

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 643 746**

51 Int. Cl.:

G10L 19/035 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.11.2013 PCT/JP2013/006948**

87 Fecha y número de publicación internacional: **19.06.2014 WO14091694**

96 Fecha de presentación y número de la solicitud europea: **26.11.2013 E 13862073 (7)**

97 Fecha y número de publicación de la concesión europea: **12.07.2017 EP 2933799**

54 Título: **Dispositivo de codificación de audio de voz, dispositivo de descodificación de audio de voz, método de codificación de audio de voz y método de descodificación de audio de voz**

30 Prioridad:

13.12.2012 JP 2012272571

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

24.11.2017

73 Titular/es:

**PANASONIC INTELLECTUAL PROPERTY CORPORATION OF AMERICA (100.0%)
20000 Mariner Avenue Suite 200
Torrance CA 90503, US**

72 Inventor/es:

**LIU, ZONGXIAN;
NAGISETTY, SRIKANTH y
OSHIKIRI, MASAHIRO**

74 Agente/Representante:

MILTENYI, Peter

ES 2 643 746 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

Dispositivo de codificación de audio de voz, dispositivo de descodificación de audio de voz, método de codificación de audio de voz y método de descodificación de audio de voz

Campo técnico

- 5 La presente invención se refiere a un aparato de codificación de voz/audio, a un aparato de descodificación de voz/audio, a un método de codificación de voz/audio y a un método de descodificación de voz/audio que usan un esquema de codificación de transformación.

Técnica anterior

- 10 Como esquema que puede codificar de manera eficiente una señal de voz o una señal de música en una banda completa (FB) de 0,02 a 20 kHz, existe una técnica normalizada en la ITU-T (sector de normalización de las telecomunicaciones de la unión internacional de telecomunicaciones). Esta técnica transforma una señal de entrada en una señal de dominio de frecuencia y codifica una banda de hasta 20 kHz (codificación de transformación).

- 15 En el presente documento, la codificación de transformación es un esquema de codificación que transforma una señal de entrada de un dominio de tiempo a un dominio de frecuencia usando una transformación tiempo/frecuencia tal como la transformada de coseno discreta (DCT) o transformada de coseno discreta modificada (MDCT) para permitir que una señal se mapee con correspondencia precisa con características acústicas.

En la codificación de transformación, se divide un coeficiente espectral en una pluralidad de subbandas de frecuencia. En la codificación de cada subbanda, asignar más bits de cuantificación a una banda que es perceptivamente importante para los oídos humanos hace posible mejorar la calidad global del sonido.

- 20 Con el fin de conseguir este objetivo, están llevándose a cabo estudios sobre esquemas de asignación de bits eficientes, y por ejemplo, se conoce una técnica dada a conocer en la bibliografía no de patente 1 (denominada a continuación en el presente documento "NPL"). A continuación en el presente documento, se describirá el esquema de asignación de bits dado a conocer en la NPL1 usando la figura 1 y la figura 2.

- 25 La figura 1 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de voz/audio dado a conocer en la NPL 1. Una señal de entrada muestreada a 48 kHz se introduce en el detector transitorio 11 y la sección de transformación 12 del aparato de codificación de voz/audio.

- 30 El detector transitorio 11 detecta, a partir de la señal de entrada, o bien una trama transitoria correspondiente a un borde delantero o un borde final de voz o bien una trama estacionaria correspondiente a una sección de voz distinta de esta, y la sección de transformación 12 aplica, a la trama de la señal de entrada, transformación de resolución de alta frecuencia o transformación de resolución de baja frecuencia dependiendo de si la trama detectada por el detector transitorio 11 es una trama transitoria o una trama estacionaria, y adquiere un coeficiente espectral (o coeficiente de transformación).

- 35 La sección de estimación de norma 13 divide el coeficiente espectral obtenido en la sección de transformación 12 en bandas de anchos de banda diferentes. La sección de estimación de norma 13 estima una norma (o energía) de cada banda dividida.

La sección de cuantificación de norma 14 determina una envolvente espectral compuesta por las normas de todas las bandas basándose en la norma de cada banda estimada por la sección de estimación de norma 13 y cuantifica la envolvente espectral determinada.

- 40 La sección de normalización de espectro 15 normaliza el coeficiente espectral obtenido por la sección de transformación 12 según la norma cuantificada por la sección de cuantificación de norma 14.

La sección de ajuste de norma 16 ajusta la norma cuantificada por la sección de cuantificación de norma 14 basándose en la ponderación espectral adaptativa.

La sección de asignación de bits 17 asigna bits disponibles para cada banda en una trama usando la norma de cuantificación ajustada por la sección de ajuste de norma 16.

- 45 La sección de codificación de vectores de red 18 realiza la codificación de vectores de red en el coeficiente espectral normalizado por la sección de normalización de espectro 15 usando bits asignados para cada banda por la sección de asignación de bits 17.

- 50 La sección de ajuste de nivel de ruido 19 estima el nivel del coeficiente espectral antes de codificarlo en la sección de codificación de vectores de red 18 y codifica el nivel estimado. De esta manera se obtiene un índice de ajuste de nivel de ruido.

El multiplexor 20 multiplexa una configuración de trama de la señal de entrada adquirida por la sección de

- transformación 12, es decir, un indicador de señal transitoria que indica si la trama es una trama estacionaria o una trama transitoria, la norma cuantificada por la sección de cuantificación de norma 14, el vector de codificación de red obtenido por la sección de codificación de vectores de red 18 y el índice de ajuste de nivel de ruido obtenido por la sección de ajuste de nivel de ruido 19, y forma un flujo de bits y transmite el flujo de bits a un aparato de descodificación de voz/audio.
- 5 La figura 2 es un diagrama de bloques que ilustra una configuración del aparato de descodificación de voz/audio dado a conocer en la NPL 1. El aparato de descodificación de voz/audio recibe el flujo de bits transmitido desde el aparato de codificación de voz/audio y el demultiplexor 21 desmultiplexa el flujo de bits.
- 10 La sección de descuantificación de norma 22 descuantifica la norma cuantificada, adquiere una envolvente espectral compuesta por normas de todas las bandas, y la sección de ajuste de norma 23 ajusta la norma descuantificada por la sección de descuantificación de norma 22 basándose en una ponderación espectral adaptativa.
- La sección de asignación de bits 24 asigna bits disponibles para cada banda en una trama usando las normas ajustadas por la sección de ajuste de norma 23. Es decir, la sección de asignación de bits 24 recalcula la asignación de bits indispensable para descodificar el código de vector de red del coeficiente espectral normalizado.
- 15 La sección de descodificación de red 25 descodifica un indicador de señal transitoria, descodifica el vector de codificación de red basándose en una configuración de trama indicada por el indicador de señal transitoria descodificado y los bits asignados por la sección de asignación de bits 24 y adquiere un coeficiente espectral.
- El generador de llenado espectral 26 regenera un coeficiente espectral de baja frecuencia al que no se han asignado bits usando un libro de códigos creado basándose en el coeficiente espectral descodificado por la sección de descodificación de red 25. El generador de llenado espectral 26 ajusta el nivel del coeficiente espectral regenerado usando un índice de ajuste de nivel de ruido. Además, el generador de llenado espectral 26 regenera un coeficiente espectral sin codificar de alta frecuencia usando un coeficiente espectral codificado de baja frecuencia.
- 20 El sumador 27 suma el coeficiente espectral descodificado y el coeficiente espectral regenerado, y genera un coeficiente espectral normalizado.
- 25 La sección de conformación de envolvente 28 aplica la envolvente espectral descuantificada por la sección de descuantificación de norma 22 al coeficiente espectral normalizado generado por el sumador 27 y genera un coeficiente espectral de banda completa.
- La sección de transformación inversa 29 aplica una transformación inversa tal como la transformada de coseno discreta modificada inversa (IMDCT) al coeficiente espectral de banda completa generado por la sección de conformación de envolvente 28 para transformarlo en una señal de dominio de tiempo. En este caso, la transformación inversa con resolución de alta frecuencia se aplica a un caso con una trama estacionaria y la transformación inversa con resolución de baja frecuencia se aplica a un caso con una trama transitoria.
- 30 En G.719, los coeficientes espectrales se dividen en grupos espectrales. Cada grupo espectral se divide en bandas con subvectores de igual longitud tal como se muestra en la figura 3. Los subvectores tienen diferente longitud de un grupo a otro y esta longitud aumenta a medida que aumenta la frecuencia. Con respecto a la resolución de transformación, se usa resolución de frecuencia más alta para frecuencias bajas, mientras que se usa una resolución de frecuencia más baja para frecuencias altas. Tal como se describe en G.719, la agrupación permite un uso eficiente de los bits disponibles durante la codificación.
- 35 En G.719, el esquema de asignación de bits es idéntico en un aparato de codificación y un aparato de descodificación. En este caso, el esquema de asignación de bits se describirá usando la figura 4.
- 40 Tal como se muestra en la figura 4, en la etapa (abreviada a continuación en el presente documento como "ST") 31, se ajustan normas cuantificadas antes de la asignación de bits para ajustar la ponderación psicoacústica y los efectos de enmascaramiento.
- En ST32, se identifican subbandas que tienen una norma máxima de entre todas las subbandas y en ST33, se asigna un bit a cada coeficiente espectral para las subbandas que tienen la norma máxima. Es decir, se asignan tantos bits como coeficientes espectrales.
- 45 En ST34, se reducen las normas según los bits asignados, y en ST35, se determina si el número de bits asignables restante es de 8 o más. Cuando el número de bits asignables restante es de 8 o más, el flujo vuelve a ST32 y cuando el número de bits asignables restante es menor que 8, se termina el procedimiento de asignación de bits.
- 50 Por tanto, en el esquema de asignación de bits, los bits disponibles dentro de una trama se asignan entre subbandas usando las normas de cuantificación ajustadas. Los coeficientes espectrales normalizados se codifican mediante codificación de vectores de red usando los bits asignados a cada subbanda.

