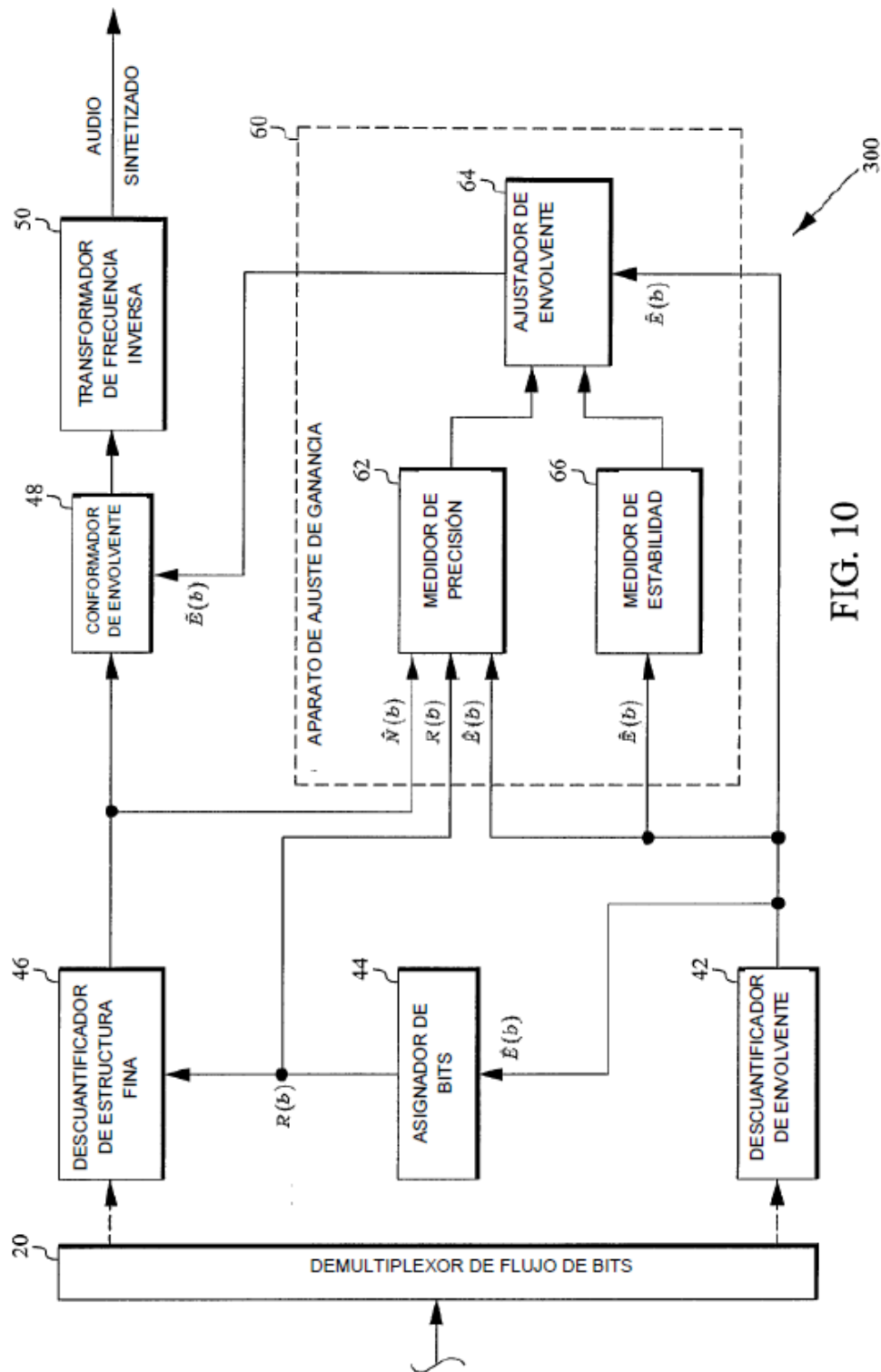


FIG. 9



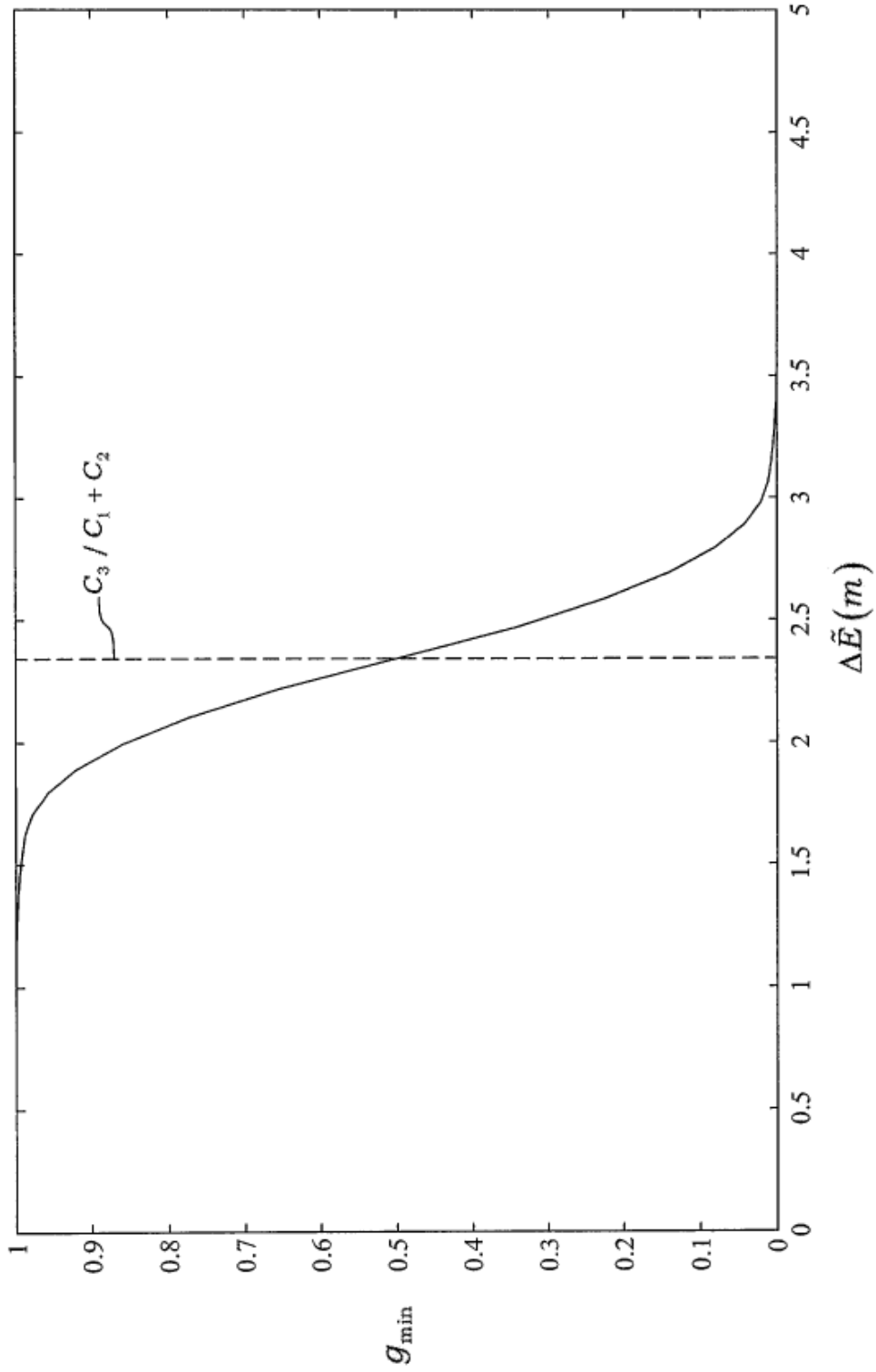


