

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 970 676**

51 Int. Cl.:

G10L 19/035 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.11.2013** **E 18202397 (8)**

97 Fecha y número de publicación de la concesión europea: **16.08.2023** **EP 3457400**

54 Título: **Dispositivo de codificación de audio vocal, dispositivo de decodificación de audio vocal, procedimiento decodificación de audio vocal, y procedimiento de decodificación de audio vocal**

30 Prioridad:

13.12.2012 JP 2012272571

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

30.05.2024

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastr. 27c
80686 München, DE**

72 Inventor/es:

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

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 970 676 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 vocal, dispositivo de decodificación de audio vocal, procedimiento de codificación de audio vocal, y procedimiento de decodificación de audio vocal

Campo técnico

[0001] La presente invención se refiere a un aparato de codificación de voz/audio, un aparato de decodificación de voz/audio, un procedimiento de codificación de voz/audio y un procedimiento de decodificación de voz/audio usando un esquema de codificación por transformada.

Antecedentes de la técnica

[0002] Como esquema capaz de codificar eficientemente una señal de voz o señal de música en una banda completa (FB, del inglés "full band") de 0,02 a 20 kHz, existe una técnica estandarizada en el ITU-T (Sector de estandarización de las telecomunicaciones de la Unión Internacional de las Telecomunicaciones). Esta técnica transforma una señal de entrada en una señal en el dominio de la frecuencia y codifica una banda de hasta 20 kHz (codificación por transformada).

[0003] Aquí, la codificación por transformada es un esquema de codificación que transforma una señal de entrada de un dominio del tiempo a un dominio de la frecuencia usando transformación de tiempo/frecuencia tal como la transformada discreta del coseno (DCT, del inglés "Discrete Cosine Transform") o la transformada discreta del coseno modificada (MDCT, del inglés "Modified Discrete Cosine Transform") para permitir que una señal sea correlacionada en correspondencia precisa con características auditivas.

[0004] En la codificación por transformada, un coeficiente espectral es dividido 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 sonora global.

[0005] Con el fin de alcanzar este objeto, se están llevando a cabo estudios sobre esquemas de asignación de bits eficiente y, por ejemplo, se conoce una técnica descrita en la bibliografía no relacionada con patentes (denominada en lo sucesivo "NPL", del inglés "Non-Patent Literature") 1. En lo sucesivo, el esquema de asignación de bits descrito en la bibliografía de patentes (denominada en lo sucesivo "PTL", del inglés "Patent Literature") 1 se describirá usando la FIG. 1 y la FIG. 2.

[0006] La FIG. 1 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de voz/audio descrito en la PTL 1. Una señal de entrada muestreada a 48 kHz es introducida en un detector de transitorios 11 y la sección de transformación 12 del aparato de codificación de voz/audio.

[0007] El detector de transitorios 11 detecta, a partir de la señal de entrada, o bien una trama transitoria que corresponde a un borde anterior o un borde final de voz o bien una trama estacionaria que corresponde a una sección de voz distinta de esa, 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 de transitorios 11 es una trama transitoria o una trama estacionaria, y adquiere un coeficiente espectral (o coeficiente de transformada).

[0008] La sección de estimación de norma 13 divide el coeficiente espectral obtenido en la sección de transformación 12 en bandas de diferentes anchuras de banda. La sección de estimación de norma 13 estima una norma (o energía) de cada banda dividida.

[0009] La sección de cuantificación de norma 14 determina una envolvente espectral constituida 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.

[0010] 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.

[0011] La sección de ajuste de norma 16 ajusta la norma cuantificada por la sección de cuantificación de norma 14 basándose en ponderación espectral adaptativa.

[0012] La sección de asignación de bits 17 asigna los bits disponibles para cada banda en una trama usando la norma de cuantificación ajustada por la sección de ajuste de norma 16.

[0013] La sección de codificación de vector retícula-vector 18 realiza codificación de retícula-vector sobre el coeficiente espectral normalizado por la sección de normalización de espectro 15 usando los bits asignados para cada banda por la sección de asignación de bits 17.

5 **[0014]** La sección de ajuste de nivel de ruido 19 estima el nivel del coeficiente espectral antes de la codificación en la sección de codificación de retícula-vector 18 y codifica el nivel estimado. De esta manera se obtiene un índice de ajuste de nivel de ruido.

10 **[0015]** El multiplexor 20 multiplexa una configuración de trama de la señal de entrada adquirida por la sección de transformación 12, es decir, una bandera 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 retícula obtenido por la sección de codificación de retícula-vector 18 y el índice de ajuste de nivel de ruido obtenido por la sección de ajuste de nivel de ruido 19, y forma un tren de bits y transmite el tren de bits a un aparato de decodificación de voz/audio.

15 **[0016]** La FIG. 2 es un diagrama de bloques que ilustra una configuración del aparato de decodificación de voz/audio descrito en PTL 1. El aparato de decodificación de voz/audio recibe el tren de bits transmitido desde el aparato de codificación de voz/audio y el demultiplexor 21 demultiplexa el tren de bits.

20 **[0017]** La sección de decuantificación de norma 22 decuantifica la norma cuantificada, adquiere una envolvente espectral constituida por normas de todas las bandas, y la sección de ajuste de norma 23 ajusta la norma decuantificada por la sección de decuantificación de norma 22 basándose en ponderación espectral adaptativa.

25 **[0018]** La sección de asignación de bits 24 asigna los 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 decodificar el código de retícula-vector del coeficiente espectral normalizado.

30 **[0019]** La sección de decodificación de retícula 25 decodifica una bandera de señal transitoria, decodifica el vector de codificación de retícula basándose en una configuración de trama indicada por la bandera de señal transitoria decodificada y los bits asignados por la sección de asignación de bits 24 y adquiere un coeficiente espectral.

35 **[0020]** El generador de relleno espectral 26 regenera un coeficiente espectral de baja frecuencia al que no se ha asignado ningún bit usando un libro de códigos creado basándose en el coeficiente espectral decodificado por la sección de decodificación de retícula 25. El generador de relleno espectral 26 ajusta el nivel del coeficiente espectral regenerado usando un índice de ajuste de nivel de ruido. Además, el generador de relleno espectral 26 regenera un coeficiente espectral no codificado de alta frecuencia usando un coeficiente espectral codificado de baja frecuencia.

40 **[0021]** El sumador 27 suma el coeficiente espectral decodificado y el coeficiente espectral regenerado, y genera un coeficiente espectral normalizado.

45 **[0022]** La sección de modelado de envolvente 28 aplica la envolvente espectral decuantificada por la sección de decuantificación de norma 22 al coeficiente espectral normalizado generado por el sumador 27 y genera un coeficiente espectral de banda completa.

50 **[0023]** La sección de transformación inversa 29 aplica transformada inversa tal como la transformada discreta del coseno modificada inversa (IMDCT, del inglés "inverse modified discrete cosine transform") al coeficiente espectral de banda completa generado por la sección de modelado de envolvente 28 para transformarlo en una señal en el dominio del tiempo. Aquí, la transformada inversa con resolución de alta frecuencia se aplica a un caso con una trama estacionaria y la transformada inversa con resolución de baja frecuencia se aplica a un caso con una trama transitoria.

55 **[0024]** En G.719, los coeficientes espectrales son divididos en grupos de espectros. Cada grupo de espectros es dividido en bandas de subvectores de la misma longitud como se muestra en la FIG. 3. Los subvectores son de diferente longitud de un grupo a otro y esta longitud aumenta a medida que aumenta la frecuencia. Respecto a la resolución de la transformada, se usa resolución de frecuencia más alta para bajas frecuencias, mientras que se usa resolución de frecuencia más baja para altas frecuencias. Como se describe en G.719, la agrupación permite un uso eficiente del presupuesto de bits disponible durante la codificación.

60 **[0025]** En G.719, el esquema de asignación de bits es idéntico en un aparato de codificación y un aparato de decodificación. Aquí, el esquema de asignación de bits se describirá usando la FIG. 4.

65 **[0026]** Como se muestra en la FIG. 4, en la etapa (abreviada en lo sucesivo como "ST") 31, las normas cuantificadas son ajustadas antes de la asignación de bits para ajustar la ponderación psicoacústica y los efectos de enmascaramiento.

[0027] En ST32, las subbandas que tienen una norma máxima son identificadas de entre todas las subbandas y en ST33, es asignado un bit a cada coeficiente espectral para las subbandas que tienen la norma máxima. Es decir, son asignados tantos bits como coeficientes espectrales.

[0028] En ST34, las normas son reducidas según los bits asignados, y en ST35, se determina si el número restante de bits asignables es 8 o más. Cuando el número restante de bits asignables es 8 o más, el flujo vuelve a ST32 y cuando el número restante de bits asignables es inferior a 8, se termina el procedimiento de asignación de bits.

[0029] Así, en el esquema de asignación de bits, los bits disponibles dentro de una trama son asignados entre subbandas usando las normas de cuantificación ajustadas. Los coeficientes espectrales normalizados son codificados mediante codificación de retícula-vector usando los bits asignados a cada subbanda.

Lista de menciones

Bibliografía de patentes

[0030] NPL 1

ITU-T Recomendación G.719, "Codificación de audio en la banda completa y de baja complejidad para aplicaciones conversacionales de alta calidad", ITU-T, 2009.

Resumen de la invención

Problema técnico

[0031] Sin embargo, el esquema de asignación de bits anterior no tiene en consideración las características de la señal de entrada al agrupar las bandas espectrales, y por lo tanto tiene un problema porque no es posible la asignación de bits eficiente y no puede esperarse mejora adicional de la calidad de sonido.