Lista de referencias

Bibliografía no de patente

NPL1

Recomendación G.719 de ITU-T, "Low-complexity full-band audio coding for high-quality conversational applications", ITU-T, 2009. **Sumario de la invención**

5 **Problema técnico**

Sin embargo, el esquema de asignación de bits anterior no tiene en cuenta características de la señal de entrada al agrupar bandas espectrales, y por lo tanto tiene un problema en cuanto a que no es posible una asignación de bits eficiente y no puede esperarse una mejora adicional de la calidad de sonido.

10 La patente europea EP0259553 describe la designación de bits (asignación de bits) a una pluralidad de subbandas, en las que los niveles de energía de pico (correspondientes a picos en las envolventes de energía) de las subbandas respectivas se agrupan en una pluralidad de grupos de banda. La designación de bits se realiza de manera que algunas bandas que tienen energías de pico más grandes reciben designaciones de bits más grandes (truncamiento de tablas). 5

15 Un objeto de la presente invención es proporcionar un aparato de codificación de voz/audio, un aparato de decodificación de voz/audio, un método de codificación de voz/audio y un método de decodificación de voz/audio que puedan realizar una asignación de bits eficiente y mejorar la calidad de sonido.

Solución al problema

Un aparato de codificación de voz/audio de la presente invención incluye las características de la reivindicación 1.

20 Un aparato de decodificación de voz/audio según la presente invención incluye las características de la reivindicación 4.

Un método de codificación de voz/audio según la presente invención incluye las características de la reivindicación 7

Un método de decodificación de voz/audio según la presente invención incluye las características de la reivindicación 8.

Efectos ventajosos de la invención

25 Según la presente invención, es posible realizar una asignación de bits eficiente y mejorar la calidad de sonido.

Breve descripción de los dibujos

La figura 1 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de voz/audio dado a conocer en la NPL 1;

30 la figura 2 es un diagrama de bloques que ilustra una configuración de un aparato de decodificación de voz/audio dado a conocer en la NPL 1;

la figura 3 es un diagrama que ilustra la agrupación de coeficientes espectrales en un modo estacionario dado a conocer en la NPL 1;

la figura 4 es un diagrama de flujo que ilustra un esquema de asignación de bits dado a conocer en la NPL 1;

35 la figura 5 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de voz/audio según una realización de la presente invención;

la figura 6 es un diagrama de bloques que ilustra una configuración de un aparato de decodificación de voz/audio según una realización de la presente invención;

la figura 7 es un diagrama de bloques que ilustra una configuración interna de la sección de asignación de bits mostrada en la figura 5;

40 las figuras 8A a 8C son diagramas proporcionados para describir un método de agrupación según una realización de la presente invención; y

la figura 9 es un diagrama que ilustra una variación de norma.

Descripción de realizaciones

45 A continuación en el presente documento, se describirán en detalle realizaciones de la presente invención con referencia a los dibujos adjuntos.

(Realización)

La figura 5 es un diagrama de bloques que ilustra una configuración del aparato de codificación de voz/audio 100 según una realización de la presente invención. Una señal de entrada muestreada a 48 kHz se introduce en el detector transitorio 101 y la sección de transformación 102 del aparato de codificación de voz/audio 100.

5 El detector transitorio 101 detecta, a partir de una señal de entrada, o bien una trama transitoria correspondiente a un borde delantero o un borde final de voz o bien una trama estacionaria correspondiente a una sección de voz distinta de esta, y emite el resultado de detección a la sección de transformación 102. La sección de transformación 102 aplica, a la trama de la señal de entrada, transformación de resolución de alta frecuencia o transformación de resolución de baja frecuencia dependiendo de si el resultado de detección emitido desde el detector transitorio 101 es una trama transitoria o una trama estacionaria, y adquiere un coeficiente espectral (o coeficiente de transformación) y emite el coeficiente espectral a la sección de estimación de norma 103 y la sección de normalización de espectro 105. La sección de transformación 102 emite una configuración de trama que es el resultado de detección emitido desde el detector transitorio 101, es decir, un indicador de señal transitoria que indica si la trama es una trama estacionaria o una trama transitoria al multiplexor 110.

15 La sección de estimación de norma 103 divide el coeficiente espectral emitido desde la sección de transformación 102 en bandas de anchos de banda diferentes y estima una norma (o energía) de cada banda dividida. La sección de estimación de norma 103 emite la norma estimada de cada banda a la sección de cuantificación de norma 104.

La sección de cuantificación de norma 104 determina una envolvente espectral compuesta por normas de todas las bandas basándose en normas de bandas respectivas emitidas desde la sección de estimación de norma 103, cuantifica la envolvente espectral determinada y emite la envolvente espectral cuantificada a la sección de normalización de espectro 105 y la sección de ajuste de norma 106.

20 La sección de normalización de espectro 105 normaliza el coeficiente espectral emitido desde la sección de transformación 102 según la envolvente espectral cuantificada emitida desde la sección de cuantificación de norma 104 y emite el coeficiente espectral normalizado a la sección de codificación de vectores de red 108.

25 La sección de ajuste de norma 106 ajusta la envolvente espectral cuantificada emitida desde la sección de cuantificación de norma 104 basándose en la ponderación espectral adaptativa y emite la envolvente espectral cuantificada ajustada a la sección de asignación de bits 107.

La sección de asignación de bits 107 asigna bits disponibles para cada banda en una trama usando la envolvente espectral cuantificada ajustada emitida desde la sección de ajuste de norma 106 y emite los bits asignados a la sección de codificación de vectores de red 108. Más adelante se describirán detalles de la sección de asignación de bits 107.

30 La sección de codificación de vectores de red 108 realiza la codificación de vectores de red en el coeficiente espectral normalizado por la sección de normalización de espectro 105 usando los bits asignados para cada banda en la sección de asignación de bits 107 y emite el vector de codificación de red a la sección de ajuste de nivel de ruido 109 y el multiplexor 110.

La sección de ajuste de nivel de ruido 109 estima el nivel del coeficiente espectral antes de la codificación en la sección de codificación de vectores de red 108 y codifica el nivel estimado. De esta manera se determina un índice de ajuste de nivel de ruido. El índice de ajuste de nivel de ruido se emite al multiplexor 110.

40 El multiplexor 110 multiplexa el indicador de señal transitoria emitido desde la sección de transformación 102, la envolvente espectral cuantificada emitida desde la sección de cuantificación de norma 104, el vector de codificación de red emitido desde la sección de codificación de vectores de red 108 y el índice de ajuste de nivel de ruido emitido desde la sección de ajuste de nivel de ruido 109, y forma un flujo de bits y transmite el flujo de bits a un aparato de descodificación de voz/audio.

La figura 6 es un diagrama de bloques que ilustra una configuración del aparato de descodificación de voz/audio 200 según una realización de la presente invención. Un flujo de bits transmitido desde el aparato de codificación de voz/audio 100 se recibe mediante el aparato de descodificación de voz/audio 200 y se demultiplexa mediante el demultiplexor 201.

45 La sección de descuantificación de norma 202 descuantifica la envolvente espectral cuantificada (es decir, la norma) emitida desde el multiplexor, obtiene una envolvente espectral compuesta por normas de todas las bandas y emite la envolvente espectral obtenida a la sección de ajuste de norma 203.

50 La sección de ajuste de norma 203 ajusta la envolvente espectral emitida desde la sección de descuantificación de norma 202 basándose en la ponderación espectral adaptativa y emite la envolvente espectral ajustada a la sección de asignación de bits 204.

La sección de asignación de bits 204 asigna bits disponibles para cada banda en una trama usando la envolvente

espectral emitida desde la sección de ajuste de norma 203. Es decir, la sección de asignación de bits 204 recalcula la asignación de bits indispensable para descodificar el código de vectores de red del coeficiente espectral normalizado. Los bits asignados se emiten a la sección de descodificación de red 205.