FIG. 11

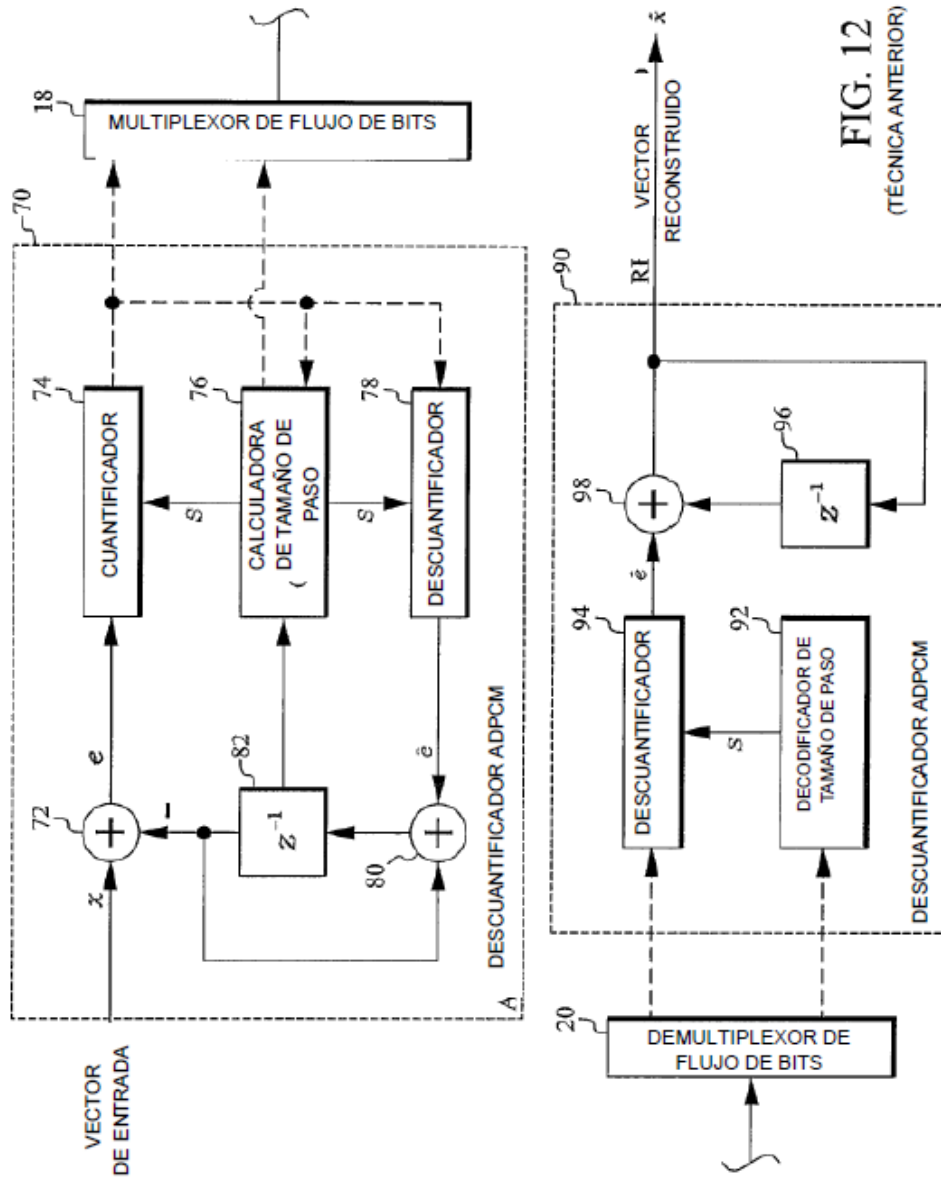


FIG. 12
