

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 404 408**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 19/10 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.02.2008** **E 08720311 (3)**

97 Fecha y número de publicación de la concesión europea: **10.04.2013** **EP 2128858**

54 Título: **Dispositivo de codificación y método de codificación**

30 Prioridad:

02.03.2007 JP 2007053497

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

27.05.2013

73 Titular/es:

**PANASONIC CORPORATION (100.0%)
1006, OAZA KADOMA KADOMA-SHI
OSAKA 571-8501, JP**

72 Inventor/es:

**MORII, TOSHIYUKI;
OSHIKIRI, MASAHIRO y
YAMANASHI, TOMOFUMI**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 404 408 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 y método de codificación

5 Campo técnico

La presente invención se refiere a un aparato de codificación y método de codificación para codificar señales del habla y señales de audio.

10 Antecedentes de la invención

En comunicaciones móviles, es necesario comprimir y codificar información digital tal como el habla e imágenes para uso eficaz de capacidad del canal de radio y medios de almacenamiento para ondas de radio, y se han desarrollado muchos esquemas de codificación y decodificación hasta la fecha.

15 Entre estos, se ha mejorado significativamente el rendimiento de la tecnología de codificación del habla mediante el esquema fundamental de "CELP (Predicción Lineal con Excitación por Código)", que adopta hábilmente cuantificación vectorial modelando el sistema del tracto vocal del habla. Además, se ha mejorado significativamente el rendimiento de la tecnología de codificación de sonido tal como codificación de audio transformando técnicas de codificación (tales como los estándares MPEG ACC y MP3).

25 Por otro lado, se diseña un códec escalable, la estandarización del cual está en progreso por la ITU-T (Unión Internacional de Telecomunicaciones - Sector de Estandarización de las Telecomunicaciones) y otros, para cubrir desde la banda del habla convencional (300 Hz a 3,4 kHz) hasta la banda ancha (hasta 7 kHz), con su velocidad de bits establecida tan alta como aproximadamente 32 kbps. Es decir, un código de banda ancha tiene incluso que aplicar un cierto grado de codificación a audio y por lo tanto no se puede soportar mediante únicamente métodos de codificación del habla de baja velocidad de bits convencionales, en base al modelo de voz humana, tal como CELP. Ahora, el estándar ITU-T G.729.1, declarado anteriormente como una recomendación, usa un esquema de codificación de códec de audio de codificación por transformación, para codificar el habla de banda ancha y superior.

30 El Documento de Patente 1 desvela un esquema de codificación que utiliza parámetros espectrales y parámetros de paso, en los cuales se realiza una transformación ortogonal y codificación de una señal obtenida mediante filtrado inverso a una señal del habla en base a parámetros espectrales, y adicionalmente desvela, como un ejemplo de codificación, un método de codificación en base a libros de código de estructuras algebraicas.

35 El Documento de Patente 2 desvela un esquema de codificación de división de una señal en los parámetros de predicción lineales y los componentes residuales, realizando la transformación cuadrática de los componentes residuales y normalizando la forma de onda residual por la potencia, y a continuación cuantificando la ganancia y el residuo normalizado. Adicionalmente, el Documento de Patente 2 desvela cuantificación vectorial como un método de cuantificación para el residuo normalizado.

40 El Documento distinto de Patente 1 desvela un método de codificación en base a un libro de código algebraico formado con espectros de excitación mejorados en TCX (es decir un esquema de codificación fundamental modelado con una excitación sometida a una codificación por transformación y filtrado de parámetros espectrales), y se adopta este método de codificación en el estándar ITU-T G.729.1.

45 El Documento distinto de Patente 2 desvela la descripción del esquema del estándar MPEG "TC-WVQ". Este esquema se usa también para transformar residuo de predicción lineal en un espectro y realizar cuantificación vectorial del espectro, usando la DCT (Transformación Discreta en Coseno) como el método de transformación ortogonal.

50 Por medio de las cuatro técnicas anteriores, es posible aplicar, a codificación, cuantificación de parámetros espectrales tales como parámetros de predicción lineal, que es parte de una técnica de codificación útil de señales del habla, posibilitando de esta manera la eficacia y la baja velocidad de codificación de audio a realizar.

55 Documento de Patente 1: Solicitud de Patente Japonesa Abierta a Inspección Pública Nº HE110-260698
 Documento de Patente 2: Solicitud de Patente Japonesa Abierta a Inspección Pública Nº HE107-261800
 Documento distinto de Patente 1: Xie, Adoul, "EMBEDDED ALGEBRAIC VECTOR QUANTIZERS (EAVQ) WITH APPLICATION TO WIDEBAND SPEECH CODING" ICASSP'96
 Documento distinto de Patente 2: Moriya, Honda, "Transform Coding of Speech Using a Weighted Vector Quantizer" IEEE journal on selected areas in communications, Vol. 6, Nº 2, febrero 1988

Divulgación de la invención

Problemas a resolver mediante la invención

5 Sin embargo, el número de bits a asignar mediante un códec escalable es pequeño especialmente en una capa relativamente inferior, y, en consecuencia, el rendimiento de la codificación por transformación de excitación no es suficiente. Por ejemplo, en el estándar ITU-T G.729.1, aunque una velocidad de bits es 12 kbps en la segunda capa o inferior que soporta la banda telefónica (300 Hz a 3,4 kHz), únicamente se asigna una velocidad de bits de 2 kbps a la siguiente, tercera capa que soporta una banda ancha (50 Hz a 7 kHz). Por lo tanto, cuando hay unos pocos bits de información, no es posible conseguir suficiente rendimiento perceptivo usando un método de codificación de un espectro, que se obtiene mediante una transformación ortogonal, con cuantificación vectorial usando un libro de código.

10 Es por lo tanto un objeto de la presente invención proporcionar un aparato de codificación y método de codificación que puedan conseguir buena calidad perceptiva incluso si hay pocos bits de información.

Medios para resolver el problema

20 Se define una solución a los problemas anteriores mediante un aparato de codificación de acuerdo con la reivindicación 1 y un método de codificación de acuerdo con la reivindicación 5.

Efectos ventajosos de la invención

25 De acuerdo con la presente invención, es posible codificar frecuencias (posiciones) con precisión donde la energía está presente, de modo que es posible mejorar rendimiento cualitativo, que es único para codificación de espectro, y producir buena calidad de sonido incluso a bajas velocidades de bits.

Breve descripción de los dibujos

30 La Figura 1 es un diagrama de bloques que muestra la configuración de un aparato de codificación del habla de acuerdo con una realización de la presente invención;
La Figura 2 es un diagrama de bloques que muestra la configuración de un aparato de decodificación del habla de acuerdo con una realización de la presente invención;
35 La Figura 3 es un diagrama de flujo que muestra el algoritmo de búsqueda en una sección de búsqueda de intervalo de acuerdo con una realización de la presente invención;
La Figura 4 es un diagrama que muestra un ejemplo de un espectro representado mediante pulsos buscados en una sección de búsqueda de intervalo de acuerdo con una realización de la presente invención;
La Figura 5 es un diagrama de flujo que muestra el algoritmo de búsqueda en una sección de búsqueda minuciosa de acuerdo con una realización de la presente invención;
40 La Figura 6 es un diagrama de flujo que muestra el algoritmo de búsqueda en una sección de búsqueda minuciosa de acuerdo con una realización de la presente invención;
La Figura 7 es un diagrama que muestra un ejemplo de un espectro representado mediante pulsos buscados en una sección de búsqueda de intervalo y sección de búsqueda minuciosa de acuerdo con una realización de la presente invención;
45 La Figura 8 es un diagrama de flujo que muestra el algoritmo de decodificación en una sección de decodificación de espectro de acuerdo con una realización de la presente invención.