[0032] El documento US 2012/004918 A1 describe un códec de audio escalable para un dispositivo de procesamiento que determina primera y segunda asignaciones de bits para cada trama de audio de entrada. Se asignan primeros bits para una primera banda de frecuencia y se asignan segundos bits para una segunda banda de frecuencia. Las asignaciones se hacen trama por trama basándose en la relación de energía entre las dos bandas. Para cada trama, la transformada del códec codifica ambas bandas de frecuencia en dos conjuntos de coeficientes de transformada, que a continuación se empaquetan basándose en las asignaciones de bits. Los paquetes se transmiten a continuación con el dispositivo de procesamiento. Además, las regiones de frecuencia de los coeficientes de transformada pueden disponerse en orden de importancia determinado por los niveles de energía y el modelado perceptivo. Si se produce eliminación de bits, el decodificador en un dispositivo de recepción puede producir audio de calidad adecuada dado que se han asignado bits entre las bandas y las regiones de los coeficientes de transformada han sido ordenadas por importancia.

[0033] El documento EP 0 259 553 A2 describe codificadores de voz de subbanda que dependen de la asignación de la capacidad de bits disponible de un medio de transmisión para proporcionar codificación de voz de alta calidad para la transmisión digital son bien conocidos. Se utilizan una o más tablas de asignación de bits para distribuir dinámicamente la anchura de banda de capacidad de bits del canal entre las bandas de frecuencia según la calidad de salida deseada de la voz en lugar de por medio de algoritmos complejos o técnicas de simulación. Se proporcionan múltiples tablas de asignación de bits para permitir el intercambio de diversos niveles de calidad a medida que aumentan las demandas de velocidad binaria en el sistema de transmisión. La técnica se utiliza en un solo codificador para lograr una velocidad binaria mínima para un nivel dado deseado de calidad subjetiva en la salida de voz o puede utilizarse en un recurso binario compartido para mantener una degradación de calidad igual y mínima para todos los usuarios.

[0034] La publicación de normalización "Low Complexity coding at 24 and 32 kbits/s for hands-free operation in systems with low frame loss", TD 47 (WP316), ITU-T Draft Study, Periodo 2005-2008, vol. 10/16, 8 de abril de 2005, páginas 1-37 describe un texto completo para el códec G.722.1 revisado.

[0035] Un objeto de la presente invención es proporcionar un aparato de codificación de voz o audio, un aparato de decodificación de voz o audio, un procedimiento de codificación de voz o audio y un procedimiento de decodificación de voz o audio capaz de producir asignación de bits eficiente y mejorar la calidad del sonido.

[0036] Este objeto se logra mediante un aparato de codificación de voz o audio de la reivindicación 1, un aparato de decodificación de voz o audio de la reivindicación 18, un procedimiento de codificación de voz o audio de la reivindicación 21, o un procedimiento de decodificación de voz o audio de la reivindicación 22.

Efectos ventajosos de la invención

[0037] Según la presente invención, es posible producir asignación de bits eficiente y mejorar la calidad del sonido.

Breve descripción de los dibujos

[0038]

- 10 La FIG. 1 es un diagrama de bloques que ilustra una configuración de un aparato de codificación de voz/audio descrito en PTL 1;
La FIG. 2 es un diagrama de bloques que ilustra una configuración de un aparato de decodificación de voz/audio descrito en PTL 1;
15 La FIG. 3 es un diagrama que ilustra la agrupación de coeficientes espectrales en un modo estacionario descrito en PTL 1;
La FIG. 4 es un diagrama de flujo que ilustra un esquema de asignación de bits descrito en PTL 1;
La FIG. 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 FIG. 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;
20 La FIG. 7 es un diagrama de bloques que ilustra una configuración interna de la sección de asignación de bits mostrada en la FIG. 5;
Las FIGS. 8A a 8C son diagramas proporcionados para describir un procedimiento de agrupación según una realización de la presente invención; y
25 La FIG. 9 es un diagrama que ilustra una varianza de norma.

Descripción de las realizaciones

- [0039]** En lo sucesivo, las realizaciones de la presente invención se describirán en detalle con referencia a los dibujos adjuntos.

(Realización)

- [0040]** La FIG. 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 es introducida en el detector de transitorios 101 y la sección de transformación 102 del aparato de codificación de voz/audio 100.

- [0041]** El detector de transitorios 101 detecta, de una señal de entrada, o bien una trama transitoria que corresponde a un borde anterior o un borde final de voz o bien una trama estacionaria que corresponde a una sección de voz distinta de esa, y produce como salida el resultado de detección hacia 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 producido como salida desde el detector de transitorios 101 es una trama transitoria o una trama estacionaria, y adquiere un coeficiente espectral (o coeficiente de transformada) y produce como salida el coeficiente espectral hacia la sección de estimación de norma 103 y la sección de normalización de espectro 105. La sección de transformación 102 produce como salida una configuración de trama que es el resultado de detección producido como salida desde el detector de transitorios 101, es decir, una bandera de señal transitoria que indica si la trama es una trama estacionaria o una trama transitoria al multiplexor 10.

- [0042]** La sección de estimación de norma 103 divide el coeficiente espectral producido como salida desde la sección de transformación 102 en bandas de diferentes anchuras de banda y estima una norma (o energía) de cada banda dividida. La sección de estimación de norma 103 produce como salida la norma estimada de cada banda hacia la sección de cuantificación de norma 104.

- [0043]** La sección de cuantificación de norma 104 determina una envolvente espectral constituida por normas de todas las bandas basándose en normas de bandas respectivas producidas como salida desde la sección de estimación de norma 103, cuantifica la envolvente espectral determinada y produce como salida la envolvente espectral cuantificada hacia la sección de normalización de espectro 105 y la sección de ajuste de norma 106.

- [0044]** La sección de normalización de espectro 105 normaliza el coeficiente espectral producido como salida desde la sección de transformación 102 según la envolvente espectral cuantificada producida como salida desde la sección de cuantificación de norma 104 y produce como salida el coeficiente espectral normalizado hacia la sección de codificación de retícula-vector 108.

[0045] La sección de ajuste de norma 106 ajusta la envolvente espectral cuantificada producida como salida desde la sección de cuantificación de norma 104 basándose en la ponderación espectral adaptativa y produce como salida la envolvente espectral cuantificada ajustada hacia la sección de asignación de bits 107.

5 **[0046]** La sección de asignación de bits 107 asigna los bits disponibles para cada banda en una trama usando la envolvente espectral cuantificada ajustada producida como salida desde la sección de ajuste de norma 106 y produce como salida los bits asignados hacia la sección de codificación de retícula-vector 108. Los detalles de la sección de asignación de bits 107 se describirán más adelante.

10 **[0047]** La sección de codificación de retícula-vector 108 realiza codificación de retícula-vector sobre 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 produce como salida el vector de codificación de retícula hacia la sección de ajuste de nivel de ruido 109 y el multiplexor 110.

15 **[0048]** 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 retícula-vector 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 es producido como salida hacia el multiplexor 110.

20 **[0049]** El multiplexor 110 multiplexa la bandera de señal transitoria producida como salida desde la sección de transformación 102, la envolvente espectral cuantificada producida como salida desde la sección de cuantificación de norma 104, el vector de codificación de retícula producido como salida desde la sección de codificación de retícula-vector 108 y el índice de ajuste de nivel de ruido producido como salida desde la sección de ajuste de nivel 109, y forma un tren de bits y transmite el tren de bits a un aparato de decodificación de voz/audio.

25 **[0050]** La FIG. 6 es un diagrama de bloques que ilustra una configuración del aparato de decodificación de voz/audio 200 según una realización de la presente invención. Un tren de bits transmitido desde el aparato de codificación de voz/audio 100 es recibido por el aparato de decodificación de voz/audio 200 y demultiplexado por el demultiplexor 201.

30 **[0051]** La sección de decuantificación de norma 202 decuantifica la envolvente espectral cuantificada (es decir, la norma) producida como salida desde el multiplexor, obtiene una envolvente espectral constituida por normas de todas las bandas y produce como salida la envolvente espectral obtenida hacia la sección de ajuste de norma 203.

35 **[0052]** La sección de ajuste de norma 203 ajusta la envolvente espectral producida como salida desde la sección de decuantificación de norma 202 basándose en la ponderación espectral adaptativa y produce como salida la envolvente espectral ajustada hacia la sección de asignación de bits 204.

40 **[0053]** La sección de asignación de bits 204 asigna los bits disponibles para cada banda en una trama usando la envolvente espectral producida como salida desde la sección de ajuste de norma 203. Es decir, la sección de asignación de bits 204 recalcula la asignación de bits indispensable para decodificar el código de retícula-vector del coeficiente espectral normalizado. Los bits asignados son producidos como salida hacia la sección de decodificación de retícula 205.

45 **[0054]** La sección de decodificación de retícula 205 decodifica el vector de codificación de retícula producido como salida desde el demultiplexor 201 basándose en una configuración de trama indicada por la bandera de señal transitoria producida como salida desde el demultiplexor 201 y los bits producidos como salida desde la sección de asignación de bits 204 y adquiere un coeficiente espectral. El coeficiente espectral es producido como salida hacia el generador de relleno espectral 206 y el sumador 207.

50 **[0055]** El generador de relleno espectral 206 regenera un coeficiente espectral de baja frecuencia al que no se ha asignado ningún bit usando un libro de códigos creado basándose en el coeficiente espectral producido como salida desde la sección de decodificación de retícula 205. El generador de relleno espectral 206 ajusta el nivel del coeficiente espectral regenerado usando un índice de ajuste de nivel de ruido producido como salida desde el demultiplexor 201. Además, el generador de relleno 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 con nivel ajustado y el coeficiente espectral de alta frecuencia regenerado son producidos como salida hacia el sumador 207.

60 **[0056]** El sumador 207 suma el coeficiente espectral producido como salida desde la sección de decodificación de retícula 205 y el coeficiente espectral producido como salida desde el generador de relleno espectral 206, genera un coeficiente espectral normalizado y produce como salida el coeficiente espectral normalizado hacia la sección de modelado de envolvente 208.