(TÉCNICA ANTERIOR)

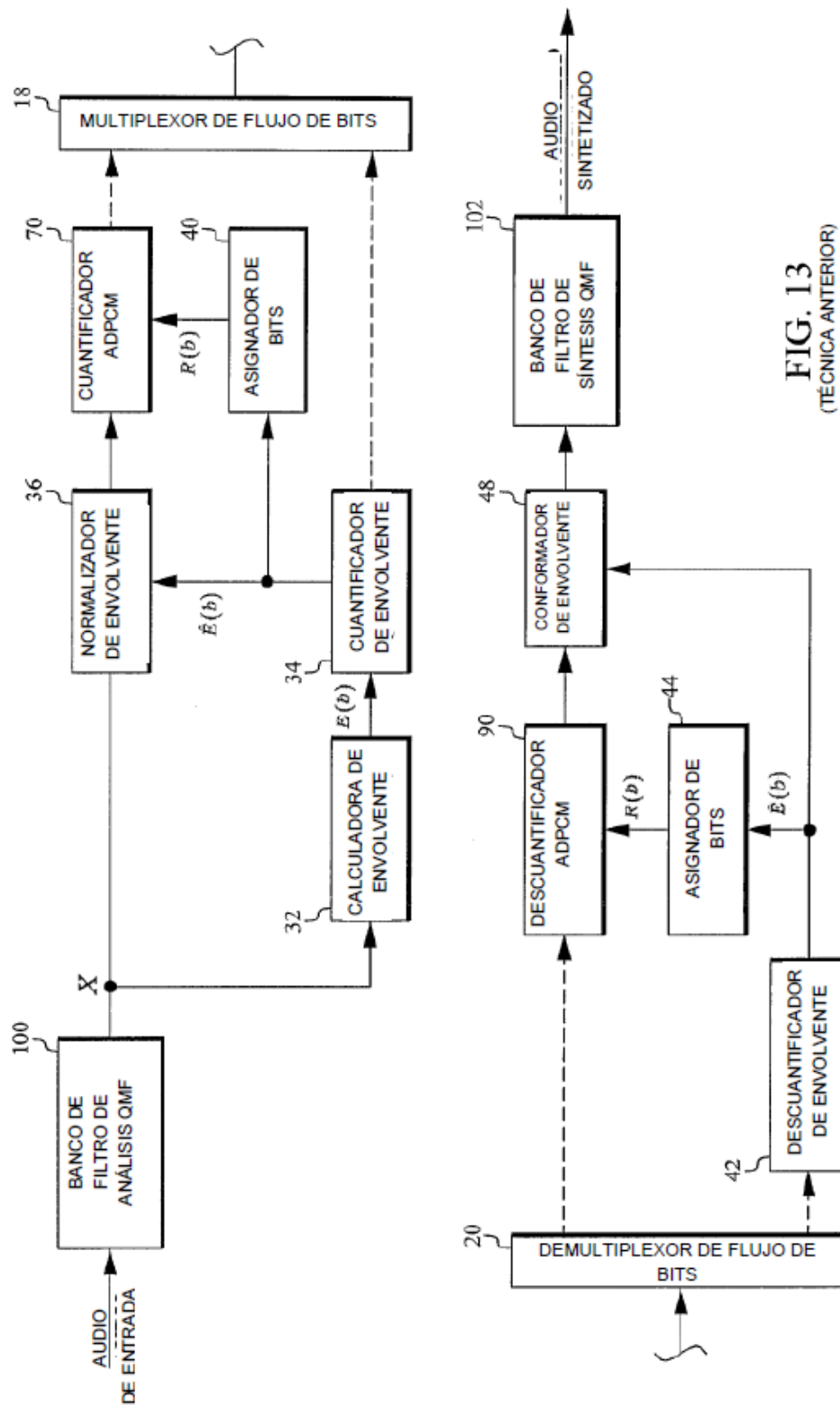


FIG. 13
(TÉCNICA ANTERIOR)