5 La sección de descodificación de red 205 descodifica el vector de codificación de red emitido desde el demultiplexor 201 basándose en una configuración de trama indicada por el indicador de señal transitoria emitido desde el demultiplexor 201 y los bits emitidos desde la sección de asignación de bits 204 y adquiere un coeficiente espectral. El coeficiente espectral se emite al generador de llenado espectral 206 y el sumador 207.

10 El generador de llenado espectral 206 regenera un coeficiente espectral de baja frecuencia al que no se han asignado bits usando un libro de códigos creado basándose en el coeficiente espectral emitido desde la sección de descodificación de red 205. El generador de llenado espectral 206 ajusta el nivel del coeficiente espectral regenerado usando el índice de ajuste de nivel de ruido emitido desde el demultiplexor 201. Además, el generador de llenado espectral 206 regenera el coeficiente espectral no sometido a codificación de alta frecuencia usando un coeficiente espectral codificado de baja frecuencia. El coeficiente espectral de baja frecuencia de nivel ajustado y coeficiente espectral de alta frecuencia regenerado se emiten al sumador 207.

15 El sumador 207 suma el coeficiente espectral emitido desde la sección de descodificación de red 205 y el coeficiente espectral emitido desde el generador de llenado espectral 206, genera un coeficiente espectral normalizado y emite el coeficiente espectral normalizado a la sección de conformación de envolvente 208.

20 La sección de conformación de envolvente 208 aplica la envolvente espectral emitida desde la sección de descuantificación de norma 202 al coeficiente espectral normalizado generado por el sumador 207 y genera un coeficiente espectral de banda completa (correspondiente al espectro descodificado). El coeficiente espectral de banda completa generado se emite a la sección de transformación inversa 209.

25 La sección de transformación inversa 209 aplica una transformación inversa tal como transformada de coseno discreta modificada inversa (IMDCT) al coeficiente espectral de banda completa emitido desde la sección de conformación de envolvente 208, lo transforma en una señal de dominio de tiempo y emite una señal de salida. En este caso, se aplica transformación inversa con resolución de alta frecuencia a un caso de una trama estacionaria y se aplica transformación inversa con resolución de baja frecuencia a un caso de una trama transitoria.

30 A continuación, se describirán los detalles de la sección de asignación de bits 107 usando la figura 7. Obsérvese que la sección de asignación de bits 107 del aparato de codificación de voz/audio 100 es idéntica en cuanto a la configuración a la sección de asignación de bits 204 del aparato de descodificación de voz/audio 200, y por lo tanto sólo se describirá la sección de asignación de bits 107 y la descripción de la sección de asignación de bits 204 se omitirá en el presente documento.

35 La figura 7 es un diagrama de bloques que ilustra una configuración interna de la sección de asignación de bits 107 mostrada en la figura 5. La sección de identificación de banda de frecuencia dominante 301 identifica, basándose en la envolvente espectral cuantificada emitida desde la sección de ajuste de norma 106, una banda de frecuencia dominante que es una subbanda en la que un valor de coeficiente de norma en el espectro tiene un valor máximo local, y emite cada banda de frecuencia dominante identificada a las secciones de determinación de grupo dominante 302-1 a 302N. Además de designar una banda de frecuencia para la que un valor de coeficiente de norma tiene un valor máximo local, los ejemplos del método de determinación de una banda de frecuencia dominante pueden incluir designar una banda de entre todas las subbandas en las que un valor de coeficiente de norma tiene un valor máximo como banda de frecuencia dominante o designar como banda de frecuencia dominante una banda que tiene un valor de coeficiente de norma que supera un umbral predeterminado o un umbral calculado a partir de normas de todas las subbandas.

45 Las secciones de determinación de grupo dominante 302-1 a 302N determinan de manera adaptativa las anchuras de grupo según las características de la señal de entrada centrada en la banda de frecuencia dominante emitida desde la sección de identificación de banda de frecuencia dominante 301. Más específicamente, la anchura de grupo se define como la anchura de un grupo de subbandas centrado en y a ambos lados de la banda de frecuencia dominante hasta subbandas en las que se detiene una pendiente descendente del valor de coeficiente de norma. Las secciones de determinación de grupo dominante 302-1 a 302N determinan bandas de frecuencia incluidas en anchuras de grupo como grupos dominantes y emiten los grupos dominantes determinados a la sección de determinación de grupo no dominante 303. Obsérvese que cuando una banda de frecuencia dominante está ubicada en un borde (extremo de una frecuencia disponible), sólo se incluye en el grupo un lado de la pendiente descendente.

55 La sección de determinación de grupo no dominante 303 determina subbandas continuas emitidas desde las secciones de determinación de grupo dominante 302-1 a 302N distintas de los grupos dominantes como grupos no dominantes sin bandas de frecuencia dominantes. La sección de determinación de grupo no dominante 303 emite los grupos dominantes y los grupos no dominantes a la sección de cálculo de energía de grupo 304 y la sección de cálculo de variación de norma 306.

La sección de cálculo de energía de grupo 304 calcula la energía específica de grupo de los grupos dominantes y

los grupos no dominantes emitidos desde la sección de determinación de grupo no dominante 303 y emite la energía calculada a la sección de cálculo de energía total 305 y la sección de distribución de bits de grupo 308. La energía específica de grupo se calcula mediante la siguiente ecuación 1.

[1]

$$Energía(G(k)) = \sum_{i=1}^M Norma(i) \dots (Ecuación 1)$$

En este caso, k indica un índice de cada grupo, Energía(G(k)) indica la energía del grupo k, i indica un índice de subbanda del grupo k, M indica el número total de subbandas del grupo k y Norma(i) indica un valor de coeficiente de norma de la subbanda i del grupo k.

La sección de cálculo de energía total 305 suma toda la energía específica de grupo emitida desde la sección de cálculo de energía de grupo 304 y calcula la energía total de todos los grupos. La energía total calculada se emite a la sección de distribución de bits de grupo 308. La energía total se calcula mediante la siguiente ecuación 2.

[2]

$$Energía_{total} = \sum_{k=1}^N Energía(G(k)) \dots (Ecuación 2)$$

En este caso, Energía_{total} indica la energía total de todos los grupos, N indica el número total de grupos en un espectro, k indica un índice de cada grupo y Energía(G(k)) indica la energía del grupo k.

La sección de cálculo de variación de norma 306 calcula la variación de norma específica de grupo para los grupos dominantes y los grupos no dominantes emitidos desde la sección de determinación de grupo no dominante 303, y emite la variación de norma calculada a la sección de cálculo de variación de norma total 307 y la sección de distribución de bits de grupo 308. La variación de norma específica de grupo se calcula mediante la siguiente ecuación 3.

[3]

$$Norma_{var}(G(k)) = Norma_{máx}(G(k)) - Norma_{mín}(G(k)) \dots (Ecuación 3)$$

En este caso, k indica un índice de cada grupo, Norma_{var}(G(k)) indica una variación de norma del grupo k, Norma_{máx}(G(k)) indica un valor de coeficiente de norma máximo del grupo k y Norma_{mín}(G(k)) indica un valor de coeficiente de norma mínimo del grupo k.

La sección de cálculo de variación de norma total 307 calcula una variación de norma total de todos los grupos basándose en la variación de norma específica de grupo emitida desde la sección de cálculo de variación de norma 306. La variación de norma total calculada se emite a la sección de distribución de bits de grupo 308. La variación de norma total se calcula mediante la siguiente ecuación 4.

[4]

$$Norma_{vartotal} = \sum_{k=1}^N Norma_{var}(G(k)) \dots (Ecuación 4)$$

En este caso, Norma_{vartotal} indica una variación de norma total de todos los grupos, N indica el número total de grupos en un espectro, k indica un índice de cada grupo, y Norma_{var}(G(k)) indica una variación de norma del grupo k.

La sección de distribución de bits de grupo 308 (correspondiente a una primera sección de asignación de bits) distribuye bits de grupo en grupo basándose en la energía específica de grupo emitida desde la sección de cálculo de energía de grupo 304, la energía total de todos los grupos emitida desde la sección de cálculo de energía total 305, la variación de norma específica de grupo emitida desde la sección de cálculo de variación de norma 306 y la variación de norma total de todos los grupos emitida desde la sección de cálculo de variación de norma total 307, y emite bits distribuidos de grupo en grupo a la sección de distribución de bits de subbanda 309. Los bits distribuidos de grupo en grupo se calculan mediante la siguiente ecuación 5.

[5]

$$Bits(G(k)) = Bits_{total} \times \left(escala1 \times \frac{Energía(G(k))}{Energía_{total}} + (1 - escala1) \times \frac{Norma_{var}(G(k))}{Norma_{var_{total}}} \right) \dots$$

(Ecuación 5)

En este caso, k indica un índice de cada grupo, Bits(G(k)) indica el número de bits distribuidos al grupo k, Bits_{total} indica el número total de bits disponibles, escala1 indica la proporción de bits asignados por energía, Energía(G(k)) indica la energía del grupo k, Energía_{total} indica la energía total de todos los grupos y Norma_{var}(G(k)) indica una variación de norma del grupo k.