65

[0057] La sección de modelado de envolvente 208 aplica la envolvente espectral producida como salida desde la sección de decuantificación de norma 202 al coeficiente espectral normalizado generado por el sumador 207 y genera un coeficiente espectral de banda completa (que corresponde al espectro decodificado). El coeficiente espectral de banda completa generado es producido como salida hacia la sección de transformación inversa 209.

[0058] La sección de transformación inversa 209 aplica transformada inversa tal como la transformada discreta del coseno modificada inversa (IMDCT) al coeficiente espectral de banda completa producido como salida desde la sección de modelado de envolvente 208, lo transforma en una señal en el dominio del tiempo y produce como salida una señal de salida. Aquí, la transformada inversa con resolución de alta frecuencia se aplica a un caso de una trama estacionaria y la transformada inversa con resolución de baja frecuencia se aplica a un caso de una trama transitoria.

[0059] A continuación, los detalles de la sección de asignación de bits 107 se describirán usando la FIG. 7. Obsérvese que la sección de asignación de bits 107 del aparato de codificación de voz/audio 100 es de configuración idéntica a la sección de asignación de bits 204 del aparato de decodificación de voz/audio 200, y por lo tanto sólo se describirá la sección de asignación de bits 107 y aquí se omitirá la descripción de la sección de asignación de bits 204.

[0060] La FIG. 7 es un diagrama de bloques que ilustra una realización inventiva de una configuración interna de la sección de asignación de bits 107 mostrada en la FIG. 5. La sección de identificación de banda de frecuencia dominante 301 identifica, basándose en la envolvente espectral cuantificada producida como salida desde la sección de ajuste de norma 106, una banda de frecuencia dominante que es una subbanda en la cual un valor de coeficiente de norma en el espectro tiene un valor máximo local, y produce como salida cada banda de frecuencia dominante identificada hacia las secciones de determinación de grupos dominantes 301-1 a 302N. Además de designar una banda de frecuencia dominante para la cual un valor de coeficiente de norma tiene un valor máximo local, ejemplos del procedimiento de determinación de una banda de frecuencia dominante pueden incluir designar una banda de entre todas las subbandas en la que un valor de coeficiente de norma tiene un valor máximo como una banda de frecuencia dominante o designar como una banda de frecuencia dominante una banda que tiene un valor de coeficiente de norma que excede de un umbral predeterminado o un umbral calculado a partir de normas de todas las subbandas.

[0061] Las secciones de determinación de grupos dominantes 302-1 a 302N determinan adaptativamente anchuras de grupo según las características de la señal de entrada centradas en la banda de frecuencia dominante producida como salida 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 centradas en y en ambos lados de la banda de frecuencia dominante hasta las subbandas donde se detiene una pendiente descendiente del valor de coeficiente de norma. Las secciones de determinación de grupos dominantes 302-1 a 302N determinan bandas de frecuencia incluidas en anchuras de grupo como grupos dominantes y produce como salida los grupos dominantes determinados hacia la sección de determinación de grupos no dominantes 303. Obsérvese que cuando una banda de frecuencia dominante está situada en un borde (extremo de una frecuencia disponible), sólo un lado de la pendiente descendente está incluido en el grupo.

[0062] La sección de determinación de grupos no dominantes 303 determina subbandas continuas producidas como salida desde las secciones de determinación de grupos dominantes 302-1 a 302N distintas de los grupos dominantes como grupos no dominantes sin bandas de frecuencia dominante. La sección de determinación de grupos no dominantes 303 produce como salida los grupos dominantes y los grupos no dominantes hacia la sección de cálculo de energía de grupo 304 y la sección de cálculo de varianza de norma 306.

[0063] La sección de cálculo de energía de grupo 304 calcula la energía específica de grupo de los grupos dominantes y los grupos no dominantes producidos como salida desde la sección de determinación de grupos no dominantes 303 y produce como salida la energía calculada hacia 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 por la siguiente ecuación 1.

[1]

$$Energy(G(k)) = \sum_{i=1}^M Norm(i) \quad \dots \quad \text{(Ecuación 1)}$$

[0064] Aquí, k indica un índice de cada grupo, Energy(G(k)) indica la energía del grupo k, i indica un índice de subbanda del grupo 2, M indica el número total de subbandas del grupo k y Norm (i) indica un valor de coeficiente de norma de la subbanda i del grupo n.

[0065] La sección de cálculo de energía total 305 suma toda la energía específica de grupo producida como salida desde la sección de cálculo de energía de grupo y calcula la energía total de todos los grupos. La energía total calculada es producida como salida hacia la sección de distribución de bits de grupo 308. La energía total se calcula por la siguiente ecuación 2.

5 [2]

$$Energy_{total} = \sum_{k=1}^N Energy(G(k)) \quad \dots \quad \text{(Ecuación 2)}$$

[0066] Aquí, $Energy_{total}$ 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 $Energy(G(k))$ indica la energía del grupo k .

[0067] La sección de cálculo de varianza de norma 306 calcula la varianza de norma específica de grupo para los grupos dominantes y los grupos no dominantes producidos como salida desde la sección de determinación de grupos no dominantes 303, y produce como salida la varianza de norma calculada hacia la sección de cálculo de varianza de norma 307 y la sección de distribución de bits de grupo 308. La varianza de norma específica de grupo se calcula por la siguiente ecuación 3.

[3]

$$Norm_{var}(G(k)) = Norm_{max}(G(k)) - Norm_{min}(G(k)) \quad \dots \quad \text{(Ecuación 3)}$$

20

[0068] Aquí, k indica un índice de cada grupo, $Norm_{var}(G(k))$ indica una varianza de norma del grupo k , $Norm_{max}(G(k))$ indica un valor de coeficiente de norma máximo del grupo k , y $Norm_{min}(G(k))$ indica un valor de coeficiente de norma mínimo del grupo k .

[0069] La sección de cálculo de varianza de norma 307 calcula una varianza de norma total de todos los grupos basándose en la varianza de norma específica de grupo producida como salida desde la sección de cálculo de varianza de norma 306. La varianza de norma total calculada es producida como salida hacia la sección de distribución de bits de grupo 308. La varianza de norma total se calcula por la siguiente ecuación 4.

[4]

30

$$Norm_{var_{total}} = \sum_{k=1}^N Norm_{var}(G(k)) \quad \dots \quad \text{(Ecuación 4)}$$

[0070] Aquí, $Norm_{var_{total}}$ indica una varianza 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 $Norm_{var}(G(k))$ indica una varianza de norma del grupo k .

35

[0071] La sección de distribución de bits de grupo 308 (que corresponde a una primera sección de asignación de bits) distribuye los bits grupo por grupo basándose en la energía específica de grupo producida como salida desde la sección de cálculo de energía de grupo 304, la energía total de todos los grupos producida como salida desde la sección de cálculo de energía total 305, la varianza de norma específica de grupo producida como salida desde la sección de cálculo de varianza de norma 306 y la varianza de norma total de todos los grupos producida como salida desde la sección de cálculo de varianza de norma total 307, y produce como salida los bits distribuidos grupo por grupo hacia la sección de distribución de bits de subbanda 309. Los bits distribuidos grupo por grupo se calculan por la siguiente ecuación 5.

[5]

45

$$Bits(G(k)) = Bits_{total} \times \left(scale1 \times \frac{Energy(G(k))}{Energy_{total}} + (1 - scale1) \times \frac{Norm_{var}(G(k))}{Norm_{var_{total}}} \right) \quad \text{(Ecuación 5)}$$

[0072] Aquí, 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, $scale1$ indica la relación de bits asignados por energía, $Energy(G(k))$ indica la energía del grupo k , $Energy_{total}$ indica la energía total de todos los grupos, y $Norm_{var}(G(k))$ indica una varianza de norma del grupo k .

50

[0073] Además, $scale1$ en la ecuación 5 anterior toma un valor dentro de un intervalo de $[0, 1]$ y ajusta la relación de bits asignados por energía o varianza de norma. Cuanto mayor es el valor de $scale1$, más bits son asignados por energía, y en un caso extremo, si el valor es 1, todos los bits son asignados por energía. Cuanto menor es el valor de $scale1$, más bits son asignados por varianza de norma, y en un caso extremo, si el valor es 0, todos los bits son asignados por varianza de norma.

55

[0074] Distribuyendo los bits grupo por grupo como se describe anteriormente, la sección de distribución de bits de grupo 308 puede distribuir más bits a los grupos dominantes y distribuir menos bits a los grupos no dominantes.

- 5 **[0075]** Así, la sección de distribución de bits de grupo 308 puede determinar la importancia perceptiva de cada grupo por energía y varianza de norma y aumentar más los grupos dominantes. La varianza de norma equivale a una teoría de enmascaramiento y puede determinar la importancia perceptiva con más exactitud.

- 10 **[0076]** La sección de distribución de bits de subbanda 309 (que corresponde a una segunda sección de asignación de bits) distribuye los bits a las subbandas en cada grupo basándose en bits específicos de grupo producidos como salida desde la sección de distribución de bits de grupo 308 y produce como salida los bits asignados a subbandas específicas de grupo hacia la sección de codificación de retícula-vector 108 como el resultado de asignación de bits. Aquí, se distribuyen más bits a las subbandas perceptivamente importantes y se distribuyen menos bits a las subbandas perceptivamente menos importantes. Los bits distribuidos a cada subbanda en un grupo se calculan por la siguiente ecuación 6.
- 15 [6]

$$Bits_{G(k)sb(i)} = Bits(G(k)) \times \frac{Norm(i)}{Energy(G(k))} \quad \dots \quad (\text{Ecuación 6})$$

- 20 **[0077]** Aquí, $Bits_{G(k)sb(i)}$ 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, $Energy(G(k))$ indica la energía del grupo k, y $Norm(i)$ indica un valor de coeficiente de norma de la subbanda i del grupo k.