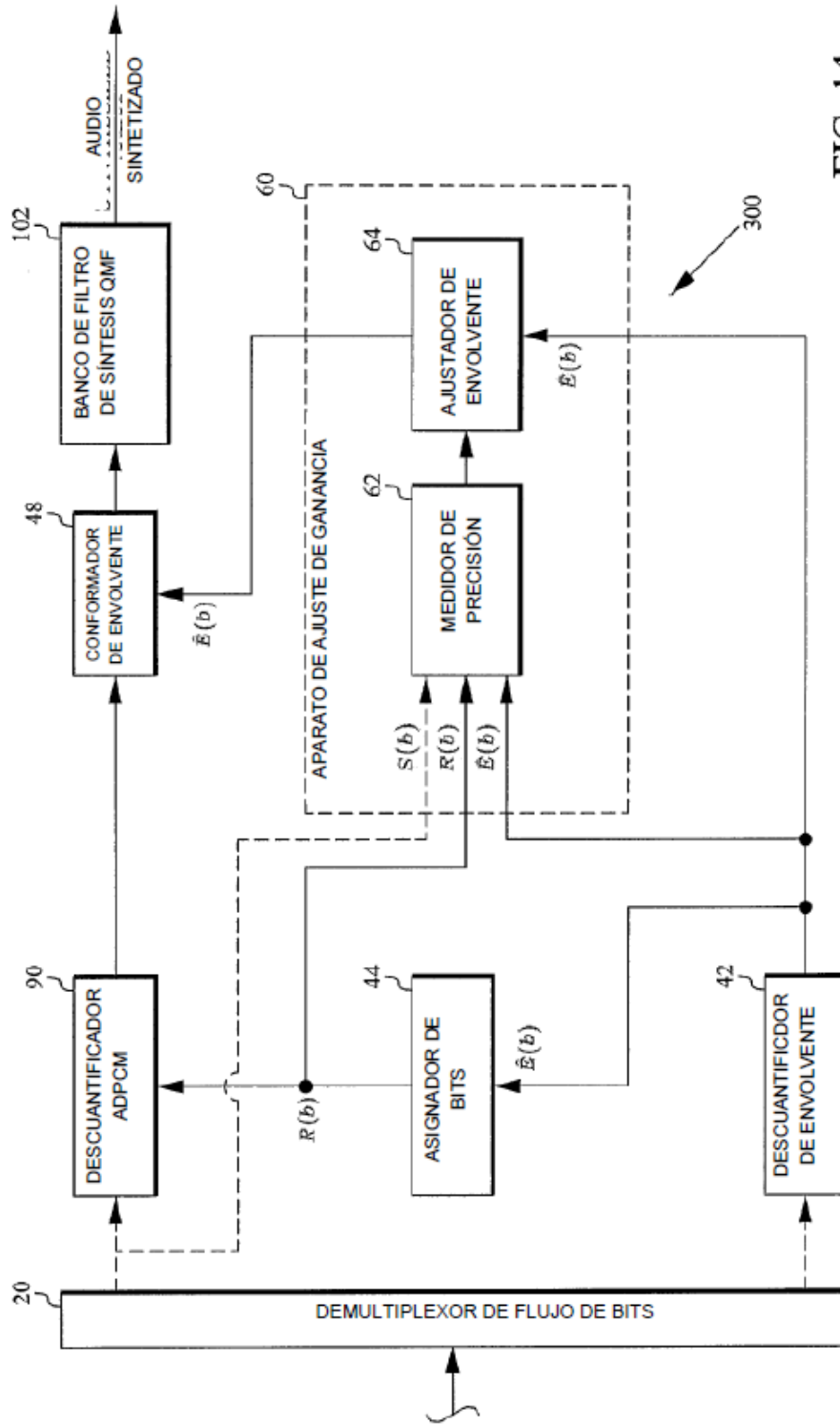


FIG. 14