Además, escala1 en la ecuación 5 anterior adopta un valor dentro de un intervalo de [0, 1] y ajusta la proporción de bits asignados por energía o variación de norma. Cuanto mayor es el valor de escala1, más bits se asignan por energía, y en un caso extremo, si el valor es 1, todos los bits se asignan por energía. Cuanto menor es el valor de escala1, más bits se asignan por variación de norma, y en un caso extremo, si el valor es 0, todos los bits se asignan por variación de norma.

Distribuyendo bits de grupo en grupo tal como se describió anteriormente, la sección de distribución de bits de grupo 308 puede distribuir más bits a grupos dominantes y distribuir menos bits a grupos no dominantes.

Por tanto, la sección de distribución de bits de grupo 308 puede determinar la importancia perceptiva de cada grupo por energía y variación de norma y potencia más los grupos dominantes. La variación de norma cumple una teoría de enmascaramiento y puede determinar la importancia perceptiva de manera más precisa.

La sección de distribución de bits de subbanda 309 (correspondiente a una segunda sección de asignación de bits) distribuye bits a subbandas en cada grupo basándose en los bits específicos de grupo emitidos desde la sección de distribución de bits de grupo 308 y emite los bits asignados a subbandas específicas de grupo a la sección de codificación de vectores de red 108 como el resultado de la asignación de bits. En este caso, se distribuyen más bits a subbandas perceptivamente importantes y se distribuyen menos bits a subbandas perceptivamente menos importantes. Los bits distribuidos a cada subbanda en un grupo se calculan mediante la siguiente ecuación 6.

[6]

$$Bits_{G(k)sb(i)} = Bits(G(k)) \times \frac{Norma(i)}{Energía(G(k))} \dots \text{(Ecuación 6)}$$

En este caso, Bits_{G(k)sb(i)} indica un bit asignado a la subbanda i del grupo k, i indica un índice de subbanda del grupo k, Bits(G(k)) indica un bit asignado al grupo k, Energía(G(k)) indica la energía del grupo k y Norma(i) indica un valor de coeficiente de norma de la subbanda i del grupo k.

A continuación, se describirá un método de agrupación usando las figuras 8A a 8C. Se supone que se introduce una envolvente espectral cuantificada mostrada en la figura 8A a la sección de identificación de banda de frecuencia de pico 301. La sección de identificación de banda de frecuencia de pico 301 identifica las bandas de frecuencia dominantes 9 y 20 basándose en la envolvente espectral cuantificada introducida (véase la figura 8B).

Las secciones de generación de grupo dominante 302-1 a 302-N determinan subbandas centradas en y a ambos lados de las bandas de frecuencia dominantes 9 y 20 hasta las subbandas en las que se detiene una pendiente descendente del valor de coeficiente de norma como grupo dominante idéntico. En los ejemplos en las figuras 8A a 8C, al igual que para la banda de frecuencia dominante 9, las subbandas 6 a 12 se determinan como grupo dominante (grupo 2), mientras que al igual que para la banda de frecuencia dominante 20, las subbandas 17 a 22 se determinan como grupo dominante (grupo 4) (véase la figura 8C).

La sección de determinación de grupo no dominante 303 determina bandas de frecuencia continuas distintas de los grupos dominantes como grupos no dominantes sin las bandas de frecuencia dominantes. En el ejemplo en las figuras 8A a 8C, las subbandas 1 a 5 (grupo 1), las subbandas 13 a 16 (grupo 3) y las subbandas 23 a 25 (grupo 5) se determinan como grupos no dominantes respectivamente (véase la figura 8C).

Como resultado, las envolventes espectrales cuantificadas se dividen en cinco grupos, es decir, dos grupos dominantes (grupos 2 y 4) y tres grupos no dominantes (grupos 1, 3 y 5).

Usando un método de agrupación de este tipo, es posible determinar de manera adaptativa anchuras de grupo según las características de la señal de entrada. Según este método, el aparato de descodificación de voz/audio también usa coeficientes de norma cuantificados disponibles, y por lo tanto no necesita transmitirse información adicional al aparato de descodificación de voz/audio.

Obsérvese que la sección de cálculo de variación de norma 306 calcula una variación de norma específica de grupo. En los ejemplos en las figuras 8A a 8C, se muestra la variación de norma $Energía_{var}(G(2))$ en el grupo 2 en la figura 9 como referencia.

5 A continuación, se describirá la importancia perceptiva. Un espectro de una señal de voz/audio incluye generalmente una pluralidad de picos (montañas) y valles. Un pico está compuesto por una componente espectral ubicada en una frecuencia dominante de la señal de voz/audio (componente de sonido dominante). El pico es perceptivamente muy importante. La importancia perceptiva del pico puede determinarse mediante una diferencia entre la energía del pico y la energía del valle, es decir, mediante una variación de norma. Teóricamente, cuando un pico tiene una energía suficientemente grande en comparación con bandas de frecuencia vecinas, el pico debe codificarse con un número de bits suficiente, pero si el pico se codifica con un número de bits insuficiente, el ruido de codificación que se mezcla se vuelve destacado, provocando que se deteriore la calidad de sonido. Por otra parte, un valle no está compuesto por ninguna componente de sonido dominante de una señal de voz/audio y no es perceptivamente importante.

15 Según el método de agrupación de banda de frecuencia de la presente realización, una banda de frecuencia dominante corresponde a un pico de un espectro y agrupar bandas de frecuencia significa separar los picos (grupos dominantes que incluyen bandas de frecuencia dominantes) de los valles (grupos no dominantes sin bandas de frecuencia dominantes).

20 La sección de distribución de bits de grupo 308 determina la importancia perceptiva de un pico. Al contrario que la técnica de G.719 en la que se determina la importancia perceptiva sólo mediante la energía, la presente realización determina la importancia perceptiva basándose tanto en la energía como en distribuciones de la norma (energía) y determina los bits que van a distribuirse a cada grupo basándose en la importancia perceptiva determinada.

25 En la sección de distribución de bits de subbanda 309, cuando una variación de norma en un grupo es grande, significa que este grupo es uno de picos, el pico es perceptivamente más importante y debe codificarse de manera precisa un coeficiente de norma que tiene un valor máximo. Por este motivo, se distribuyen más bits a cada subbanda de este pico. Por otra parte, cuando una variación de norma en un grupo es muy pequeña, significa que este grupo es uno de valles, y el valle no es perceptivamente importante y no necesita codificarse de manera precisa. Por este motivo, se distribuyen menos bits a cada subbanda de este grupo.

30 Por tanto, la presente realización identifica una banda de frecuencia dominante en la que un valor de coeficiente de norma en un espectro de una señal de voz/audio de entrada tiene un valor máximo local, agrupa todas las subbandas en grupos dominantes que incluyen una banda de frecuencia dominante y grupos no dominantes que no incluyen ninguna banda de frecuencia dominante, distribuye bits a cada grupo basándose en la energía específica de grupo y las variaciones de norma, y distribuye adicionalmente los bits distribuidos de grupo en grupo a cada subbanda según una razón de una norma con respecto a la energía de cada grupo. De esta manera, es posible asignar más bits a grupos y subbandas perceptivamente importantes y realizar una distribución de bits eficiente.

35 Como resultado, puede mejorarse la calidad de sonido.

Obsérvese que el coeficiente de norma en la presente realización representa la energía de subbanda y también se denomina "envolvente de energía".

40 La divulgación de la solicitud de patente japonesa n.º 2012-272571, presentada el 13 de diciembre de 2012, que incluye la memoria descriptiva, los dibujos y el resumen se incorpora en el presente documento como referencia en su totalidad.

Aplicabilidad industrial

45 El aparato de codificación de voz/audio, el aparato de decodificación de voz/audio, el método de codificación de voz/audio y el método de decodificación de voz/audio según la presente invención pueden aplicarse a un aparato de terminal de comunicación por radio, un aparato de estación base de comunicación por radio, un aparato de terminal de conferencia telefónica, un aparato de terminal de conferencia de vídeo y un aparato de terminal de protocolo de voz por Internet (VoIP) o similares.