Mejor modo de llevar a cabo la invención

50 En la codificación de señal del habla en base al esquema CELP y otros, a menudo se representa una señal del habla mediante una excitación y filtro de síntesis. Si se puede decodificar un vector que tiene una forma similar a una señal de excitación, que es una secuencia vectorial en el dominio del tiempo, es posible producir una forma de onda similar al habla de entrada a través de un filtro de síntesis, y conseguir buena calidad perceptiva. Esta es la característica cualitativa que ha conducido al éxito del libro de código algebraico usado en CELP.

55 Por otro lado, en el caso de codificación de espectro (vector) de frecuencia, un filtro de síntesis tiene como sus componentes ganancias espectrales, y por lo tanto la distorsión de las frecuencias (es decir posiciones) de los componentes de gran potencia es más significativa que la distorsión de esas ganancias. Es decir, buscando posiciones de alta energía y decodificando los pulsos en las posiciones de alta energía, en lugar de decodificar un vector que tenga una forma similar a un espectro de entrada, es más probable conseguir buena calidad perceptiva.

60 Los presentes inventores se centraron en este punto y llegaron a la presente invención. Es decir, en base a un modelo de codificación de un espectro de frecuencia mediante un pequeño número de pulsos, la presente invención transforma una señal del habla para codificar (es decir secuencia vectorial en el dominio del tiempo) en una señal de dominio en frecuencia mediante una transformación ortogonal, divide el intervalo de frecuencia del objetivo de codificación en una pluralidad de bandas, y busca un pulso en cada banda, y, además, busca varios pulsos sobre el

intervalo de frecuencia completo del objetivo de codificación.

Además, la presente invención separa la cuantificación de forma (apariencia) y cuantificación de ganancia (cantidad), y, en cuantificación de forma, supone una ganancia ideal y busca pulsos que tengan una amplitud “1” y una polaridad “+” o “-”, en un bucle abierto. En este punto, especialmente tras una búsqueda sobre el intervalo de frecuencia completo del objetivo de codificación, la presente invención no permite que aparezcan dos pulsos en la misma posición y permite combinaciones de las posiciones de una pluralidad de pulsos a codificar como información de transmisión acerca de posiciones de pulso.

10 Se explicará a continuación una realización de la presente invención usando los dibujos adjuntos.

La Figura 1 es un diagrama de bloques que muestra la configuración del aparato de codificación del habla de acuerdo con la presente realización. Se proporciona el aparato de codificación del habla mostrado en la Figura 1 con la sección 101 de análisis LPC, la sección 102 de cuantificación LPC, el filtro 103 inverso, la sección 104 de transformación ortogonal, la sección 105 de codificación de espectro y la sección 106 de multiplexación. Se proporciona la sección 105 de codificación de espectro con la sección 111 de cuantificación de forma y la sección 112 de cuantificación de ganancia.

La sección 101 de análisis LPC realiza un análisis de predicción lineal de una señal del habla de entrada y emite un parámetro envolvente espectral a la sección 102 de cuantificación LPC como un resultado de análisis. La sección 102 de cuantificación LPC realiza procesamiento de cuantificación del parámetro envolvente espectral (LPC: Coeficiente de Predicción Lineal) emitido desde la sección 101 de análisis LPC, y emite un código que representa la cuantificación LPC, a la sección 106 de multiplexación. Además, la sección 102 de cuantificación LPC emite parámetros decodificados obtenidos decodificando el código que representa el LPC cuantificado, al filtro 103 inverso. En este punto, la cuantificación de parámetro puede emplear la cuantificación vectorial (“VQ”), cuantificación de predicción, VP multi-etapa, VQ de división y otros modos.

El filtro 103 inverso filtra de manera inversa el habla de entrada usando los parámetros decodificados y emite el componente residual resultante a la sección 104 de transformación ortogonal.

La sección 104 de transformación ortogonal aplica una ventana de correspondencia, tal como una ventana seno, al componente residual, realiza una transformación ortogonal usando MDCT, y emite un espectro transformado en un espectro en dominio de frecuencia (en lo sucesivo “espectro de entrada”), a la sección 105 de codificación de espectro. En este punto, la transformación ortogonal puede emplear otras transformaciones tales como la FFT, KLT y transformación Wavelet, y, aunque su uso varía, es posible transformar el componente residual en un espectro de entrada usando cualquiera de estas.

En este punto, se puede invertir el orden de procesamiento entre el filtro 103 inverso y la sección 104 de transformación ortogonal. Es decir, dividiendo el habla de entrada sometida a una transformación ortogonal mediante el espectro de frecuencia de un filtro inverso (es decir resta en el eje logarítmico), es posible producir el mismo espectro de entrada.

La sección 105 de codificación de espectro divide el espectro de entrada cuantificando la forma y ganancia del espectro por separado, y emite los códigos de cuantificación resultantes a la sección 106 de multiplexación. La sección 111 de cuantificación de forma cuantifica la forma del espectro de entrada usando un pequeño número de posiciones de pulso y polaridades, y la sección 112 de cuantificación de ganancia calcula y cuantifica las ganancias de los pulsos averiguados mediante la sección 111 de cuantificación de forma, para cada banda. Se describirán más adelante en detalle la sección 111 de cuantificación de forma y la sección 112 de cuantificación de ganancia.

La sección 106 de multiplexación recibe como entrada un código que representa la cuantificación LPC desde la sección 102 de cuantificación LPC y un código que representa el espectro de entrada cuantificado desde la sección 105 de codificación de espectro, multiplexa esta información y emite el resultado al canal de transmisión como información de codificación.

La Figura 2 es un diagrama de bloques que muestra la configuración del aparato de decodificación del habla de acuerdo con la presente realización. Se proporciona el aparato de decodificación del habla mostrado en la Figura 2 con la sección 201 de demultiplexación, la sección 202 de decodificación de parámetro, la sección 203 de decodificación de espectro, la sección 204 de transformación ortogonal y el filtro 205 de síntesis.

En la Figura 2, se demultiplexa la información de codificación en códigos individuales en la sección 201 de demultiplexación. Se emite el código que representa el LPC cuantificado a la sección 202 de decodificación de parámetro, y el código del espectro de entrada se emite a la sección 203 de decodificación de espectro.

La sección 202 de decodificación de parámetro decodifica el parámetro envolvente espectral y emite el parámetro decodificado resultante al filtro 205 de síntesis.