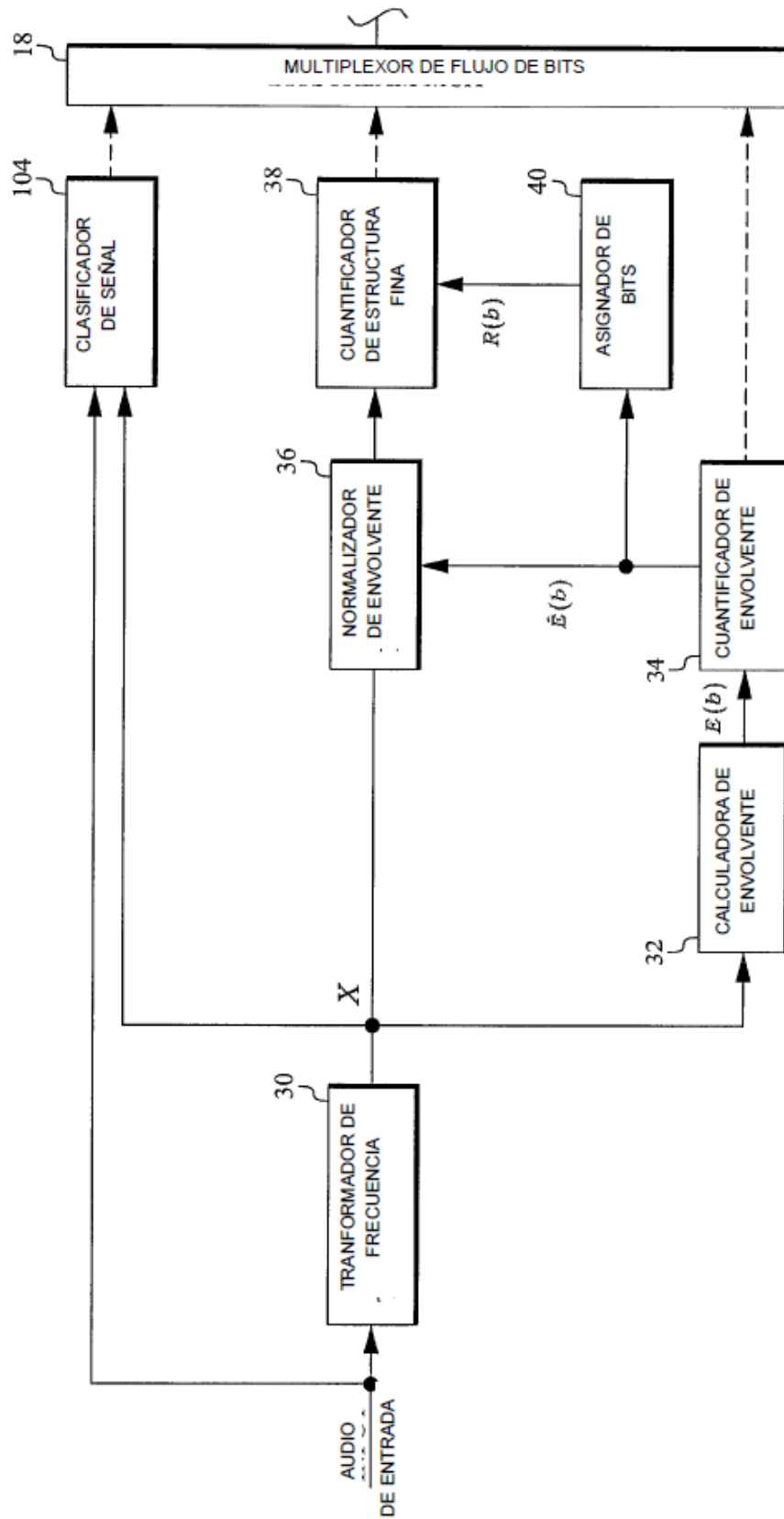


FIG. 15