- [0078]** A continuación, se describirá un procedimiento de agrupamiento de acuerdo con una realización de la presente invención usando las FIGS. 8A a 8C. Supongamos que una envolvente espectral cuantificada mostrada en la FIG. 8A es introducida en la sección de identificación de banda de frecuencia pico 301. La sección de identificación de banda de frecuencia pico 301 identifica las bandas de frecuencia dominante 9 y 20 basándose en la envolvente espectral cuantificada introducida (véase la FIG. 8B).

- 30 **[0079]** Las secciones de generación de grupos dominantes 302-1 a 302N determinan subbandas centradas en y en ambos lados de la banda de frecuencia dominante 9 y 20 hasta las subbandas donde se detiene una pendiente descendiente del valor de coeficiente de norma como un grupo dominante idéntico. En los ejemplos de las FIGS. 8A a 8C, en cuanto a la banda de frecuencia dominante 9, las subbandas 6 a 12 se determinan como grupo dominante (grupo 2), mientras que en cuanto a la banda de frecuencia dominante 20, las subbandas 17 a 22 se determinan como grupo dominante (grupo 4) (véase la FIG. 8C).

- [0080]** La sección de determinación de grupos dominantes 303 determina bandas de frecuencia continuas distintas de los grupos dominantes como grupos no dominantes sin las bandas de frecuencia dominante. En el ejemplo de las FIGS. 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 FIG. 8C).

- [0081]** Como resultado, las envolventes espectrales cuantificadas son divididas en cinco grupos, es decir, dos grupos dominantes (grupos 2 y 4) y tres grupos no dominantes (grupos 1, 3 y 5).

- 45 **[0082]** Usando tal procedimiento de agrupación, es posible determinar adaptativamente anchuras de grupo según las características de la señal de entrada. Según este procedimiento, el aparato de decodificación de voz/audio también usa coeficientes de norma cuantificada disponibles, y por lo tanto no tiene que transmitirse información adicional al aparato de decodificación de voz/audio.

- 50 **[0083]** Obsérvese que la sección de cálculo de varianza de norma 306 calcula una varianza de norma específica de grupo. En los ejemplos de las FIGS. 8A a 8C, la varianza de norma $Energy_{var}(G(2))$ del grupo 2 se muestra en la FIG. 9 como referencia.

- [0084]** 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 constituido por una componente de espectro situado en una frecuencia dominante de la señal de voz/audio (componente de sonido dominante). El pico es perceptivamente muy importante. La importancia perceptiva de pico puede determinarse por una diferencia entre la energía del pico y la energía del valle, es decir, por una varianza de norma. Teóricamente, cuando un pico tiene una energía suficientemente grande en comparación con las bandas de frecuencia vecinas, el pico debería ser codificado con un número suficiente de bits, pero si el pico es codificado con un número insuficiente de bits, el ruido de codificación que se mezcla se vuelve destacado, haciendo que la calidad del sonido se deteriore. Por otra parte, un valle no está constituido por ninguna componente de sonido dominante de una señal de voz/audio y no es importante perceptivamente.
- 60

[0085] Según el procedimiento de agrupamiento de bandas de frecuencia de la presente realización, una banda de frecuencia dominante corresponde a un pico de un espectro y agrupar las bandas de frecuencia significa separar los picos (grupos dominantes que incluyen bandas de frecuencia dominante) de los valles (grupos no dominantes sin bandas de frecuencia dominante).

[0086] La sección de distribución de bits de grupo 308 determina la importancia perceptiva de un pico. A diferencia de la técnica de G.719 en la que la importancia perceptiva se determina sólo por la energía, la presente realización determina la importancia perceptiva basándose tanto en la energía como en las distribuciones de norma (energía) y determina los bits que han de ser distribuidos a cada grupo basándose en la importancia perceptiva determinada.

[0087] En la sección de distribución de bits de subbanda 309, cuando una varianza de norma en un grupo es grande, esto significa que este grupo es uno de picos, el pico es perceptivamente más importante y debería codificarse con exactitud un coeficiente de norma que tiene un valor máximo. Por esta razón, se distribuyen más bits a cada subbanda de este pico. Por otra parte, cuando una varianza de norma en un grupo es muy pequeña, esto significa que este grupo es uno de valles, y el valle no es perceptivamente importante y no es necesario que sea codificado con exactitud. Por esta razón, se distribuyen menos bits a cada subbanda de este grupo.

[0088] Así, 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 los bits a cada grupo basándose en la energía específica de grupo y las varianzas de norma, y además distribuye los bits distribuidos grupo por grupo a cada subbanda según una relación de una norma 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. Como resultado, la calidad del sonido puede mejorarse.

[0089] Obsérvese que el coeficiente de norma en la presente realización representa la energía de subbanda y también se denomina "envolvente de energía".

Aplicación industrial

[0090] El aparato de codificación de voz/audio, el aparato de decodificación de voz/audio, el procedimiento de codificación de voz/audio y el procedimiento de decodificación de voz/audio según la presente invención son aplicables a un aparato terminal de radiocomunicación, aparato de estación de base de radiocomunicación, aparato terminal de conferencia telefónica, aparato terminal de videoconferencia y aparato terminal de voz sobre Protocolo Internet (VoIP, del inglés "Voice over Internet Protocol") o similares.

Lista de signos de referencia

[0091]

- 101 Detector de transitorios
- 102 Sección de transformación
- 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 retícula-vector
- 109 Sección de ajuste de nivel de ruido
- 110 Multiplexor
- 201 Demultiplexor
- 202 Sección de decuantificación de norma
- 205 Sección de decodificación de retícula
- 206 Sección de relleno espectral
- 207 Sumador
- 208 Sección de modelado de envolvente
- 209 Sección de transformación inversa
- 301 Sección de identificación de banda de frecuencia dominante
- 302-1 a 302N Sección de determinación de grupos dominantes
- 303 Sección de determinación de grupos no dominantes
- 304 Sección de cálculo de energía de grupo
- 305 Sección de cálculo de energía total

306 Sección de cálculo de varianza de norma
307 Sección de cálculo de varianza de norma total
308 Sección de distribución de bits de grupo
309 Sección de distribución de bits de subbanda

5

REIVINDICACIONES

1. Un aparato de codificación de voz o audio que comprende:
 - 5 una sección de transformación (102) configurada para transformar una señal de entrada de un dominio del tiempo a un dominio de la frecuencia,
 - una sección de estimación (103) configurada para estimar una envolvente de energía que representa un nivel de energía para cada subbanda de una pluralidad de subbandas obtenidas dividiendo un espectro de frecuencia de la señal de entrada;
 - 10 una sección de cuantificación (104) configurada para cuantificar la envolvente de energía para obtener una envolvente de energía cuantificada que comprende energías de subbanda cuantificadas para la pluralidad de subbandas;
 - una sección de identificación de banda de frecuencia dominante (301) configurada para identificar una o más bandas de frecuencia dominante, siendo una banda de frecuencia dominante una subbanda en la que la
 - 15 envolvente de energía cuantificada del espectro de frecuencia tiene un valor máximo local,
 - una sección de determinación de grupos (302-1, 302-N, 303) configurada para agrupar las energías de subbanda cuantificadas en una pluralidad de grupos, en el que la sección de determinación de grupos (302-1, 302-N, 303) está configurada para determinar adaptativamente anchuras de grupos de la pluralidad de grupos según las características de la señal de entrada, y en el que la sección de determinación de grupos (302-1, 302-N, 303)
 - 20 está configurada para determinar cada banda de frecuencia dominante y las subbandas en ambos lados de la banda de frecuencia dominante que forman cada una una pendiente descendiente de la envolvente de energía cuantificada como un grupo dominante correspondiente y para determinar subbandas continuas aparte del uno o más grupos dominantes como uno o más grupos no dominantes;
 - una primera sección de asignación de bits (308) configurada para asignar bits a cada grupo de la pluralidad de grupos para obtener un número de bits específico de grupo para cada grupo de la pluralidad de grupos;
 - 25 una segunda sección de asignación de bits (309) configurada para asignar, para cada grupo de la pluralidad de grupos, el número de bits específico de grupo asignados a un grupo respectivo de la pluralidad de grupos a la pluralidad de subbandas que pertenecen al grupo respectivo; y
 - una sección de codificación (108) configurada para codificar, para cada subbanda de la pluralidad de subbandas,
 - 30 el espectro de frecuencia en la subbanda respectiva usando los bits asignados a la subbanda respectiva.
2. El aparato de codificación de voz o audio según la reivindicación 1, que comprende además:
 - una sección de cálculo de energía (304) configurada para calcular una energía específica de grupo; y
 - 35 una sección de cálculo de varianza de energía (306) configurada para calcular una varianza de energía específica de grupo, en el que la primera sección de asignación de bits (308) configurada para asignar, basándose en la energía específica de grupo y la varianza de energía específica de grupo calculadas, más bits a un grupo cuando al menos una de la energía y la varianza de energía es mayor y para asignar menos bits a un grupo cuando al menos una de la energía y la varianza de energía es menor.
 - 40
3. El aparato de codificación de voz o audio según la reivindicación 1, donde la segunda sección de asignación de bits (309) está configurada para asignar 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.
- 45 4. El aparato de codificación de voz o audio según la reivindicación 1, donde la segunda sección de asignación de bits (309) está configurada para asignar más bits a una subbanda perceptivamente más importante y menos bits a una subbanda perceptivamente menos importante.
5. El aparato de codificación de voz o audio según la reivindicación 1, donde la segunda sección de
- 50 asignación de bits (309) está configurada para asignar más bits a las subbandas de un grupo que tiene una varianza de energía más alta y para asignar menos bits a las subbandas de un grupo que tiene una varianza de energía más baja.
6. El aparato de codificación de voz o audio según la reivindicación 1, donde la segunda sección de
- 55 asignación de bits (309) está configurada para asignar más bits a las subbandas de un grupo que tiene un pico en el espectro de frecuencia y para asignar menos bits a las subbandas de un grupo que tiene un valle en el espectro de frecuencia.
7. El aparato de codificación de voz o audio según la reivindicación 1, donde la segunda sección de
- 60 asignación de bits (30) está configurada para operar basándose en la siguiente ecuación:

$$Bits_{G(k)sb(i)} = Bits(G(k)) \times \frac{Norm(i)}{Energy(G(k))} \quad \dots$$

donde $Bits_{G(k)sb(i)}$ indica un bit asignado a una subbanda i de un grupo k , i indica un índice de subbanda del grupo k , $Bits_{G(k)}$ indica un bit asignado al grupo k , $Energy(G(k))$ indica una energía del grupo k , y $Norm(i)$ indica un valor de energía de subbanda de la subbanda i del grupo k .

