

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 623 551**

51 Int. Cl.:

**G10L 19/04** (2013.01)

**G10L 19/008** (2013.01)

**G10L 19/032** (2013.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.03.2006 PCT/JP2006/305871**

87 Fecha y número de publicación internacional: **05.10.2006 WO06104017**

96 Fecha de presentación y número de la solicitud europea: **23.03.2006 E 06729819 (0)**

97 Fecha y número de publicación de la concesión europea: **25.01.2017 EP 1858006**

---

54 Título: **Dispositivo de codificación de sonido y procedimiento de codificación de sonido**

---

30 Prioridad:

**25.03.2005 JP 2005088808**

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

**11.07.2017**

73 Titular/es:

**III HOLDINGS 12, LLC (100.0%)  
2711 Centerville Road, Suite 400  
Wilmington, DE 19808, US**

72 Inventor/es:

**YOSHIDA, KOJI**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 623 551 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 sonido y procedimiento de codificación de sonido

### Campo técnico

5 La presente invención se refiere a un aparato de codificación de voz y a un procedimiento de codificación de voz. Más particularmente, la presente invención se refiere a un aparato de codificación de voz y a un procedimiento de codificación de voz para voz estéreo.

### Antecedentes de la técnica

10 A medida que la transmisión de banda ancha en comunicación móvil y la comunicación de IP se ha convertido en la norma y se han diversificado los servicios en tales comunicaciones, se demanda calidad de sonido alta y comunicación de voz de fidelidad superior. Por ejemplo, de ahora en adelante, se espera que se demande la comunicación en servicio de teléfono de vídeo de manos libres, la comunicación de voz en conferencia de vídeo, la comunicación de voz multi-punto donde un número de llamantes mantienen una conversación simultáneamente en un número de diferentes localizaciones y la comunicación de voz que puede transmitir sonido de fondo sin perder alta fidelidad. En este caso, se prefiere implementar la comunicación de voz mediante una señal estéreo que tiene 15 fidelidad superior a usar señales monoaurales y que hace posible identificar las localizaciones de una pluralidad de partes llamantes. Para implementar la comunicación de voz usando una señal estéreo, la codificación de voz estéreo es esencial.

20 Además, para implementar el control de tráfico y la comunicación de multidifusión a través de una red en comunicación de datos de voz a través de una red de IP, se prefiere la codificación de voz que emplea una configuración escalable. Una configuración escalable incluye una configuración que puede decodificar datos de voz en el lado de recepción incluso a partir de datos parcialmente codificados.

Incluso cuando se codifica voz estéreo, se prefiere implementar la codificación en una configuración escalable monoaural-estéreo donde es posible seleccionar la decodificación de una señal estéreo o la decodificación de una señal monoaural usando parte de datos codificados en el lado de recepción.

25 Los procedimientos de codificación de voz que emplean una configuración escalable monoaural-estéreo incluyen, por ejemplo, predecir señales entre canales (abreviados de manera apropiada como "ch") (predecir una segunda señal de canal a partir de una primera señal de canal o predecir la primera señal de canal a partir de la segunda señal de canal) usando predicción de tono entre canales, es decir, realizar la codificación utilizando la correlación entre 2 canales (véase el Documento no de Patente 1).

30 Documento no de Patente 1: Ramprashad, S.A., "Stereocephonic CELP coding using cross channel prediction", Proc. IEEE Workshop on Speech Coding, págs. 136-138, septiembre de 2000.  
Documento de patente 1: WO 03/090208 A1 (KONINKL PHILIPS ELECTRONICS NV).

### Divulgación de la invención

#### Problemas a resolver mediante la invención

35 Sin embargo, el procedimiento de codificación de voz desvelado en el Documento no de Patente 1 y en el Documento de Patente 1 anteriores codifica por separado los parámetros de predicción inter-canal (retardo y ganancia de predicción de tono inter-canal) entre canales y por lo tanto la eficacia de codificación no es alta.

Es un objeto de la presente invención proporcionar un aparato de codificación de voz y un procedimiento de codificación de voz que posibilitan la codificación eficaz de señales estéreo.

#### Medios para resolver el problema

40 El aparato de codificación de voz de acuerdo con la presente invención emplea una configuración que incluye: una sección de análisis de parámetro de predicción que calcula una diferencia de retardo y una relación de amplitud entre una primera señal y una segunda señal como parámetros de predicción; y una sección de cuantificación que calcula parámetros de predicción cuantificados a partir de los parámetros de predicción basándose en una 45 correlación entre la diferencia de retardo y la relación de amplitud.

#### Efecto ventajoso de la invención

La presente invención posibilita la codificación eficaz de voz estéreo.

#### Breve descripción de los dibujos

50 La Figura 1 es un diagrama de bloques que muestra una configuración del aparato de codificación de voz de acuerdo con la realización 1;

la Figura 2 es un diagrama de bloques que muestra una configuración de la segunda sección de predicción de canal de acuerdo con la realización 1;

la Figura 3 es un diagrama de bloques (ejemplo de configuración 1) que muestra una configuración de la sección de cuantificación de parámetro de predicción de acuerdo con la realización 1;

5 la Figura 4 muestra un ejemplo de características de un libro de códigos de parámetro de predicción de acuerdo con la realización 1;

la Figura 5 es un diagrama de bloques (ejemplo de configuración 2) que muestra una configuración de la sección de cuantificación de parámetro de predicción de acuerdo con la realización 1;

10 la Figura 6 muestra características que indican un ejemplo de la función usada en la sección de estimación de relación de amplitud de acuerdo con la realización 1;

la Figura 7 es un diagrama de bloques (ejemplo de configuración 3) que muestra una configuración de la sección de cuantificación de parámetro de predicción de acuerdo con la realización 2;

la Figura 8 muestra características que indican un ejemplo de la función usada en la sección de cálculo de distorsión de acuerdo con la realización 2;

15 la Figura 9 es un diagrama de bloques (ejemplo de configuración 4) que muestra una configuración de la sección de cuantificación de parámetro de predicción de acuerdo con la realización 2;

la Figura 10 muestra características que indican un ejemplo de las funciones usadas en la sección de corrección de relación de amplitud y la sección de estimación de relación de amplitud de acuerdo con la realización 2; y

20 la Figura 11 es un diagrama de bloques (ejemplo de configuración 5) que muestra una configuración de la sección de cuantificación de parámetro de predicción de acuerdo con la realización 2.