Lista de signos de referencia

- 101 Detector transitorio
- 102 Sección de transformación
- 50 103 Sección de estimación de norma
- 104 Sección de cuantificación de norma
- 105 Sección de normalización de espectro
- 106, 203 Sección de ajuste de norma

- 107, 204 Sección de asignación de bits
- 108 Sección de codificación de vectores de red
- 109 Sección de ajuste de nivel de ruido
- 110 Multiplexor
- 5 201 Demultiplexor
- 202 Sección de descuantificación de norma
- 205 Sección de descodificación de red
- 206 Generador de llenado espectral
- 207 Sumador
- 10 208 Sección de conformación de envolvente
- 209 Sección de transformación inversa
- 301 Sección de identificación de banda de frecuencia dominante
- 302-1 a 302-N Sección de determinación de grupo dominante
- 303 Sección de determinación de grupo no dominante
- 15 304 Sección de cálculo de energía de grupo
- 305 Sección de cálculo de energía total
- 306 Sección de cálculo de variación de norma
- 307 Sección de cálculo de variación de norma total
- 308 Sección de distribución de bits de grupo
- 20 309 Sección de distribución de bits de subbanda

REIVINDICACIONES

1. Aparato de codificación de voz/audio que comprende:
 - una sección de transformación que transforma una señal de entrada de un dominio de tiempo a un dominio de frecuencia;
 - 5 una sección de estimación que estima una envolvente de energía que representa un nivel de energía para cada una de una pluralidad de subbandas obtenidas dividiendo un espectro de frecuencia de la señal de entrada;
 - una sección de cuantificación que cuantifica las envolventes de energía;
 - 10 una sección de determinación de grupo que agrupa las envolventes de energía cuantificadas en una pluralidad de grupos;
 - en el que la sección de determinación de grupo determina
 - una banda de frecuencia dominante que es una subbanda en la que una envolvente de energía del espectro de frecuencia tiene un valor máximo local
 - 15 y subbandas a ambos lados de la banda de frecuencia dominante que forman, cada una, una pendiente descendente de una envolvente de energía como grupos dominantes, y determina subbandas mutuamente adyacentes distintas de los grupos dominantes como grupos no dominantes;
 - una primera sección de asignación de bits que asigna bits a la pluralidad de grupos determinados;
 - una segunda sección de asignación de bits que asigna los bits asignados a la pluralidad de grupos determinados a subbandas de grupo en grupo; y
 - 20 una sección de codificación que codifica el espectro de frecuencia usando bits asignados a las subbandas.
2. Aparato de codificación de voz/audio según la reivindicación 1, que comprende además:
 - una sección de cálculo de energía que calcula energía específica de grupo; y
 - una sección de cálculo de distribución que calcula una distribución de envolvente de energía específica de grupo, en el que
 - 25 la primera sección de asignación de bits asigna, basándose en la energía específica de grupo calculada y la distribución de envolvente de energía específica de grupo, más bits a un grupo cuando al menos una de la energía y la distribución de envolvente de energía es mayor y asigna menos bits a un grupo cuando al menos una de la energía y la distribución de envolvente de energía es menor.
3. Aparato de codificación de voz/audio según la reivindicación 1, en el que la segunda sección de asignación de bits asigna más bits a una subbanda que tiene una envolvente de energía mayor y asigna menos bits a una subbanda que tiene una envolvente de energía menor.
- 30 4. Aparato de decodificación de voz/audio que comprende:
 - una sección de descuantificación que descuantifica una envolvente espectral cuantificada;
 - 35 una sección de determinación de grupo que agrupa las envolventes espectrales cuantificadas en una pluralidad de grupos;
 - en el que la sección de determinación de grupo determina
 - una banda de frecuencia dominante que es una subbanda en la que una envolvente de energía del espectro de frecuencia tiene un valor máximo local
 - 40 y subbandas a ambos lados de la banda de frecuencia dominante que forman, cada una, una pendiente descendente de una envolvente de energía como grupos dominantes y determina subbandas mutuamente adyacentes distintas de los grupos dominantes como grupos no dominantes
 - una primera sección de asignación de bits que asigna bits a la pluralidad de grupos determinados;
 - una segunda sección de asignación de bits que asigna los bits asignados a la pluralidad de grupos determinados a subbandas de grupo en grupo;
 - 45 una sección de decodificación que decodifica un espectro de frecuencia de una señal de voz/audio

- usando los bits asignados a las subbandas;
- una sección de conformación de envolvente que aplica la envolvente espectral descuantificada al espectro de frecuencia descodificado y reproduce un espectro descodificado; y
- 5 una sección de transformación inversa que realiza la transformación inversa del espectro descodificado de un dominio de frecuencia a un dominio de tiempo.
5. Aparato de descodificación de voz/audio según la reivindicación 4, que comprende además:
- una sección de cálculo de energía que calcula energía específica de grupo; y
- una sección de cálculo de distribución que calcula una envolvente de energía específica de grupo, en el que
- 10 la primera sección de asignación de bits asigna, basándose en la energía específica de grupo calculada y la distribución de envolvente de energía específica de grupo, más bits a grupos cuando al menos una de la energía y la distribución de envolvente de energía es mayor y asigna menos bits a grupos cuando al menos una de la energía y la distribución de envolvente de energía es menor.
6. Aparato de descodificación de voz/audio según la reivindicación 4, en el que la segunda sección de
- 15 asignación de bits asigna más bits a subbandas que tienen una envolvente de energía mayor y asigna menos bits a subbandas que tienen una envolvente de energía menor.
7. Método de codificación de voz/audio que comprende:
- transformar una señal de entrada de un dominio de tiempo a un dominio de frecuencia;
- estimar una envolvente de energía que representa un nivel de energía para cada una de una pluralidad de subbandas obtenidas dividiendo un espectro de frecuencia de la señal de entrada;
- 20 cuantificar las envolventes de energía;
- agrupar las envolventes de energía cuantificadas en una pluralidad de grupos en el que, en la agrupación, se determina una banda de frecuencia dominante que es una subbanda en la que una envolvente de energía del espectro de frecuencia tiene un valor máximo local y subbandas a ambos lados de la banda de frecuencia dominante que forman, cada una, una pendiente descendente de una envolvente de energía
- 25 como grupos dominantes, y se determinan subbandas mutuamente adyacentes distintas de los grupos dominantes como grupos no dominantes;
- asignar bits a la pluralidad de grupos determinados;
- asignar los bits asignados a la pluralidad de grupos determinados a subbandas de grupo en grupo; y
- codificar el espectro de frecuencia usando bits asignados a las subbandas.
- 30 8. Método de descodificación de voz/audio que comprende:
- descuantificar una envolvente espectral cuantificada;
- agrupar la envolvente espectral cuantificada en una pluralidad de grupos;
- 35 en el que, en la agrupación, se determina una banda de frecuencia dominante que es una subbanda en la que una envolvente de energía del espectro de frecuencia tiene un valor máximo local y subbandas a ambos lados de la banda de frecuencia dominante que forman, cada una, una pendiente descendente de una envolvente de energía como grupos dominantes, y se determinan subbandas mutuamente adyacentes distintas de los grupos dominantes como grupos no dominantes;
- asignar bits a la pluralidad de grupos determinados;
- asignar los bits asignados a la pluralidad de grupos determinados a subbandas de grupo en grupo;
- 40 descodificar un espectro de frecuencia de una señal de voz/audio usando los bits asignados a las subbandas;
- aplicar la envolvente espectral descuantificada al espectro de frecuencia descodificado y reproducir un espectro descodificado; y
- 45 realizar la transformación inversa del espectro descodificado de un dominio de frecuencia a un dominio de tiempo.