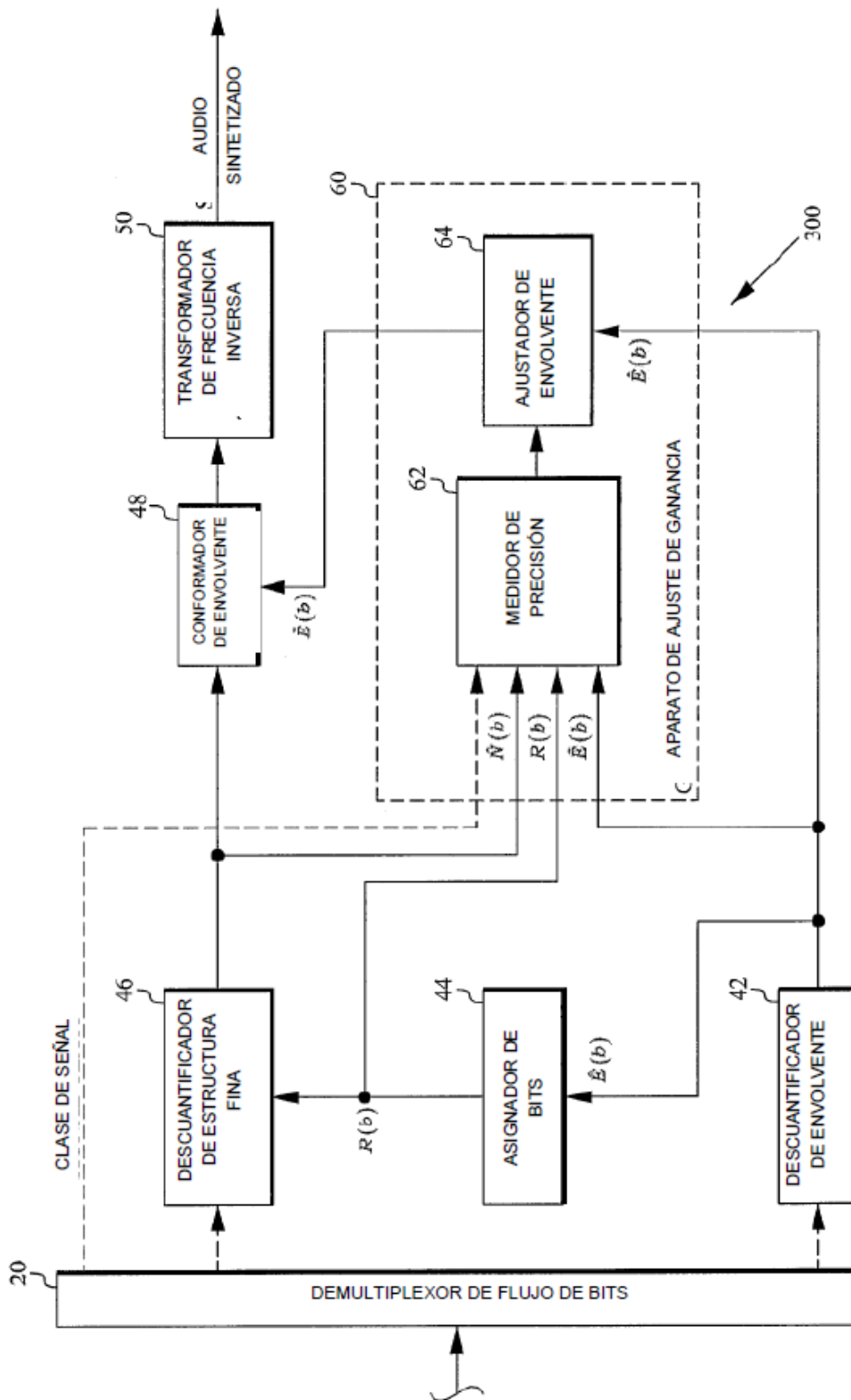


FIG. 16

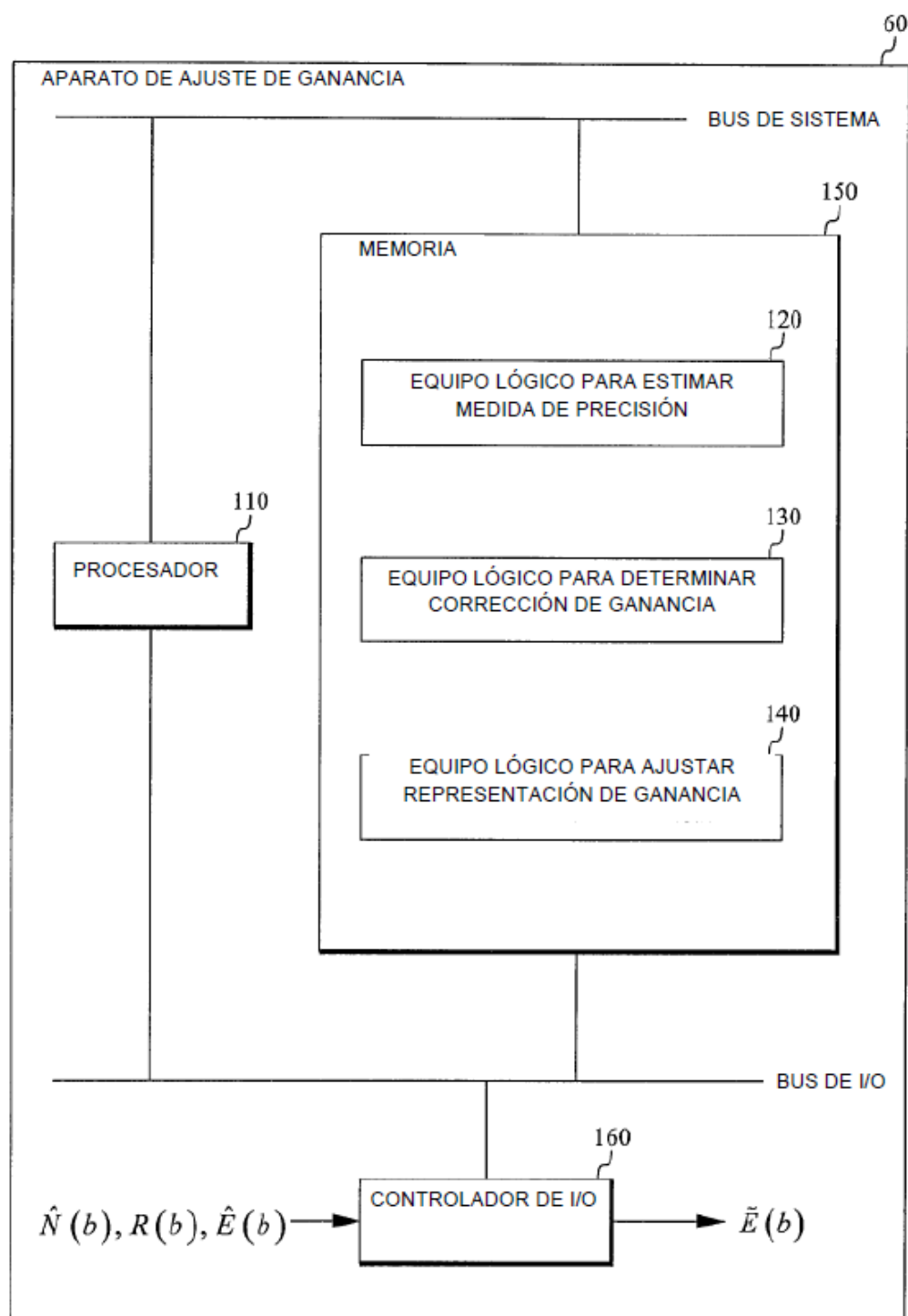