La sección 203 de decodificación de espectro decodifica el vector de forma y ganancia mediante el método que soporta el método de codificación en la sección 105 de codificación de espectro mostrada en la Figura 1, obtiene un espectro decodificado multiplicando el vector de forma decodificado por la ganancia decodificada, y emite el espectro decodificado a la sección 204 de transformación ortogonal.

5 La sección 204 de transformación ortogonal realiza una transformación inversa del espectro decodificado emitido desde la sección 203 de decodificación de espectro comparado con la sección 104 de transformación ortogonal mostrada en la Figura 1, y emite la señal residual decodificada en serie temporal resultante al filtro 205 de síntesis.

10 El filtro 205 de síntesis produce el habla de salida aplicando filtrado de síntesis a la señal residual decodificada emitida desde la sección 204 de transformación ortogonal usando el parámetro decodificado emitido desde la sección 202 de decodificación de parámetro.

15 En este punto, para invertir el orden de procesamiento entre el filtro 103 inverso y la sección 104 de transformación ortogonal mostrada en la Figura 1, el aparato de decodificación del habla en la Figura 2 multiplica el espectro decodificado por un espectro de frecuencia del parámetro decodificado (es decir adición en el eje logarítmico) y realiza una transformación ortogonal del espectro resultante.

20 A continuación, se explicarán en detalle la sección 111 de cuantificación de forma y la sección 112 de cuantificación de ganancia. Se proporciona la sección 111 de cuantificación de forma con la sección 121 de búsqueda de intervalo que busca pulsos en cada una de una pluralidad de bandas en las que se divide un intervalo de búsqueda predeterminado, y la sección 122 de búsqueda minuciosa que busca pulsos sobre el intervalo de búsqueda completo.

25 La siguiente ecuación 1 proporciona una referencia para búsqueda. En este punto, en la ecuación 1, E es la distorsión de codificación, s_i es el espectro de entrada, g es la ganancia óptima, δ es la función delta y p es la posición de pulso.

[1]

$$E = \sum_i \{s_i - g\delta(i - p)\}^2 \quad \dots(\text{Ecuación 1})$$

30 A partir de la ecuación 1 anterior, la posición de pulso para minimizar la función de coste es la posición en la que el valor absoluto $|S_p|$ del espectro de entrada en cada banda es máximo, y su polaridad es la polaridad del valor del espectro de entrada en la posición de ese pulso.

35 Se explicará a continuación un caso de ejemplo donde la longitud del vector de un espectro de entrada es ochenta muestras, el número de bandas es cinco y se codifica el espectro usando ocho pulsos, un pulso de cada banda y tres pulsos de la banda completa. En este caso, la longitud de cada banda es dieciséis muestras. Además, se fija la amplitud de los pulsos a buscar a "1", y su polaridad es "+" o "-".

40 La sección 121 de búsqueda de intervalo busca la posición de la máxima energía y la polaridad (+/-) en cada banda, y permite que aparezca un pulso por banda. En este ejemplo, el número de bandas es cinco, y cada banda requiere cuatro bits para mostrar la posición de pulso (entradas de posiciones: 16) y un bit para mostrar la polaridad (+/-), que requiere veinticinco bits de información en total.

45 Se muestra el flujo del algoritmo de búsqueda de la sección 121 de búsqueda de intervalo en la Figura 3. En este punto, los símbolos usados en el diagrama de flujo de la Figura 3 representan los siguientes contenidos.

- i: posición
- b: número de banda
- máx: máximo valor
- 50 c: contador
- pos[b]: resultado de búsqueda (posición)
- pol[b]: resultado de búsqueda (polaridad)
- s[i]: espectro de entrada

55 Como se muestra en la Figura 3, la sección 121 de búsqueda de intervalo calcula el espectro de entrada $s[i]$ de cada muestra ($0 \leq c \leq 15$) por banda ($0 \leq b \leq 4$), y calcula el máximo valor "máx".

60 La Figura 4 ilustra un ejemplo de un espectro representado mediante pulsos averiguados mediante la sección 121 de búsqueda de intervalo. Como se muestra en la Figura 4, un pulso que tiene una amplitud de "1" y una polaridad de "+" o "-" aparece en cada una de cinco bandas que tienen un ancho de banda de dieciséis muestras.

La sección 122 de búsqueda minuciosa busca las posiciones que suben tres pulsos, sobre el intervalo de búsqueda completo, y codifica las posiciones y polaridades de los pulsos. En la sección 122 de búsqueda minuciosa, se realiza una búsqueda de acuerdo con las siguientes cuatro condiciones para codificación de posición precisa con una pequeña cantidad de bits de información y una pequeña cantidad de cálculos.

- 5 (1) Dos o más pulsos no deben aparecer en la misma posición. En este ejemplo, los pulsos no deben aparecer en las posiciones en las que se sube el pulso de cada banda en la sección 121 de búsqueda de intervalo. Con este ingenio, no se deben usar bits de información para representar el componente de amplitud, de modo que es posible usar bits de información eficazmente.
- 10 (2) Se buscan los pulsos en orden, de uno en uno, en un bucle abierto. Durante una búsqueda, de acuerdo con la regla de (1), no se debe someter a búsqueda las posiciones de pulso que se hayan determinado.
- (3) En una búsqueda de posición, una posición en la que mejor no debería aparecer un pulso se codifica también como una pieza de información (posición).
- (4) Dadas esas ganancias se codifican para cada banda, se buscan los pulsos evaluando distorsión de codificación con respecto a la ganancia ideal de cada banda.

15 La sección 122 de búsqueda minuciosa realiza la siguiente evaluación de coste de dos etapas para buscar un solo pulso sobre el espectro de entrada completo. En primer lugar, en la primera etapa, la sección 122 de búsqueda minuciosa evalúa el coste en cada banda y encuentra la posición y polaridad para minimizar la función de coste. A continuación, en la segunda etapa, la sección 122 de búsqueda minuciosa evalúa el coste global cada vez que se finaliza la búsqueda anterior en una banda, y almacena la posición y polaridad del pulso para minimizar el coste, como un resultado final. Se realiza esta búsqueda por banda, en orden. Además, se realiza esta búsqueda para cumplir las condiciones anteriores (1) a (4). A continuación, cuando se finaliza una búsqueda de un pulso, suponiendo la presencia de ese pulso en la posición buscada, se realiza una búsqueda del siguiente pulso. Se realiza esta búsqueda hasta que se encuentra un número de pulsos predeterminado (tres pulsos en este ejemplo), repitiendo el procesamiento anterior.

20

25

Se muestra el flujo del algoritmo de búsqueda de la sección 122 de búsqueda minuciosa en la Figura 5. La Figura 5 es un diagrama de flujo de pre-procesamiento de una búsqueda, y la Figura 6 es un diagrama de flujo de la búsqueda. Además, se muestran las partes que corresponden a las condiciones anteriores (1), (2) y (4) en el diagrama de flujo de la Figura 6.

30

Los símbolos usados en el diagrama de flujo de la Figura 5 representan los siguientes contenidos.