5 8. El aparato de codificación de voz o audio según la reivindicación 1, donde la primera sección de asignación de bits (308) está configurada para asignar más bits a un grupo dominante y menos bits a un grupo no dominante.

9. El aparato de codificación de voz o audio según la reivindicación 1,
10 donde la primera sección de asignación de bits (308) está configurada para asignar, grupo por grupo, bits basándose en una energía específica de grupo calculada por una sección de cálculo de energía de grupo (304), una energía total de todos los grupos calculada por una sección de cálculo de energía total (305), una varianza de energía específica de grupo calculada por una sección de cálculo de varianza de norma (306) y una varianza de energía total de todos los grupos calculada por una sección de cálculo de varianza de norma total (307).

15 10. El aparato de codificación de voz o audio según la reivindicación 9, en el que la primera sección de asignación de bits (308) está configurada para operar basándose en la siguiente ecuación:

$$Bits(G(k)) = Bits_{total} \times \left(scale1 \times \frac{Energy(G(k))}{Energy_{total}} + (1 - scale1) \times \frac{Norm_{var}(G(k))}{Norm_{var_{total}}} \right),$$

20 donde k indica un índice de cada grupo, $Bits(G(k))$ indica un número de bits asignados a un grupo k , $Bits_{total}$ indica un número total de bits disponibles, $scale1$ ajusta una relación de bits asignados por energía, $Energy(G(k))$ indica la energía específica de grupo del grupo k , $Energy_{total}$ indica la energía total de todos los grupos, $Norm_{var}(G(k))$ indica la varianza de energía específica de grupo del grupo k y $Norm_{var_{total}}$ indica la varianza de energía total
25 energía total de todos los grupos.

11. El aparato de codificación de voz o audio según la reivindicación 10, donde un valor de $scale1$ está entre 0 y 1.

30 12. El aparato de codificación de voz o audio según la reivindicación 1, donde la primera sección de asignación de bits (308) está configurada para determinar una importancia perceptiva de cada grupo usando una energía y una varianza de energía del grupo y para mejorar un grupo dominante más en comparación con un grupo no dominante.

35 13. El aparato de codificación de voz o audio según la reivindicación 1, donde la primera sección de asignación de bits (308) está configurada para determinar una importancia perceptiva de un grupo basándose en una energía del grupo y una distribución de energía y para determinar los bits que han de asignarse a cada grupo basándose en la importancia perceptiva para el grupo respectivo.

40 14. El aparato de codificación de voz o audio según la reivindicación 1, donde la sección de determinación de grupos (302-1, 302-N, 303) está configurada para incluir, cuando una banda de frecuencia dominante está situada en un extremo de una frecuencia disponible, sólo un lado de la pendiente descendente en el grupo correspondiente.

45 15. El aparato de codificación de voz o audio según la reivindicación 1, donde la sección de determinación de grupos (302-1, 302-N, 303) está configurada para separar los picos del espectro de frecuencia de los valles del espectro de frecuencia, en el que un pico del espectro de frecuencia está situado en un grupo dominante y un valle del espectro de frecuencia está situado en un grupo no dominante.

50 16. El aparato de codificación de voz o audio según la reivindicación 1,
donde la primera sección de asignación de bits (308) está configurada para asignar bits a un grupo respectivo basándose en una energía del grupo respectivo y una varianza de energía del grupo respectivo, y
55 donde la segunda sección de asignación de bits (309) está configurada para asignar los bits, asignados grupo por grupo al grupo respectivo, a una subbanda respectiva del grupo respectivo según una relación de una energía de la subbanda respectiva a una energía del grupo respectivo.

17. El aparato de codificación de voz o audio según la reivindicación 1,
60 donde la primera sección de asignación de bits (308) está configurada para asignar más bits a un grupo perceptivamente más importante y menos bits a un grupo perceptivamente menos importante, y
donde la segunda sección de asignación de bits (309) está configurada para asignar más bits a una subbanda

perceptivamente más importante y menos bits a una subbanda perceptivamente menos importante.

18. Un aparato de decodificación de voz o audio que comprende:

- 5 una sección de decuantificación (202) configurada para decuantificar una envolvente espectral cuantificada que comprende energías de subbanda cuantificadas para obtener una envolvente espectral decuantificada;
- una sección de identificación de banda de frecuencia dominante (301) configurada para identificar una o más bandas de frecuencia dominante, siendo una banda de frecuencia dominante una subbanda en la que la envolvente de energía cuantificada del espectro de frecuencia tiene un valor máximo local;
- 10 una sección de determinación de grupos (302-1, 302-N, 303) configurada para agrupar las energías de subbanda cuantificadas en una pluralidad de grupos, en el que la sección de determinación de grupos (302-1, 302-N, 303) está configurada para determinar adaptativamente anchuras de grupos de la pluralidad de grupos según las características de la señal de entrada, y en el que la sección de determinación de grupos (302-1, 302-N, 303) está configurada para determinar cada banda de frecuencia dominante y las subbandas en ambos lados de la
- 15 banda de frecuencia dominante que forman cada una una pendiente descendiente de la envolvente de energía cuantificada como un grupo dominante correspondiente y para determinar subbandas continuas aparte del uno o más grupos dominantes como uno o más grupos no dominantes;
- una primera sección de asignación de bits (308) configurada para asignar bits a cada grupo de la pluralidad de grupos para obtener un número de bits específico de grupo para cada grupo de la pluralidad de grupos;
- 20 una segunda sección de asignación de bits (309) configurada para asignar, para cada grupo de la pluralidad de grupos, el número de bits específico de grupo asignados a un grupo respectivo de la pluralidad de grupos a la pluralidad de subbandas que pertenecen al grupo respectivo;
- una sección de decodificación (205) configurada para decodificar, para cada subbanda de la pluralidad de subbandas, un espectro de frecuencia de una señal de voz o audio usando los bits asignados a la subbanda
- 25 respectiva para obtener un espectro de frecuencia decodificado;
- una sección de modelado de envolvente (208) configurada para aplicar la envolvente espectral decuantificada al espectro de frecuencia decodificado y para reproducir un espectro decodificado; y
- una sección de transformación inversa (209) configurada para transformar inversamente el espectro decodificado de un dominio de la frecuencia a un dominio del tiempo.

19. El aparato de decodificación voz o audio según la reivindicación 18, que además comprende:

- una sección de cálculo de energía (304) configurada para calcular una energía específica de grupo; y
- 35 una sección de cálculo de varianza de energía (306) configurada para calcular una varianza de energía específica de grupo, donde la primera sección de asignación de bits (308) está configurada para asignar, basándose en la energía específica de grupo y la varianza de energía específica de grupo calculadas, más bits a un grupo cuando al menos una de la energía y la varianza de energía es mayor y para asignar menos bits a un grupo cuando al menos una de la energía y la varianza de energía es menor.
- 40 20. El aparato de decodificación de voz o audio según la reivindicación 18, donde la segunda sección de asignación de bits (309) está configurada para asignar más bits a una subbanda que tiene una envolvente de energía mayor y para asignar menos bits a una subbanda que tiene una envolvente de energía menor.

21. Un procedimiento de codificación de voz o audio que comprende:

- 45 transformar una señal de entrada de un dominio del tiempo a un dominio de la frecuencia;
- estimar una envolvente de energía que representa un nivel de energía para cada subbanda de una de la pluralidad de subbandas obtenidas dividiendo un espectro de frecuencia de la señal de entrada;
- 50 cuantificar la envolvente de energía para obtener una envolvente de energía cuantificada que comprende energías de subbanda cuantificadas para una pluralidad de subbandas;
- identificar una o más bandas de frecuencia dominante, siendo una banda de frecuencia dominante una subbanda en la que la envolvente de energía cuantificada del espectro de frecuencia tiene un valor máximo local;
- agrupar las energías de subbanda cuantificadas para la pluralidad de subbandas en una pluralidad de grupos, teniendo cada grupo una pluralidad de al menos dos subbandas, en el que la agrupación comprende determinar adaptativamente anchuras de grupos de la pluralidad de grupos según las características de la señal de entrada,
- 55 determinar cada banda de frecuencia dominante y las subbandas en ambos lados de la banda de frecuencia dominante que forman cada una una pendiente descendiente de la envolvente de energía cuantificada como un grupo dominante correspondiente, y determinar subbandas continuas aparte del uno o más grupos dominantes como uno o más grupos no dominantes;
- 60 asignar, para cada grupo de la pluralidad de grupos, bits a cada grupo de la pluralidad de grupos para obtener un número de bits específico de grupo para cada grupo de la pluralidad de grupos;
- asignar, para cada grupo de la pluralidad de grupos, el número de bits específico de grupo asignados a un grupo respectivo de la pluralidad de grupos a la pluralidad de subbandas que pertenecen al grupo respectivo; y
- 65 codificar, para cada subbanda de la pluralidad de subbandas, el espectro de frecuencia incluido en la subbanda respectiva usando los bits asignados a la subbanda respectiva.