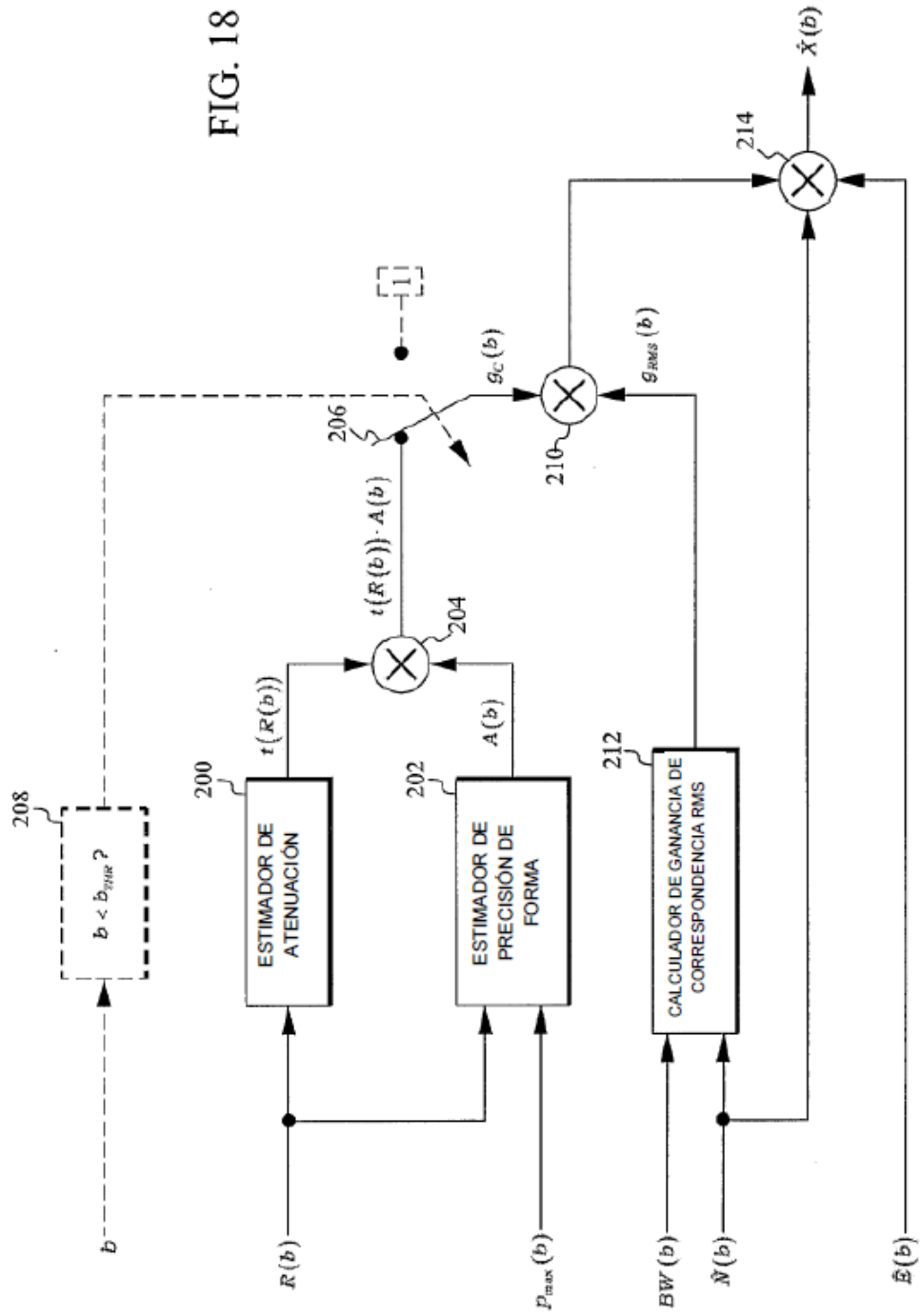