35 c: contador
 pf[*]: bandera de existencia/no existencia de pulso
 b: número de banda
 pos[*]: resultado de búsqueda (posición)
 n_s[*]: valor de correlación
 n_máx[*]: máximo valor de correlación
 40 n2_s[*]: valor de correlación cuadrado
 n2_máx[*]: máximo valor de correlación cuadrado
 d_s[*]: valor de potencia
 d_máx[*]: máximo valor de potencia
 45 s[*]: espectro de entrada

Los símbolos usados en el diagrama de flujo de la Figura 6 representan los siguientes contenidos.

i: número de pulso
 i0: posición de pulso
 50 cmáx: máximo valor de función de coste
 pf[*]: bandera de existencia/no existencia de pulso (0: no existencia, 1: existencia)
 ii0: posición de pulso relativa en una banda
 nom: amplitud espectral
 nom2: término numerador (potencia espectral)
 55 den: término denominador
 n_s[*]: valor relativo
 d_s[*]: valor de potencia
 s[*]: espectro de entrada
 n2_s[*]: valor de correlación cuadrado
 60 n_máx[*]: máximo valor de correlación
 n2_máx[*]: máximo valor de correlación cuadrado
 idx_máx[*]: resultado de búsqueda de cada pulso (posición) (en este punto, idx_máx[*] de 0 a 4 es equivalente a pos[b] de la Figura 3)
 fd0, fd1, fd2: memoria intermedia de almacenamiento temporal (tipo número real)
 65 id0, id1: memoria intermedia de almacenamiento temporal (tipo número entero)
 id0_s, id1_s: memoria intermedia de almacenamiento temporal (tipo número entero)

>>: desplazamiento de bit (a la derecha)
 &: "y" como una secuencia de bits

- 5 En este punto, en la búsqueda en la Figura 5 y la Figura 6, el caso donde $idx_máx[*]$ es "-1", corresponde al caso anterior de la condición (3) donde mejor no debería aparecer un pulso. El ejemplo detallado de esto es que, puesto que un espectro es suficientemente aproximado únicamente mediante el pulso buscado por banda y pulsos buscados en el intervalo completo, si se sube un pulso de la misma amplitud además, se produce un aumento proporcional de distorsión de código.
- 10 Las polaridades de los pulsos buscados corresponden a las polaridades del espectro de entrada en estas posiciones, y la sección 122 de búsqueda minuciosa codifica estas polaridades con $3 \text{ (pulsos)} \times 1 = 3 \text{ bits}$. En este punto, cuando la posición es "-1", es decir, cuando no aparece un pulso, no hay diferencia si la polaridad es "+" o "-". Sin embargo, se puede usar la polaridad para detectar errores de bit y generalmente se fija a "+" o "-".
- 15 Además, la sección 122 de búsqueda minuciosa codifica información de posición de pulso en base al número de combinaciones de posiciones de pulso. En este ejemplo, puesto que el espectro de entrada contiene ochenta muestras y ya se han encontrado cinco pulsos en cinco bandas individuales, si se tienen en cuenta también casos donde los pulsos no suben, las variaciones de posiciones se pueden representar usando diecisiete bits, de acuerdo con el cálculo de la siguiente ecuación 2.

[2]

$$\begin{aligned}
 {}_{75+1}C_3 &= (75+1) * (74+1) * (73+1) / 3 / 2 / 1 \\
 &= 70300 \\
 &< 131072 \\
 &= 2^{17} \dots \text{(Ecuación 2)}
 \end{aligned}$$

20

En este punto, de acuerdo con la regla de no permitir que aparezcan dos o más pulsos en la misma posición, es posible reducir el número de combinaciones, de modo que el efecto de esta regla se vuelve mayor cuando aumenta el número de pulsos a buscar en el intervalo completo.

25

Se describirá a continuación en detalle el método de codificación en base a las posiciones de pulsos buscadas en la sección 122 de búsqueda minuciosa.

30

(1) Se ordenan tres posiciones de pulso en base a su magnitud y se disponen en orden desde el valor numérico más bajo hasta el valor numérico más alto. En este punto, "-1" se deja como tal.

35

(2) Los números de pulso se alinean a la izquierda mediante el número de pulsos que hayan aparecido en bandas individuales, para reducir los valores numéricos de los números de pulso. Los valores numéricos calculados de esta manera se denominan como "números de posición". En este punto, "-1" se se deja como tal. Por ejemplo, con referencia a la posición de pulso de "66", cuando se proporciona cada pulso entre 0 y 15, entre 16 y 31, entre 32 y 47 y entre 48 y 64, se cambia el número de posición a "66-4=62".

40

(3) "-1" se establece al número de posición representado mediante "el máximo valor de un pulso + 1". En este caso, se ajusta y determina el orden de valores de manera que no se confunda el número de posición establecido con un número de posición en el que realmente se presente un pulso. De esta manera, se limita el número de pulso del pulso N° 0 al intervalo entre 0 y 73, se limita el número de posición del pulso N° 1 al intervalo entre el número de posición del pulso N° 0 y 74 y se limita el número de posición del pulso N° 2 al intervalo entre el número de posición del pulso N° 1 y 75, es decir, se diseña el número de posición de un pulso inferior para que no exceda el número de posición de un pulso superior.

45

(4) A continuación, de acuerdo con procesamiento de integración mostrado en la siguiente ecuación 3 para calcular un código de combinación, se integran los números de posición (i_0, i_1, i_2) para producir el código (c). Este procesamiento de integración es el procesamiento de cálculo de integrar todas las combinaciones cuando existe el orden de magnitud.

[3]

$$c = ((76-0)*(77-0)*(153-2*0)/3 + (74-0)*(75-0))/4 - ((76-i0)*(77-i0)*(153-2*i0)/3 + (74-i0)*(75-i0))/4;$$

$$c = c + (76-i0)*(77-i0)/2 - (76-i1)*(77-i1)/2;$$

$$c = c + 75-i2; \dots(\text{Ecuación 3})$$

(5) A continuación, combinando los 17 bits de este c y 3 bits para polaridad, se produce un código de 20 bits.

5 En este punto, en los números de posición anteriormente mencionados, el pulso N° 0 de "73", el pulso N° 1 de "74" y pulso N° 2 de "75" son números de posición en los que no aparecen pulsos. Por ejemplo, si hay tres números de posición (73, -1, -1), de acuerdo con la relación anteriormente mencionada entre un número de posición y el número de posición en el que no aparece un pulso, se reordenan estos números de posición a (-1, 73, -1) y hacen (73, 73, 75).

10 Por lo tanto, en el modelo donde se representa un aspecto de entrada mediante una secuencia de 8 pulsos (cinco pulsos en bandas individuales y tres pulsos en el intervalo completo) como se muestra en este ejemplo, es posible realizar codificación mediante 45 bits de información.

15 La Figura 7 ilustra un ejemplo de un espectro representado mediante los pulsos averiguados en la sección 121 de búsqueda de intervalo y la sección 122 de búsqueda minuciosa. También, en la Figura 7, los pulsos representados mediante líneas en negrita son pulsos averiguados en la sección 122 de búsqueda minuciosa.