22. Un procedimiento de decodificación de voz o audio que comprende:
- decuantificar una envolvente espectral cuantificada que comprende energías de subbanda cuantificadas para obtener una envolvente espectral decuantificada;
 - 5 identificar una o más bandas de frecuencia dominante, siendo una banda de frecuencia dominante una subbanda donde la envolvente de energía cuantificada del espectro de frecuencia tiene un valor máximo local;
 - agrupar las energías de subbanda cuantificadas en una pluralidad de grupos, en el que la agrupación comprende determinar adaptativamente anchuras de grupos de la pluralidad de grupos según las características de la señal de entrada, determinar cada banda de frecuencia dominante y las subbandas en ambos lados de la banda de
 - 10 frecuencia dominante que forman cada una una pendiente descendiente de la envolvente de energía cuantificada como un grupo dominante correspondiente, y determinar subbandas continuas aparte del uno o más grupos dominantes como uno o más grupos no dominantes;
 - asignar, para cada grupo de la pluralidad de grupos, bits a cada grupo de la pluralidad de grupos para obtener un número de bits específico de grupo para cada grupo de la pluralidad de grupos;
 - 15 asignar, para cada grupo de la pluralidad de grupos, el número de bits específico de grupo asignados a un grupo respectivo de la pluralidad de grupos a una pluralidad de subbandas que pertenecen al grupo respectivo;
 - decodificar, para cada subbanda de la pluralidad de subbandas, un espectro de frecuencia de una señal de voz o audio usando los bits asignados a la subbanda respectiva para obtener un espectro de frecuencia decodificado;
 - 20 aplicar la envolvente espectral decuantificada al espectro de frecuencia decodificado y reproducir un espectro decodificado; y
 - transformar inversamente el espectro decodificado de un dominio de la frecuencia a un dominio del tiempo.
23. Aparato terminal de radiocomunicación, aparato de estación de base de radiocomunicación, aparato terminal de conferencia telefónica, aparato terminal de videoconferencia, o aparato terminal de voz sobre Protocolo
- 25 Internet (VoIP) que comprende un aparato de codificación de voz o audio según la reivindicación 1, o un aparato de decodificación de voz o audio según la reivindicación 18.

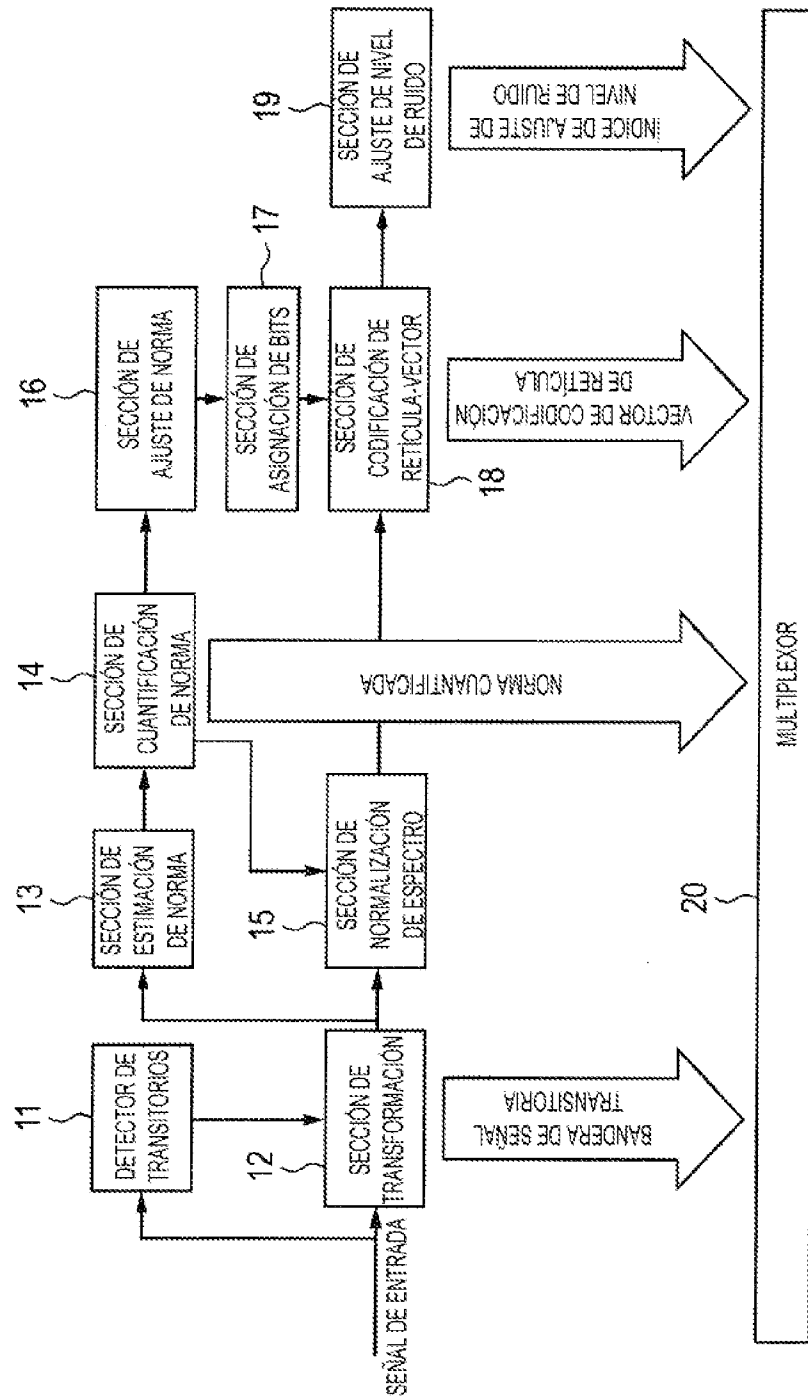


FIG. 1

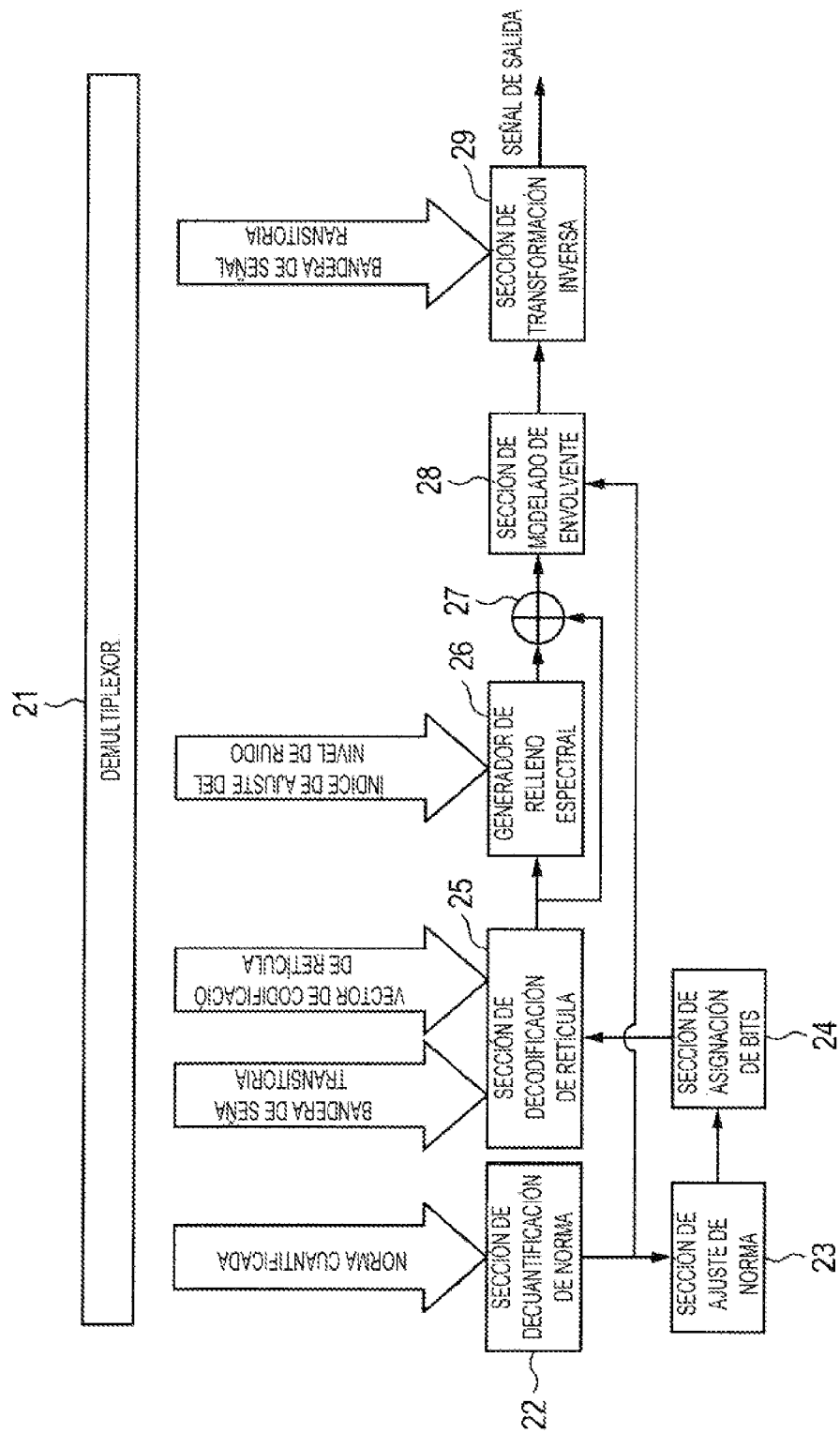


FIG. 2

GRUPO	LONGITUD DE SUBVECTOR	NÚMERO DE SUBVECTORES	NÚMERO DE COEFICIENTES	ANCHURA DE BANDA (Hz)	INICIO (Hz)	FIN (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