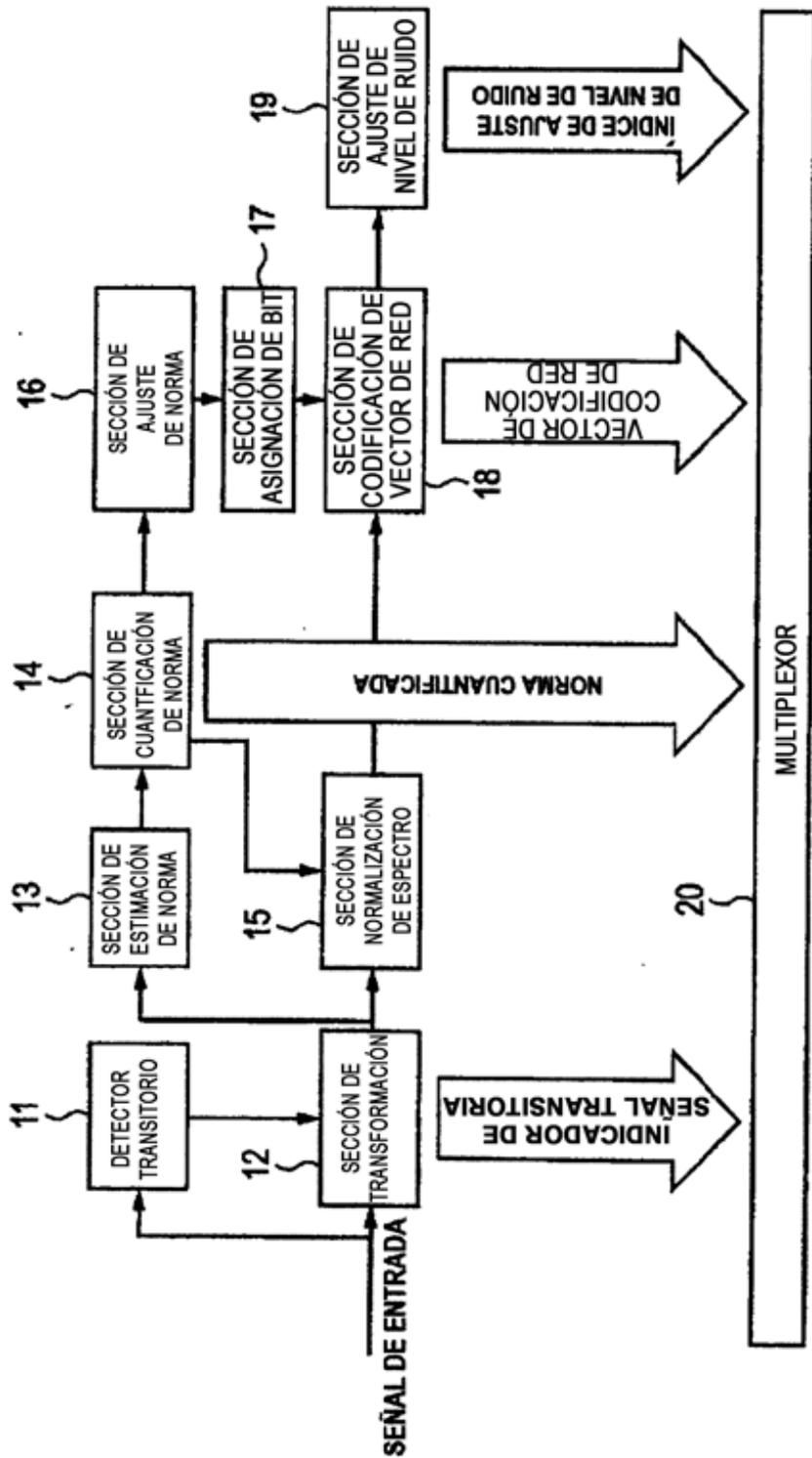


FIG. 1

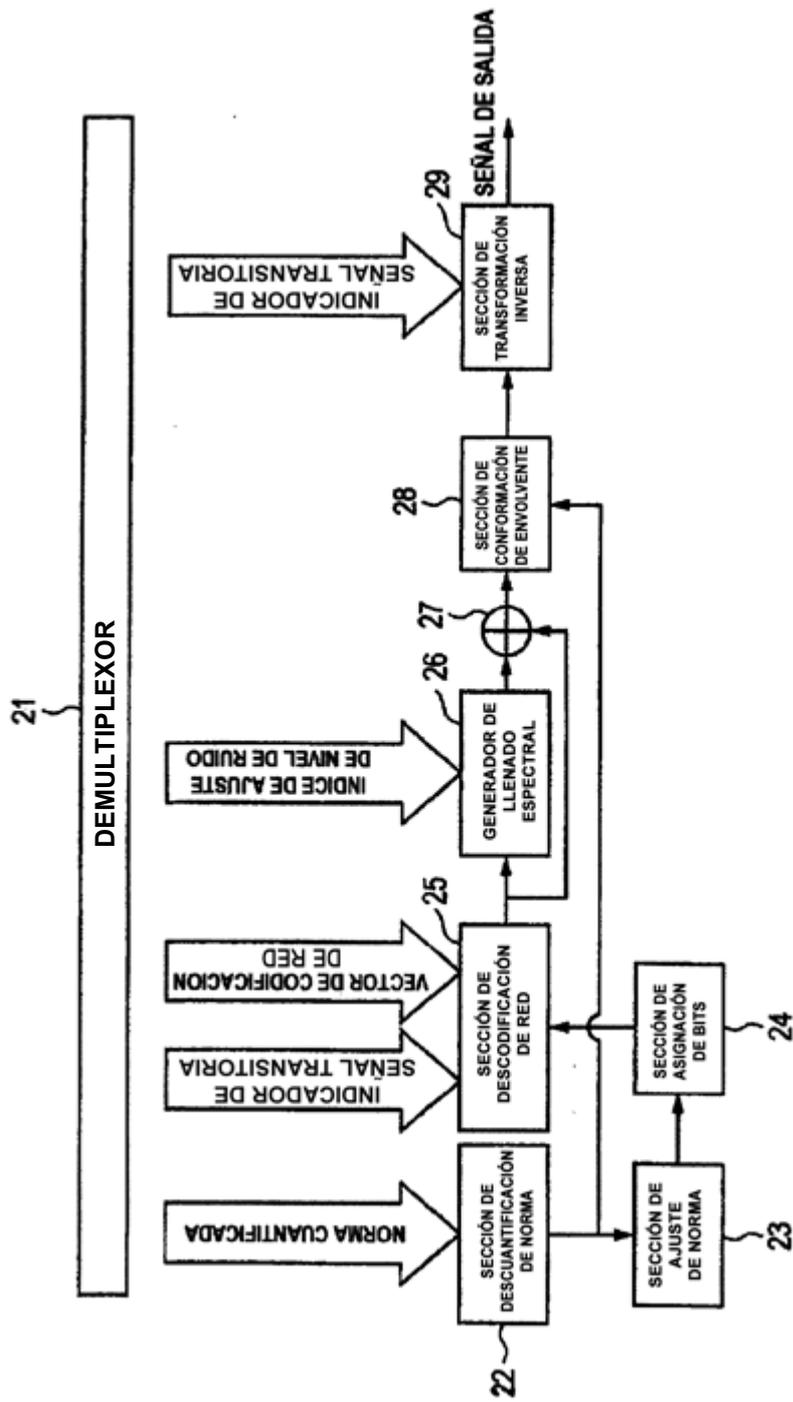


FIG. 2

GRUPO	LONGITUD DEL SUBVECTOR	NUMERO DE SUBVECTORES	NUMERO DE COEFICIENTES	ANCHO DE BANDA (Hz)	INICIO (Hz)	FINAL (Hz)
I	8	16	128	3 200	0	3 200
II	16	8	128	3 200	3 200	6 400
III	24	12	288	7 200	6 400	13 600
IV	32	8	256	6 400	13 600	20 000
TOTAL		44	800	20 000		