20 La sección 112 de cuantificación de ganancia cuantifica la ganancia de cada banda. Se asignan ocho pulsos en las bandas, y la sección 112 de cuantificación de ganancia calcula las ganancias analizando la correlación entre esos pulsos y el espectro de entrada.

25 Si la sección 112 de cuantificación de ganancia calcula las ganancias ideales y a continuación realiza codificación mediante cuantificación escalar o cuantificación vectorial, en primer lugar, la sección 112 de cuantificación de ganancia calcula las ganancias ideales de acuerdo con la siguiente ecuación 4. En este punto, en la ecuación 4, g^n es la ganancia ideal de la banda "n", $s(i+16n)$ es el espectro de entrada de la banda "n", $v^n(i)$ es el vector obtenido decodificando la forma de la banda "n".

[4]

$$g^n = \frac{\sum_i s(i+16n) \times v^n(i)}{\sum_i v^n(i) \times v^n(i)} \dots(\text{Ecuación 4})$$

30 Además, la sección 112 de cuantificación de ganancia realiza codificación realizando cuantificación escalar ("SQ") de las ganancias ideales o realizando cuantificación vectorial de estas cinco ganancias juntas. En el caso de realizar cuantificación vectorial, es posible realizar codificación eficaz mediante cuantificación de predicción, VQ multi-etapa, VQ de división y así sucesivamente. En este punto, se puede escuchar la ganancia de manera perceptiva en base a una escala logarítmica, y, en consecuencia, realizando SQ o VQ después de realizar transformación logarítmica de ganancia, es posible producir buen sonido de síntesis de manera perceptiva.

40 Además, en lugar de calcular ganancias ideales, hay un método de evaluar directamente distorsión de codificación. Por ejemplo, en el caso de realizar VQ de cinco ganancias, se calcula la distorsión de codificación para minimizar la siguiente ecuación 5. En este punto, en la ecuación 5, E_k es la distorsión del k-ésimo vector de ganancia, $s(i+16n)$ es el espectro de entrada de banda "n", $g_n^{(k)}$ es el enésimo elemento del k-ésimo vector de ganancia y $v^n(i)$ es el vector de forma obtenido decodificando la forma de la banda "n".

[5]

$$E_k = \sum_n \sum_i \{ s(i+16n) - g_n^{(k)} v^n(i) \} \dots(\text{Ecuación 5})$$

A continuación, se explicará el método de decodificar tres pulsos en la sección 203 de decodificación de espectro que se averiguan mediante la búsqueda minuciosa.

5 En la sección 122 de búsqueda minuciosa de la sección 105 de codificación de espectro, se integran los números de posición (i_0, i_1, i_2) en un código usando la ecuación 3 anteriormente descrita. En la sección 203 de decodificación de espectro se realiza procesamiento inverso. Es decir, la sección 203 de decodificación de espectro calcula secuencialmente el valor de la ecuación de integración a medida que cambia cada número de posición, fija el número de posición cuando el número de posición es inferior que el valor de integración y realiza este
10 procesamiento desde el número de posición de orden inferior hasta el número de posición de orden superior uno a uno, realizando de esta manera decodificación. La Figura 8 es un diagrama de flujo que muestra el algoritmo de decodificación de la sección 203 de decodificación de espectro.
Además, en la Figura 8, cuando el código de entrada "k" de la posición integrada implica error debido a error de bit, el flujo continúa a la etapa de procesamiento de error. Por lo tanto, en este caso, se debe encontrar la posición
15 mediante procesamiento de error predeterminado.

Además, puesto que el decodificador tiene procesamiento de bucle, la cantidad de cálculos en el decodificador es mayor que en el codificador. En este punto, cada bucle es un bucle abierto, y, en consecuencia, visto desde la cantidad global de procesamiento en el códec, la cantidad de cálculos en el decodificador no es bastante grande.
20

Por lo tanto, la presente realización puede codificar frecuencias con precisión (posiciones) en las que está presente energía, de modo que es posible mejorar el rendimiento cualitativo, que es único a codificación en espectro y producir buena calidad de sonido incluso a bajas velocidades de bits.

25 Además, aunque anteriormente se ha descrito un caso con la presente realización donde se realiza codificación de ganancia después de codificación de forma, la presente invención puede proporcionar el mismo rendimiento si se realiza codificación de forma después de codificación de ganancia. Además, puede ser posible emplear un método de realización de codificación de ganancia para cada banda y a continuación normalizar el espectro mediante ganancias decodificadas y realizar codificación de forma de la presente invención.
30

Además, aunque anteriormente se ha descrito un caso de ejemplo con la presente realización donde, en la cuantificación de la forma de un espectro, la longitud del espectro es ochenta, el número de bandas es cinco, el número de pulsos a buscar por banda es uno y el número de pulsos a buscar en el intervalo completo es tres, la presente invención no depende de ninguna manera de los valores anteriores y puede producir los mismos efectos con diferentes valores numéricos.
35

Además, si el ancho de banda es suficientemente corto, se pueden codificar relativamente muchas ganancias y el número de bits de información es suficientemente grande, la presente invención puede conseguir el rendimiento anteriormente descrito únicamente realizando una búsqueda de pulso para cada banda o realizar una búsqueda de pulso en un intervalo amplio sobre una pluralidad de bandas.
40

Además, aunque se establece la condición de no subir dos pulsos en la misma posición en la realización anteriormente descrita, la presente invención puede relajar parcialmente esta condición. Por ejemplo, si se permite aparecer el pulso a buscar para cada banda y los pulsos a buscar en un intervalo amplio sobre la pluralidad de
45 bandas en las mismas posiciones, es posible eliminar pulsos de bandas individuales o permitir que aparezcan pulsos de doble amplitud. Para relajar esa condición, el requisito esencial es no almacenar la bandera de existencia/no existencia de pulso $pf[*]$ con respecto al pulso por banda. Es decir, se necesita omitir " $pf[pos[b]]=1$ " en la última etapa en la Figura 5. Como alternativa, otro método de relajar esa condición es no almacenar una bandera de existencia/no existencia de pulso tras una búsqueda de pulso en un intervalo amplio. Es decir, se necesita omitir
50 " $pf[idx_máx[i+5]]=1$ " en la última etapa en la Figura 6. En este caso, aumentan las variaciones de las posiciones. Las combinaciones no son tan simples como se muestra en la presente realización y por lo tanto es necesario clasificar casos y codificar las combinaciones de acuerdo con los casos clasificados.

Además, aunque se realiza codificación mediante pulsos para un espectro sometido a una transformación ortogonal en la presente realización, la presente invención no está limitada a esto, y es también aplicable a otros vectores. Por ejemplo, se puede aplicar la presente invención a vectores de números complejos en la FFT o DCT compleja, y se puede aplicar a una secuencia vectorial en el dominio del tiempo en la transformación Wavelet o similares. Además, la presente invención es también aplicable a una secuencia vectorial en el dominio del tiempo tal como formas de onda de excitación de CELP. Al igual que para formas de onda de excitación en CELP, se implica un filtro de
60 síntesis, y por lo tanto una función de coste implica un cálculo de matriz. En este punto, el rendimiento no es suficiente mediante una búsqueda en un bucle abierto cuando se implica un filtro, y por lo tanto se necesita realizar una búsqueda de bucle cerrado en algún grado. Cuando hay muchos pulsos, es eficaz usar una búsqueda de haz o similares para reducir la cantidad de cálculos.