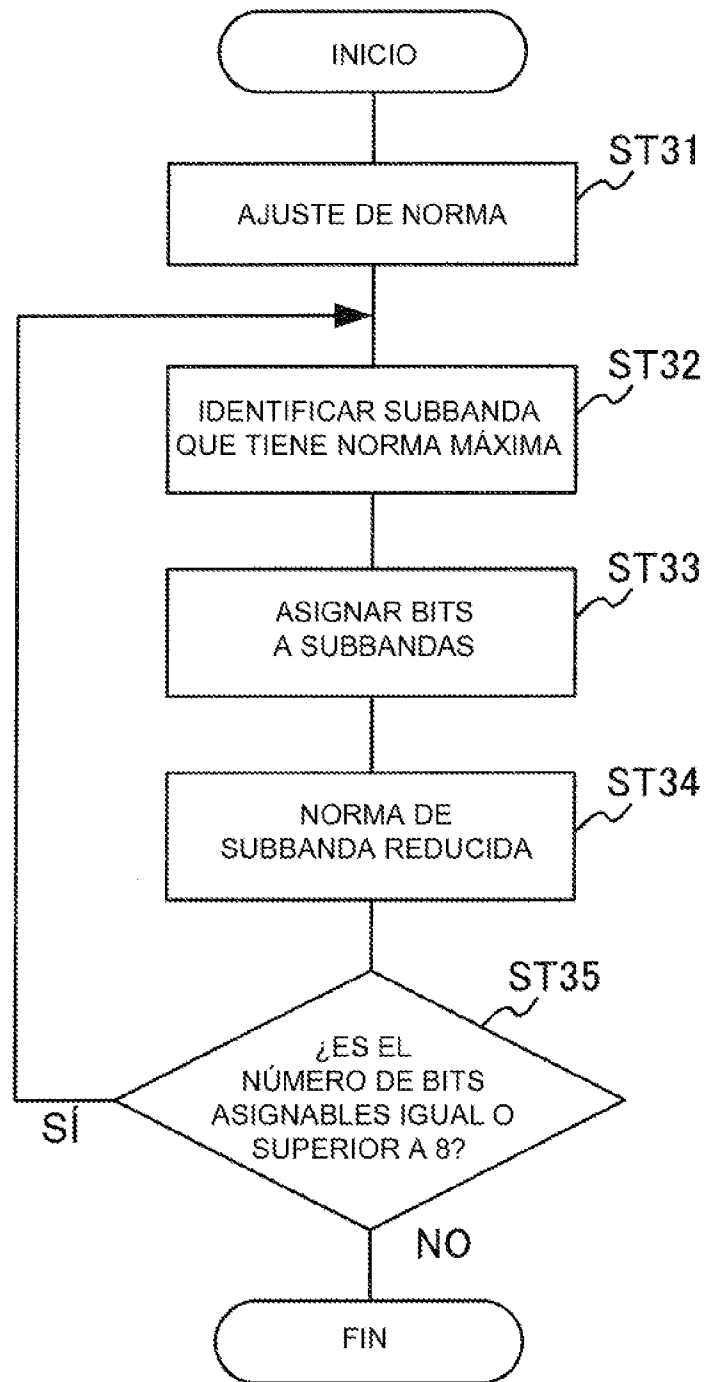


FIG. 4

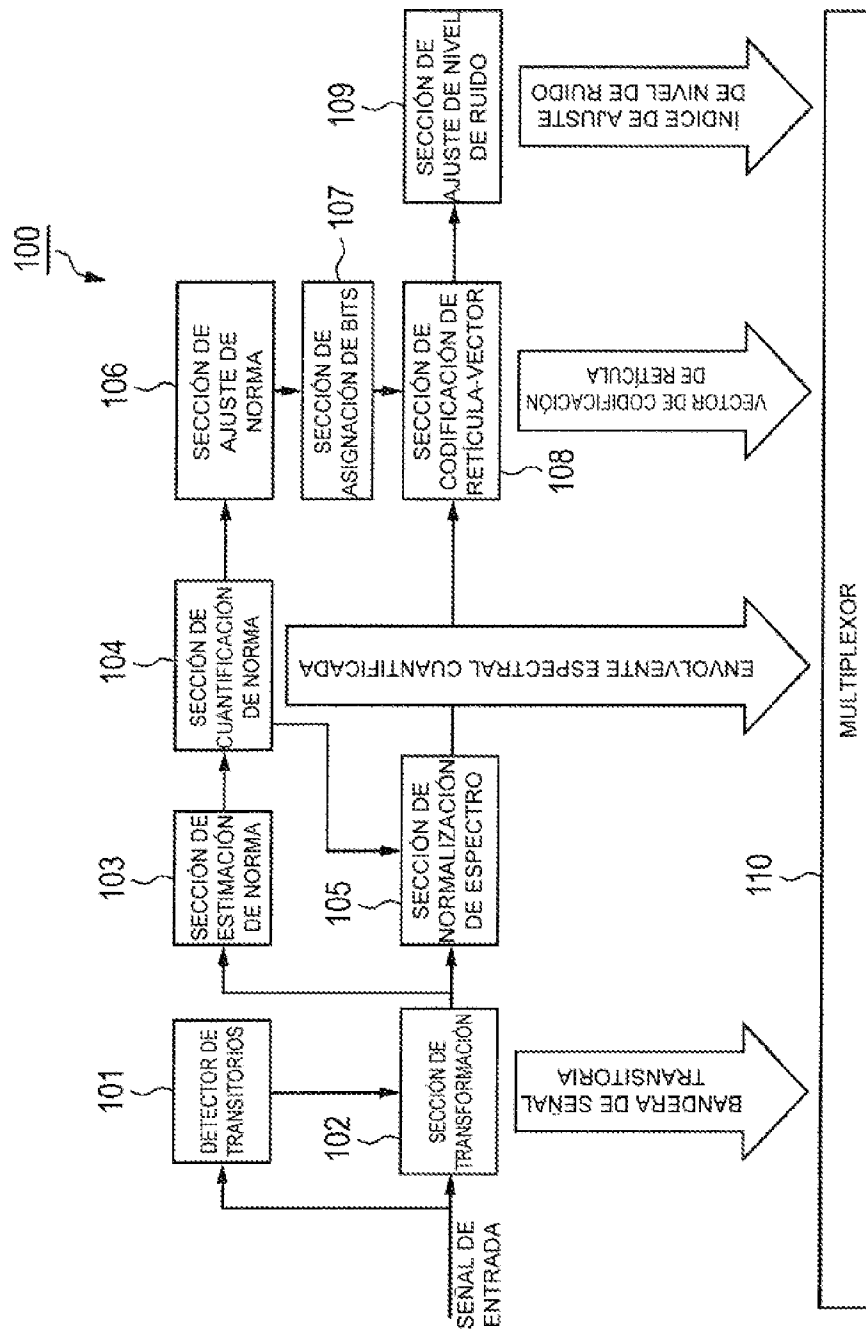


FIG. 5

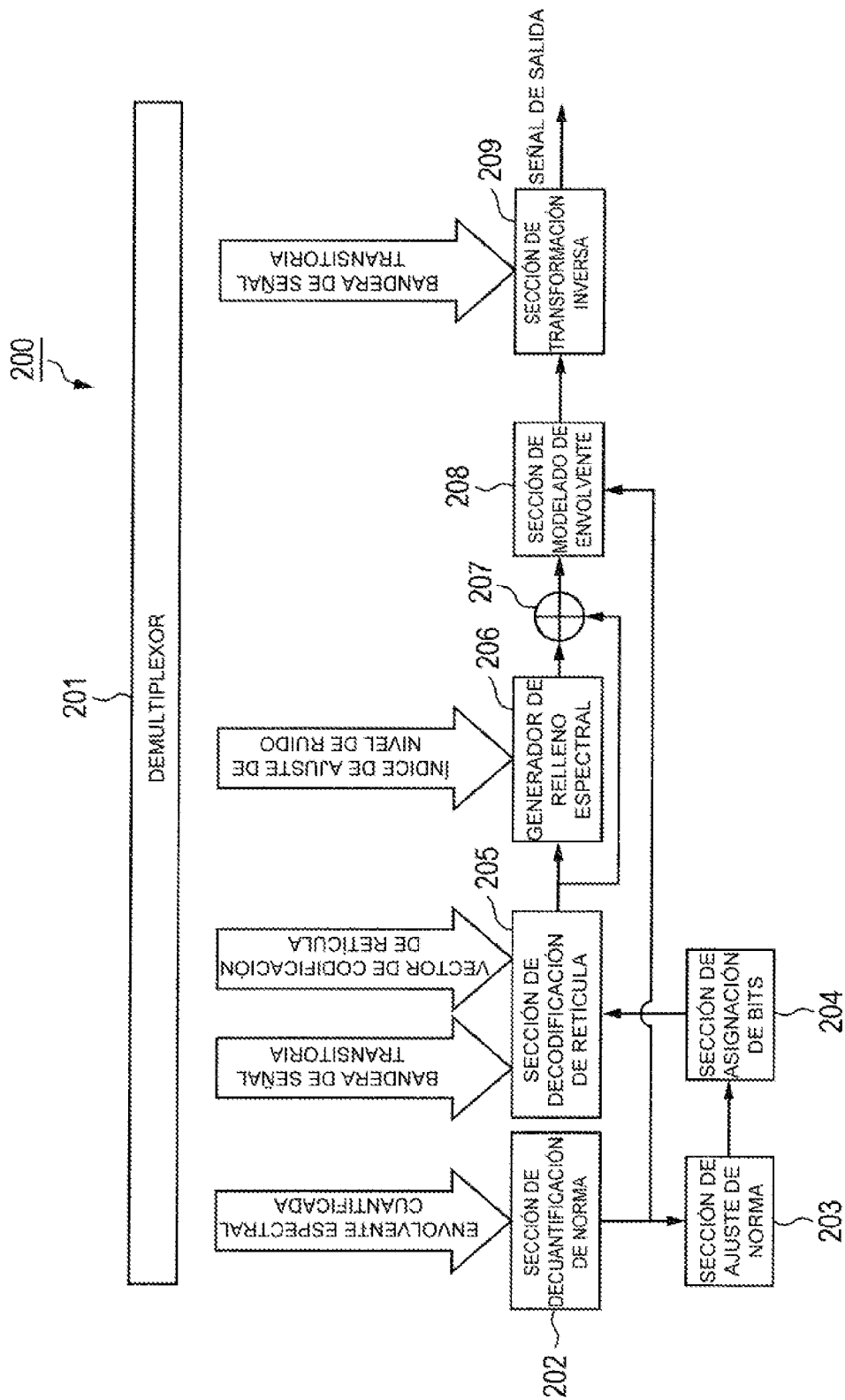


FIG. 6

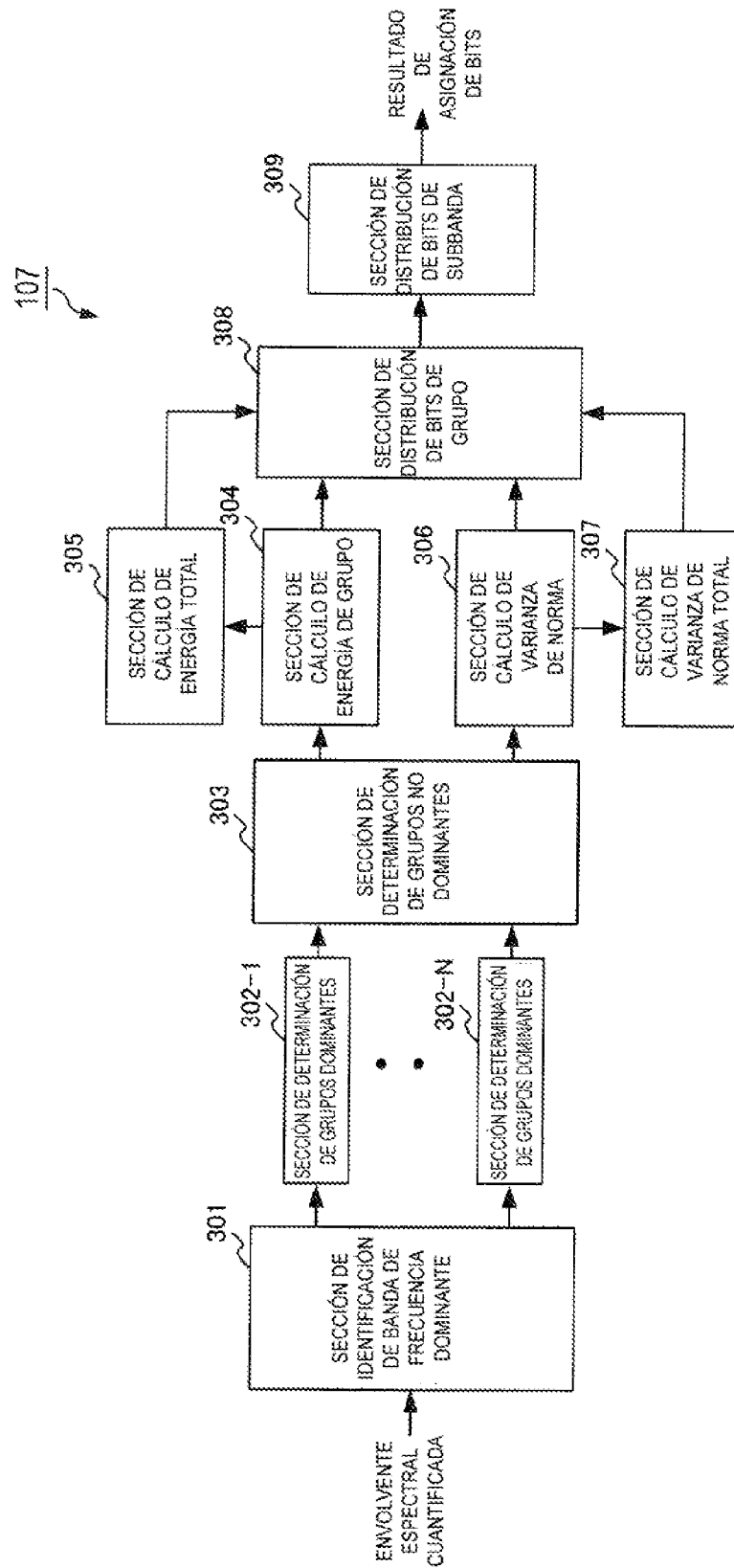


FIG. 7