FIG. 3

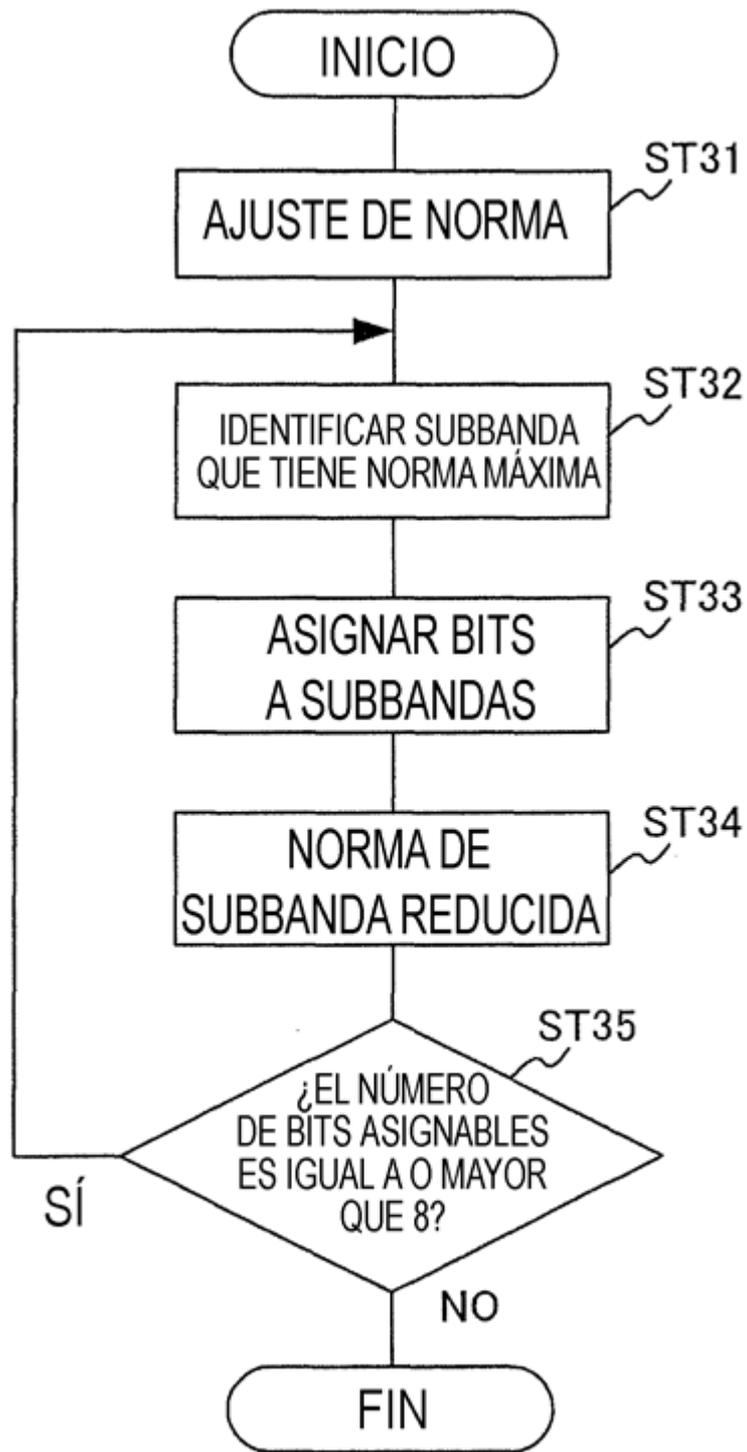


FIG. 4

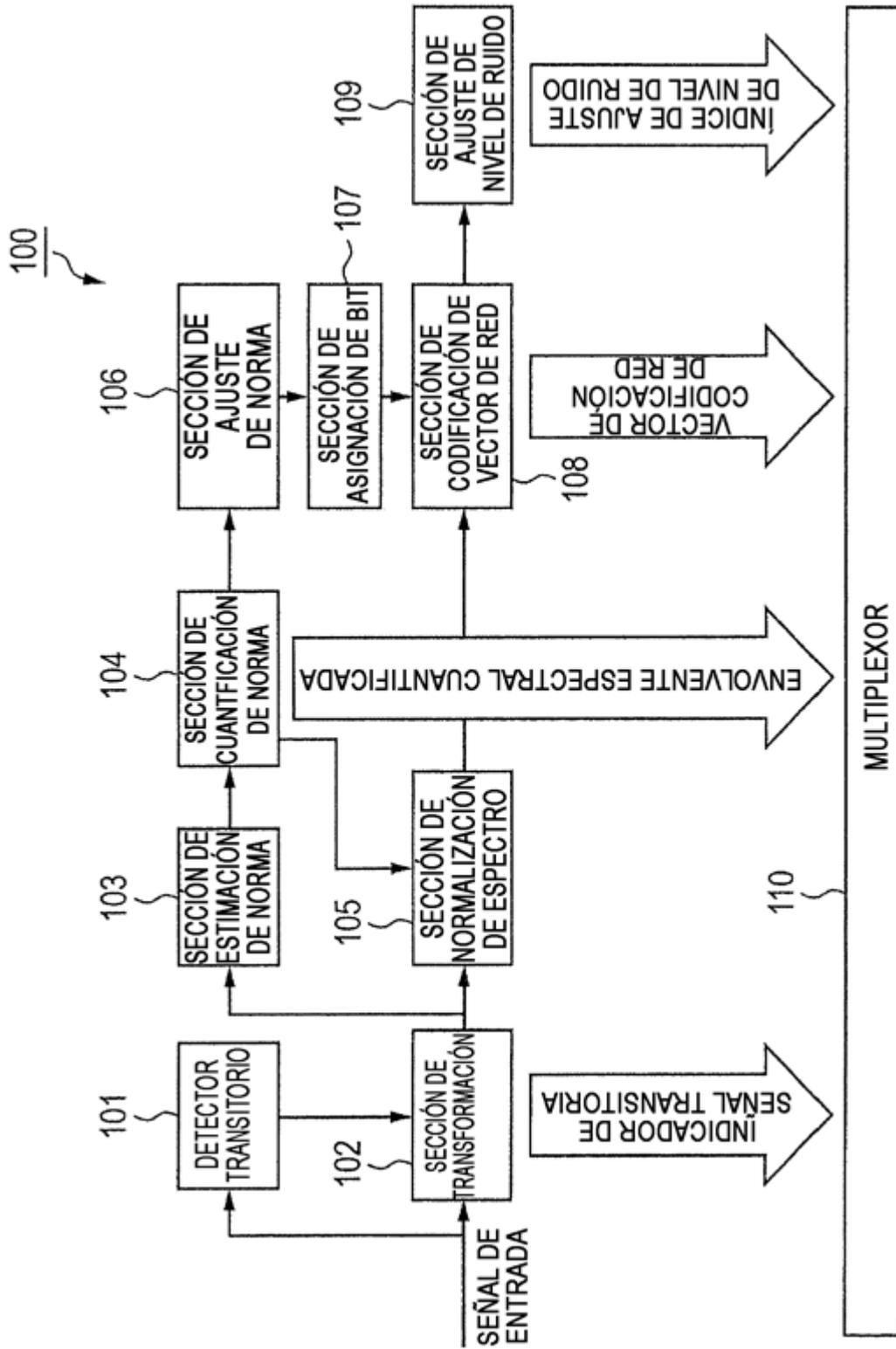


FIG. 5

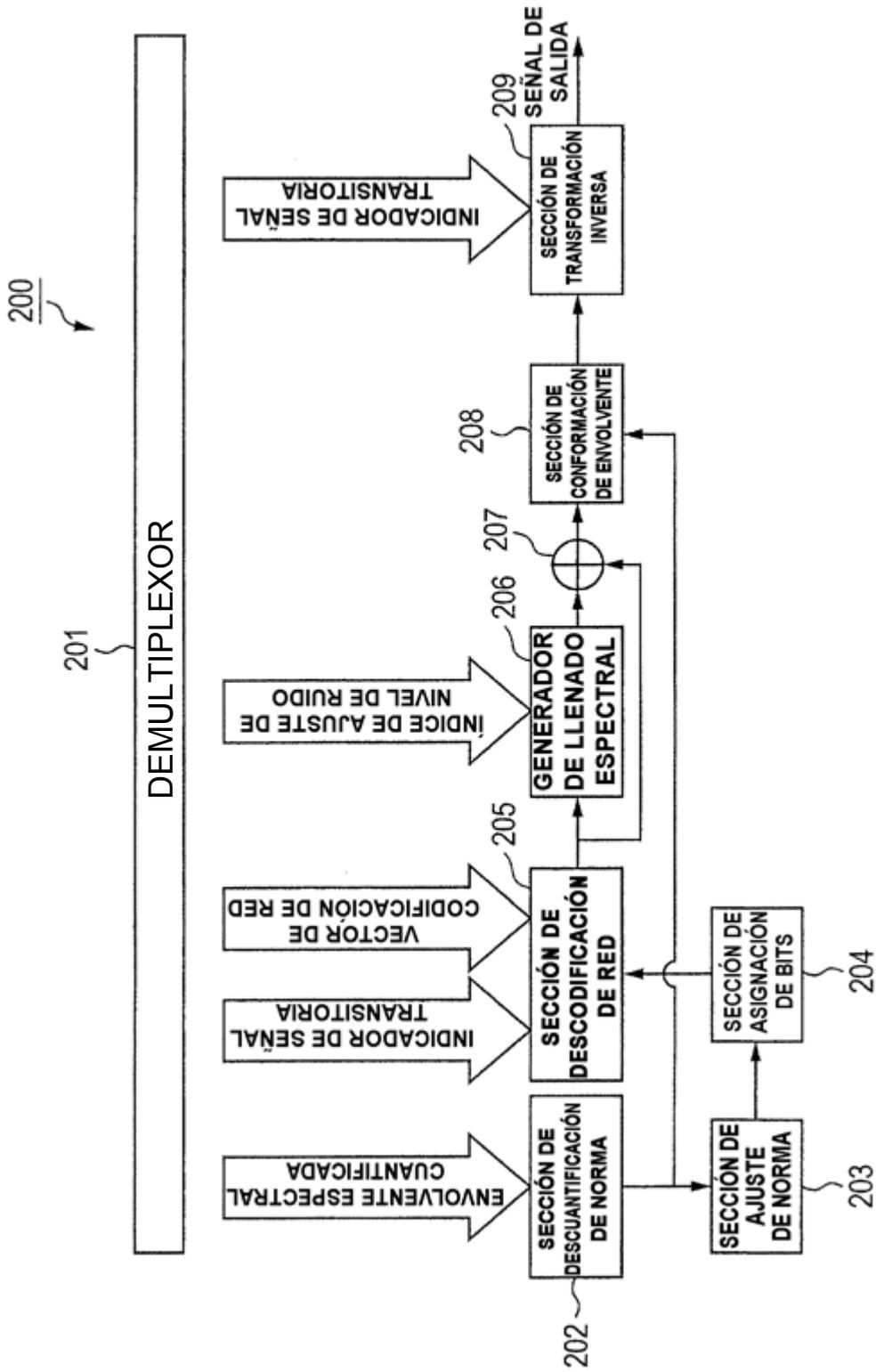


FIG. 6

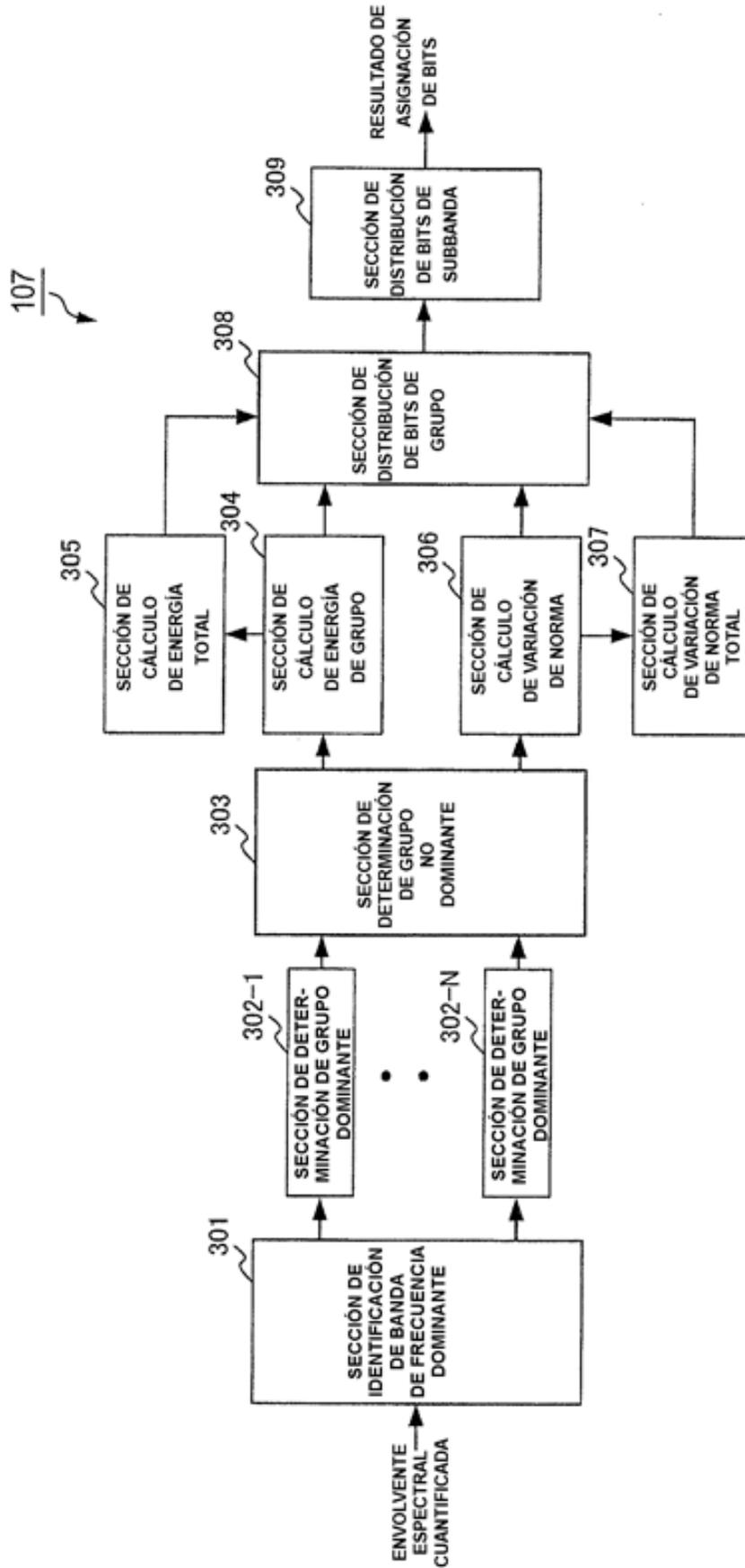


FIG. 7