65 Además, de acuerdo con la presente invención, no se limita una forma de onda a buscar a un pulso (impulso), y es igualmente posible buscar incluso otras formas de onda fijas (tales como pulso dual, onda triangular, onda finita de

respuesta de impulso, coeficiente de filtro y formas de onda fijas que cambian la forma adaptable), y producen el mismo efecto.

5 Además, aunque se ha descrito un caso con la presente realización donde se aplica la presente invención a CELP, no se limita la presente invención a esto sino que es eficaz con otros códec.

10 Además, no se puede usar únicamente una señal del habla sino también una señal de audio como la señal de acuerdo con la presente invención. Es también posible emplear una configuración en la que se aplica la presente invención a una señal residual de predicción de LPC en lugar de una señal de entrada.

15 Se puede montar el aparato de codificación y aparato de decodificación de acuerdo con la presente invención en un aparato terminal de comunicación y aparato de estación base en un sistema de comunicación móvil, de modo que es posible proporcionar un aparato terminal de comunicación, aparato de estación base y sistema de comunicación móvil que tenga el mismo efecto operacional como anteriormente.

20 Aunque se ha descrito un caso con la anterior realización como un ejemplo donde se implementa la presente invención con hardware, la presente invención se puede implementar con software. Por ejemplo, describiendo el algoritmo de acuerdo con la presente invención en un lenguaje de programación, almacenar este programa en una memoria y hacer que la sección de procesamiento de información ejecute este programa, es posible implementar la misma función que el aparato de codificación de acuerdo con la presente invención.

25 Adicionalmente, típicamente se puede implementar cada bloque de función empleado en la descripción de cada una de las realizaciones anteriormente mencionadas como un LSI constituido mediante un circuito integrado. Estos pueden ser chips individuales o parcialmente o totalmente contenidos en un solo chip.

Se adopta "LSI" en este punto pero esto se puede denominar también como "IC", "sistema LSI", "súper LSI" o "ultra LSI" dependiendo de diferentes extensiones de integración.

30 Además, no se limita el método de integración de circuito a LSI, y es también posible implementación usando circuitería dedicada o procesadores de fines generales. Después de la fabricación de LSI, también es posible utilización de un FPGA (Campo de Matriz de Puertas Programables) o un procesador reconfigurable donde se pueden reconfigurar conexiones y ajustes de celdas de circuito en un LSI.

35 Además, si la tecnología de circuitos integrados resulta que sustituye LSI como un resultado del avance de la tecnología de semiconductores u otra tecnología derivada, naturalmente es posible también llevar a cabo integración de bloque de función usando esta tecnología. También es posible la aplicación de biotecnología.

Aplicabilidad industrial

40 La presente invención es adecuada para un aparato de codificación que codifica señales del habla y señales de audio, y un aparato de decodificación que decodifica estas señales codificadas.

REIVINDICACIONES

1. Un aparato de codificación que comprende:

5 una sección de cuantificación de forma que codifica una forma de un espectro de frecuencia de audio; y una sección de cuantificación de ganancia que codifica una ganancia del espectro de frecuencia de audio, **caracterizado por que** la sección de cuantificación de forma comprende:

10 una sección de búsqueda de intervalo que busca una primera forma de onda fija en cada una de una pluralidad de bandas de frecuencia que dividen un intervalo de búsqueda de frecuencia predeterminado; y una sección de búsqueda minuciosa que busca segundas formas de onda fijas sobre una totalidad del intervalo de búsqueda de frecuencia predeterminado.

15 2. El aparato de codificación de acuerdo con la reivindicación 1, donde la sección de búsqueda minuciosa busca las segundas formas de onda fijas evaluando distorsión de codificación mediante una ganancia ideal por banda.

20 3. El aparato de codificación de acuerdo con la reivindicación 1, donde la sección de búsqueda minuciosa codifica información de posición de las segundas formas de onda fijas en base a un número de combinaciones de posiciones de las segundas formas de onda fijas.

4. El aparato de codificación de acuerdo con la reivindicación 1, donde la sección de cuantificación de ganancia calcula ganancias de la primera forma de onda fija y las segundas formas de onda fijas para cada banda.

5. Un método de codificación que comprende:

25 una etapa de cuantificación de forma de codificación de una forma de un espectro de frecuencia de audio; y una etapa de cuantificación de ganancia de codificación de una ganancia del espectro de frecuencia de audio, **caracterizado por que** la etapa de cuantificación de forma comprende:

30 una etapa de búsqueda de intervalo de búsqueda de una primera forma de onda fija en una pluralidad de bandas de frecuencia que dividen un intervalo de búsqueda de frecuencia predeterminado; y una etapa de búsqueda minuciosa de búsqueda de segundas formas de onda fijas sobre una totalidad del intervalo de búsqueda de frecuencia predeterminado.