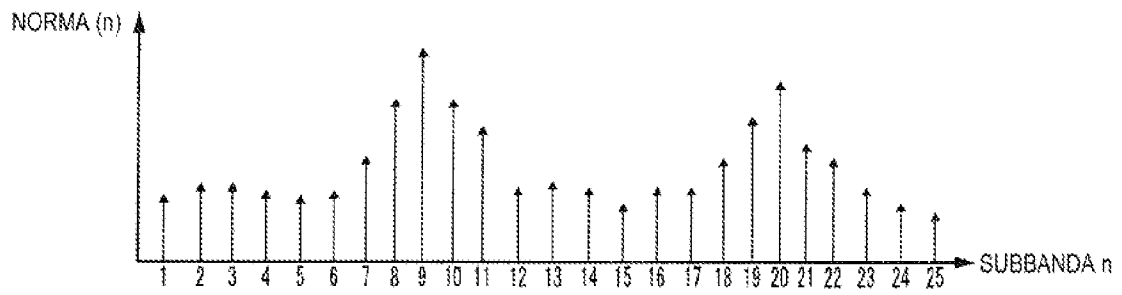


FIG. 8A

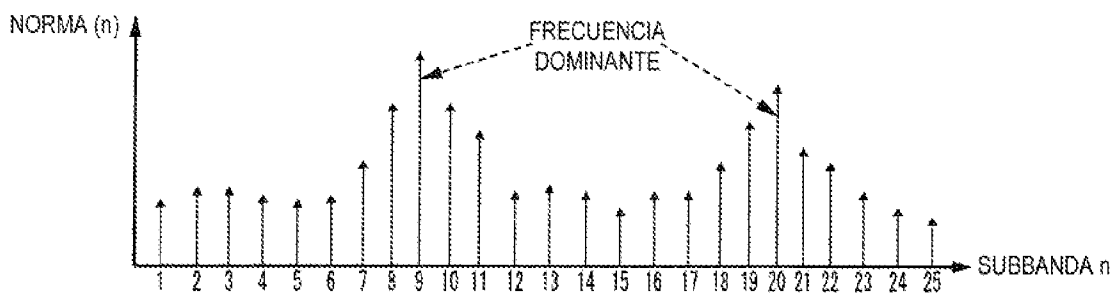


FIG. 8B

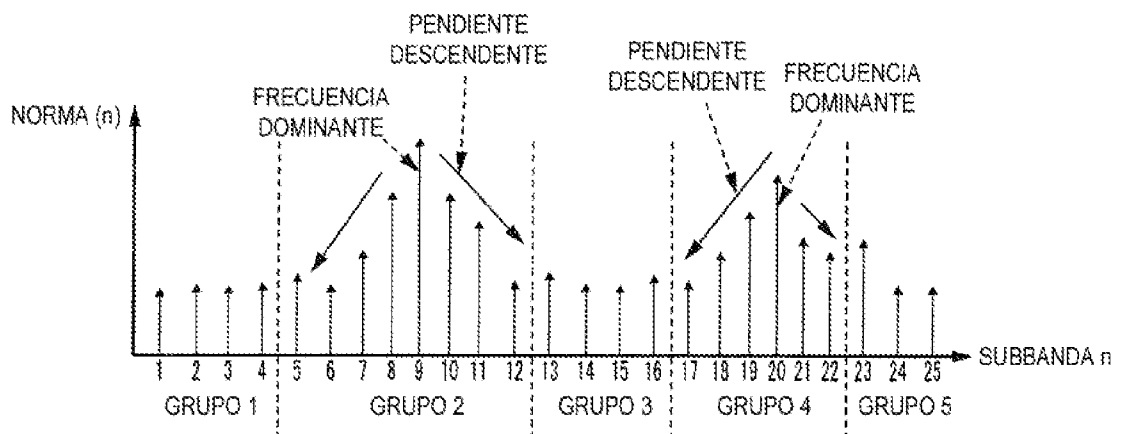


FIG. 8C

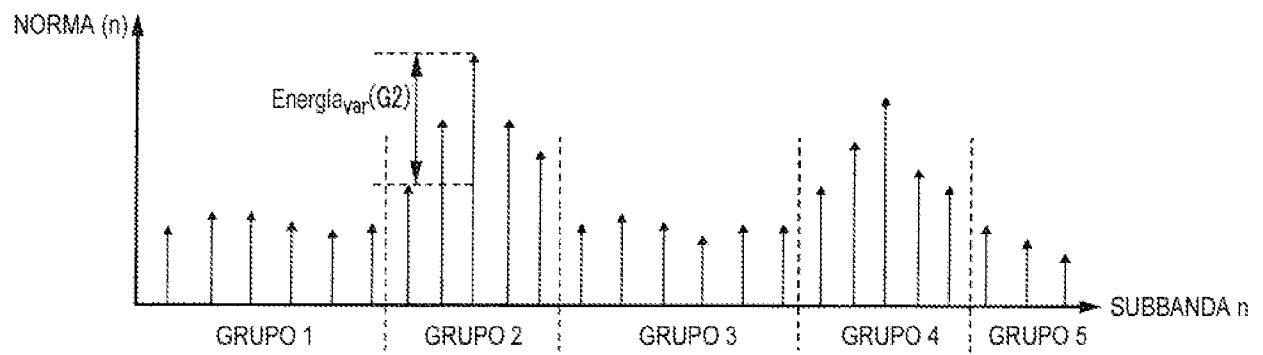


FIG. 9