FIG. 17

FIG. 18



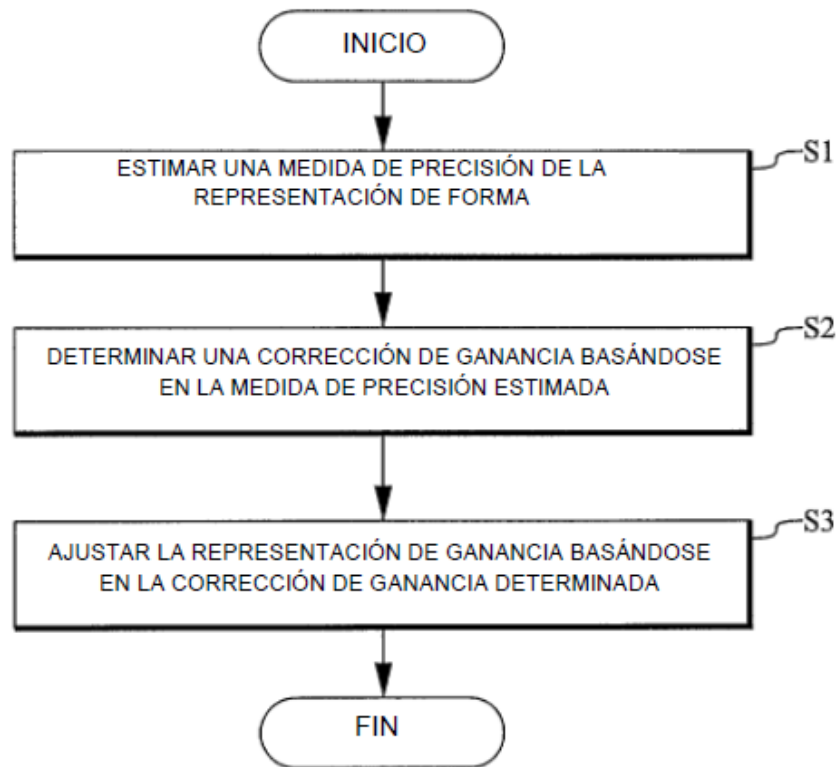


FIG. 19

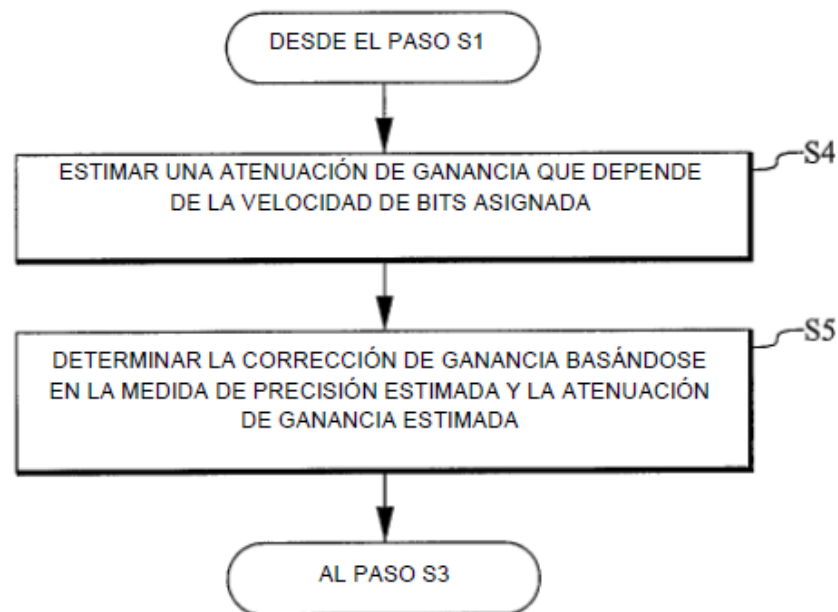


FIG. 20

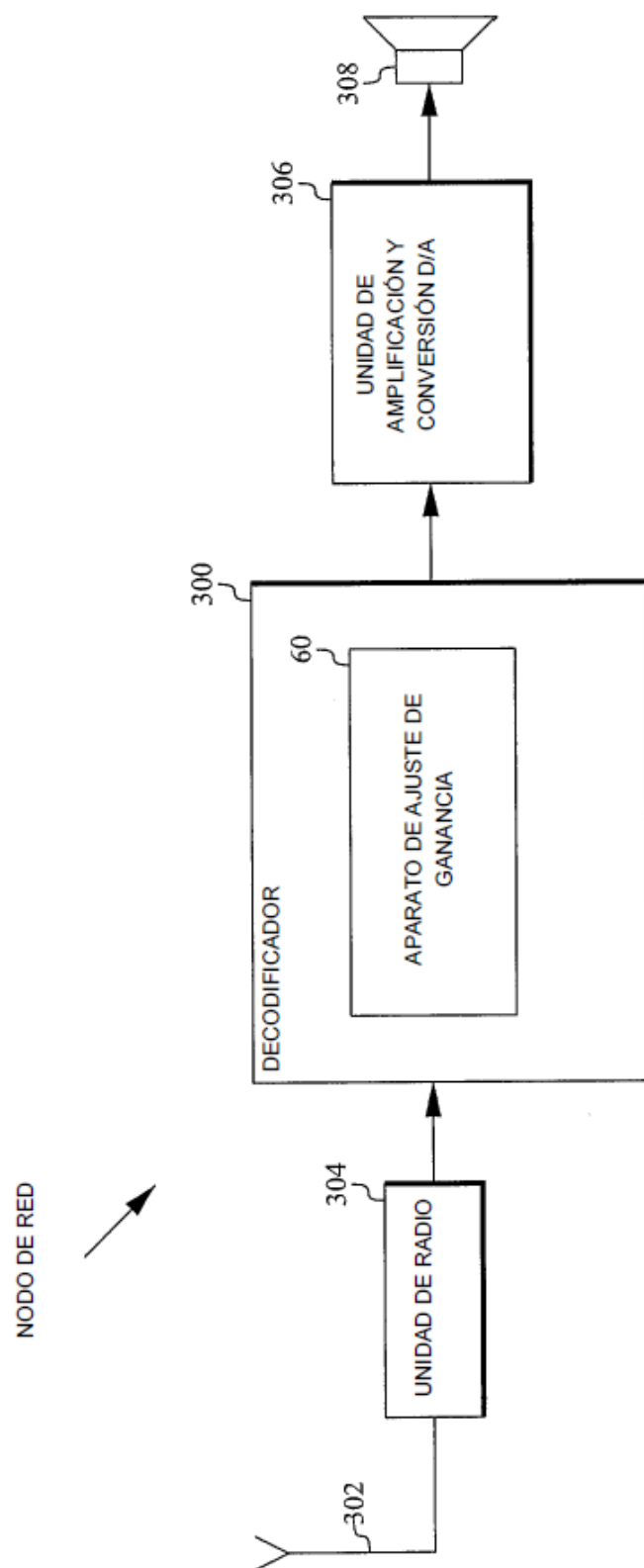


FIG. 21