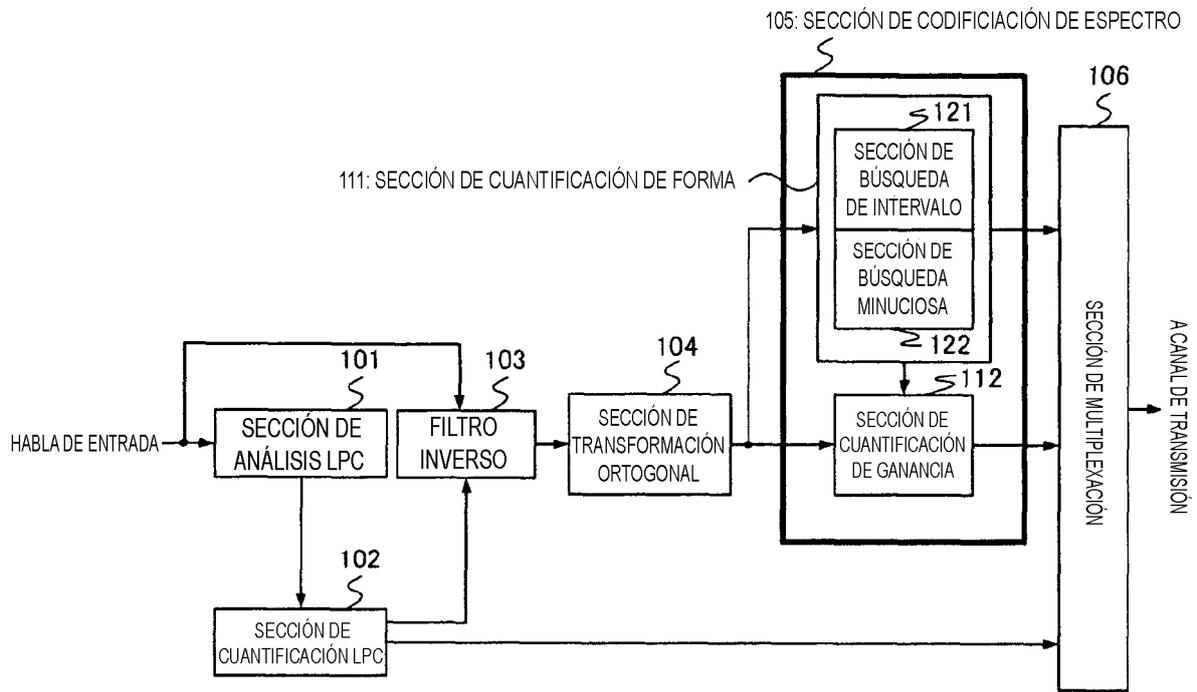


FIG.1

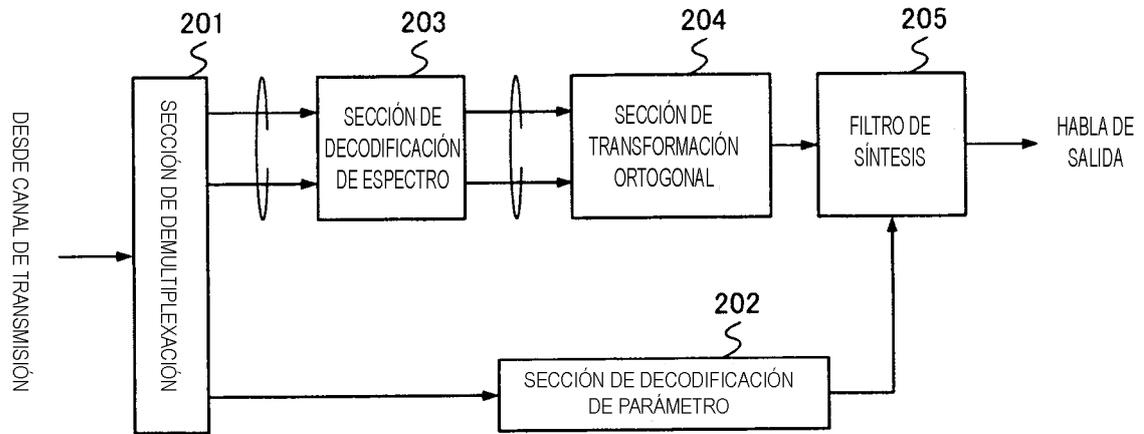


FIG.2

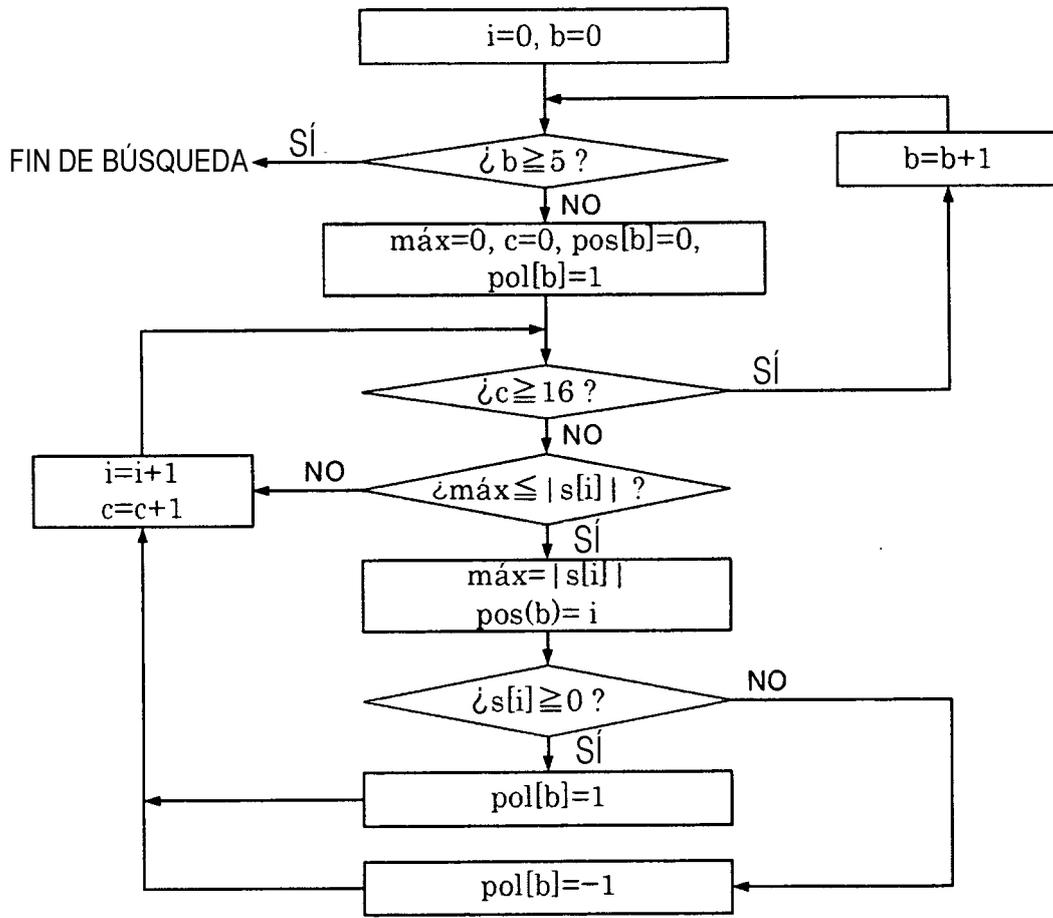


FIG.3

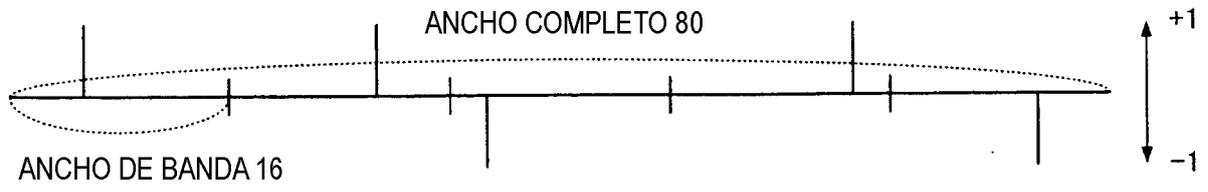