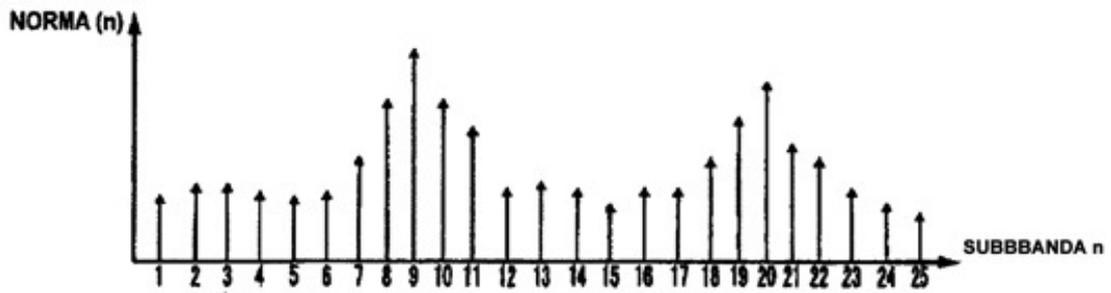


FIG. 8A

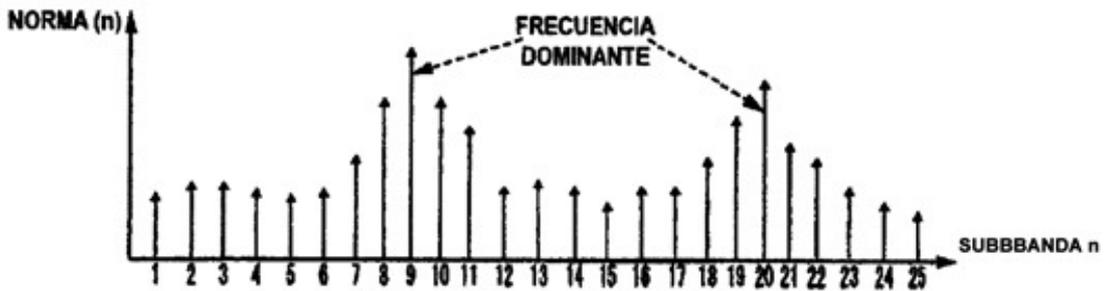


FIG. 8B

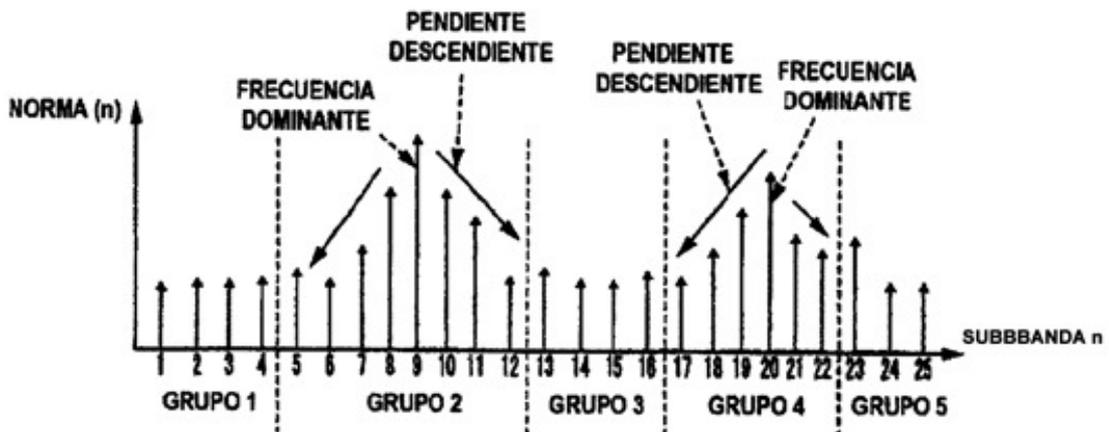


FIG. 8C

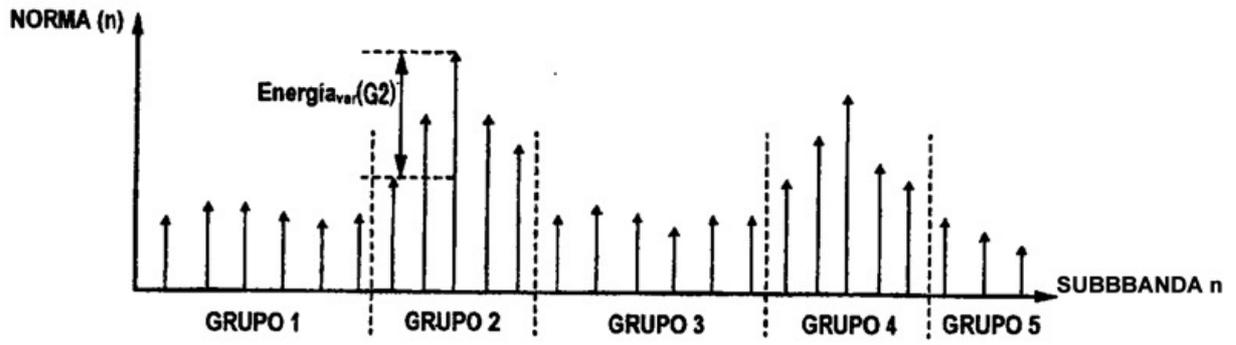


FIG. 9