FIG.4

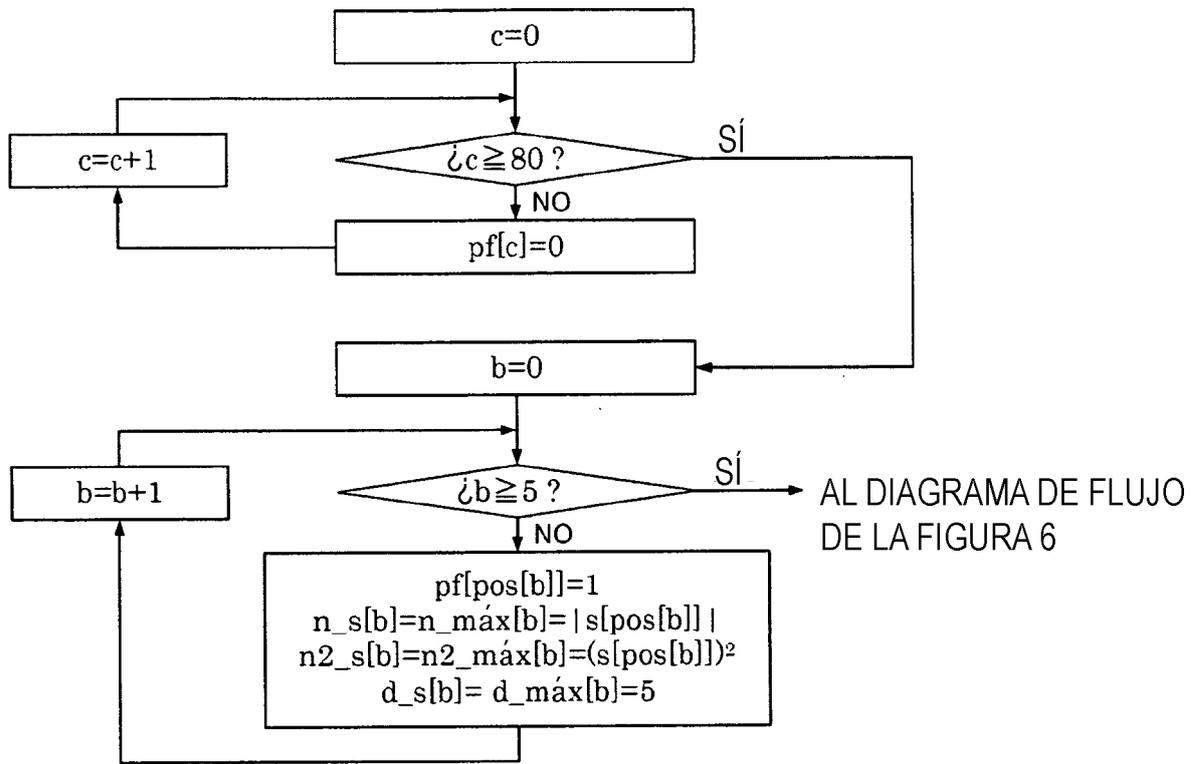


FIG.5

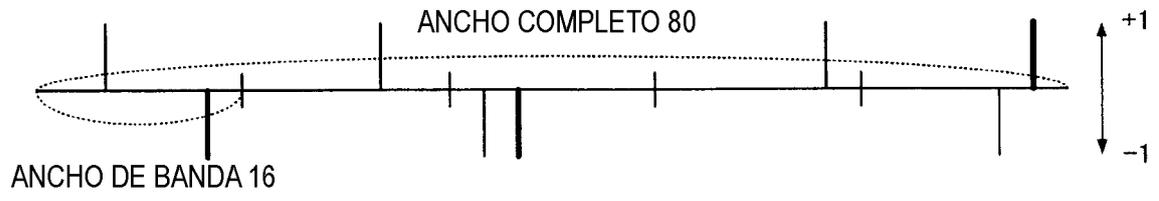


FIG.7

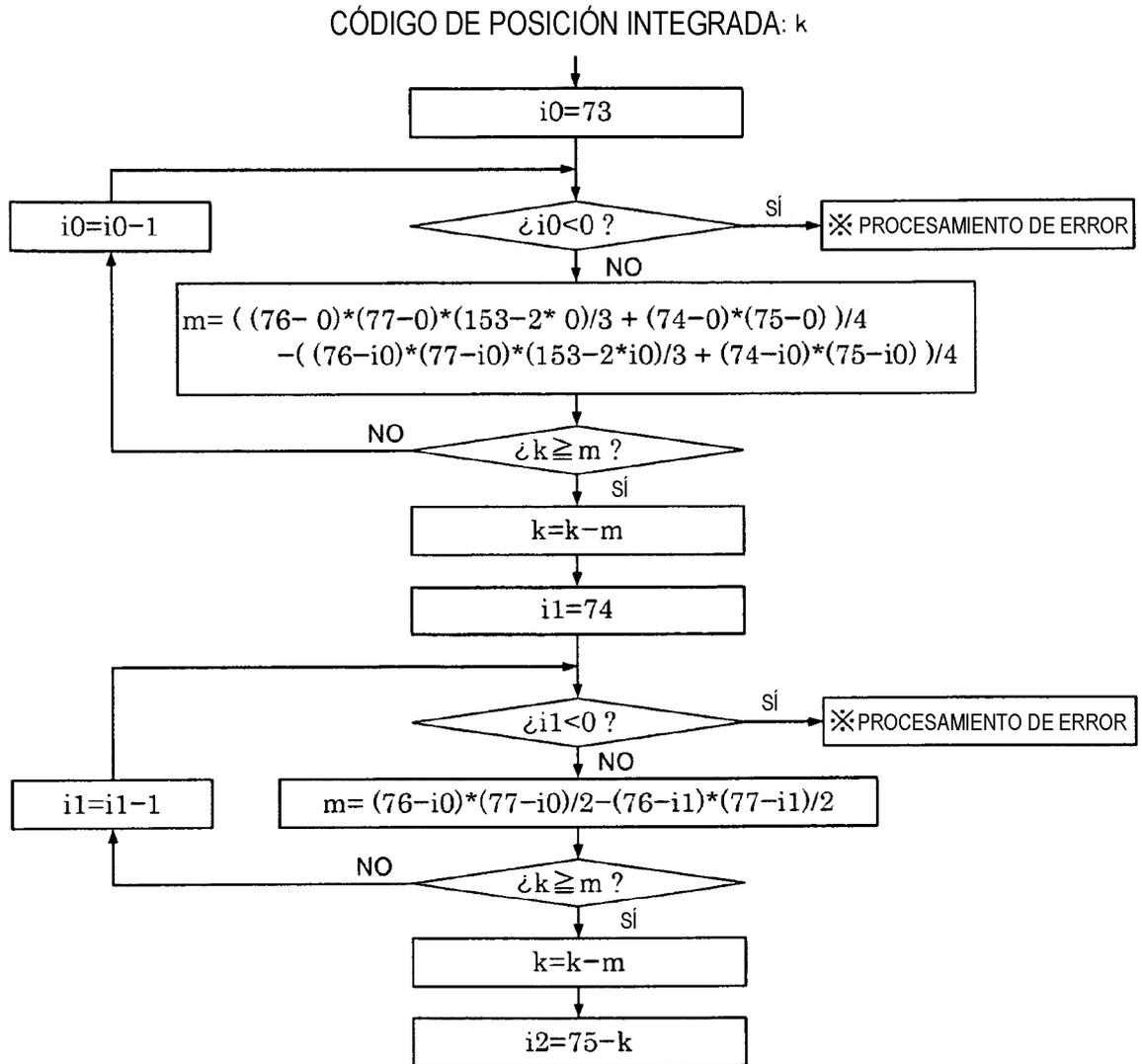


FIG.8