### **Mejor modo para llevar a cabo la invención**

Las realizaciones de la presente invención se describirán en detalle con referencia a los dibujos adjuntos.

(Realización 1)

25 La Figura 1 muestra una configuración del aparato de codificación de voz de acuerdo con la presente realización. El aparato 10 de codificación de voz mostrado en la Figura 1 tiene la primera sección 11 de codificación de canal, primera sección 12 de decodificación de canal, segunda sección 13 de predicción de canal, restador 14 y segunda sección 15 de codificación residual de predicción de canal. En la siguiente descripción, se proporciona una descripción suponiendo la operación en unidades de trama.

30 La primera sección 11 de codificación de canal codifica una primera señal de voz de canal  $s_{ch1}(n)$  (donde  $n$  es entre 0 y  $NF-1$  y  $NF$  es la longitud de trama) de una señal de estéreo de entrada, y emite datos codificados (primeros datos codificados de canal) para la primera señal de voz de canal a la primera sección 12 de decodificación de canal. Además, estos primeros datos codificados de canal se multiplexan con segundos datos codificados de parámetro de predicción de canal y segundos datos codificados de canal, y se transmiten a un aparato de decodificación de voz (no mostrado).

35 La primera sección 12 de decodificación de canal genera una primera señal decodificada de canal a partir de los primeros datos codificados de canal, y emite el resultado a segunda sección 13 de predicción de canal.

40 La segunda sección 13 de predicción de canal calcula segundos parámetros de predicción de canal a partir de la primera señal decodificada de canal y una segunda señal de voz de canal  $s_{ch2}(n)$  (donde  $n$  es entre 0 y  $NF-1$  y  $NF$  es la longitud de trama) de la señal estéreo de entrada, y emite segundos datos codificados de parámetro de predicción de canal, es decir los segundos parámetros de predicción de canal sometidos a codificación. Estos segundos datos codificados de parámetros de predicción se multiplexan con otros datos codificados, y se transmiten al aparato de decodificación de voz (no mostrado). La segunda sección 13 de predicción de canal sintetiza una segunda señal predicha de canal  $sp_{ch2}(n)$  a partir de la primera señal decodificada de canal y la segunda señal de voz de canal, y emite la segunda señal predicha de canal al restador 14. La segunda sección 13 de predicción de canal se describirá en detalle más adelante.

45 El restador 14 calcula la diferencia entre la segunda señal de voz de canal  $s_{ch2}(n)$  y la segunda señal predicha de canal  $sp_{ch2}(n)$ , es decir, la señal (segunda señal residual de predicción de canal) del componente residual de la segunda señal predicha de canal con respecto a la segunda señal de voz de canal, y emite la diferencia a segunda sección 15 de codificación residual de predicción de canal.

50 La segunda sección 15 de codificación residual de predicción de canal codifica la segunda señal residual de predicción de canal y emite segundos datos codificados de canal. Estos segundos datos codificados de canal se multiplexan con otros datos codificados y se transmiten al aparato de decodificación de voz.

55 A continuación, la segunda sección 13 de predicción de canal se describirá en detalle. La Figura 2 muestra la configuración de la segunda sección 13 de predicción de canal. Como se muestra en la Figura 2, la segunda sección 13 de predicción de canal tiene la sección 21 de análisis de parámetro de predicción, sección 22 de cuantificación de parámetro de predicción y sección 23 de predicción de señal.

Basándose en la correlación entre las señales de canal de la señal estéreo, la segunda sección 13 de predicción de canal predice la segunda señal de voz de canal a partir de la primera señal de voz de canal usando parámetros basándose en la diferencia de retardo D y la relación de amplitud g de la segunda señal de voz de canal con respecto a la primera señal de voz de canal.

5 A partir de la primera señal decodificada de canal y la segunda señal de voz de canal, la sección 21 de análisis de parámetro de predicción calcula la diferencia de retardo D y la relación de amplitud g de la segunda señal de voz de canal con respecto a la primera señal de voz de canal como parámetros de predicción inter-canal y emite los parámetros de predicción inter-canal a la sección 22 de cuantificación de parámetro de predicción.

10 La sección 22 de cuantificación de parámetro de predicción cuantifica los parámetros de predicción introducidos (diferencia de retardo D y relación de amplitud g) y emite parámetros de predicción cuantificados y segundos datos codificados de parámetro de predicción de canal. Los parámetros de predicción cuantificados se introducen a la sección 23 de predicción de señal. La sección 22 de cuantificación de parámetro de predicción se describirá en detalle más adelante.

15 La sección 23 de predicción de señal predice la segunda señal de canal usando la primera señal decodificada de canal y los parámetros de predicción cuantificados, y emite la señal predicha. La segunda señal predicha de canal  $sp\_ch2(n)$  (donde n es entre 0 y NF-1 y NF es la longitud de trama) predicha en la sección 23 de predicción de señal se expresa mediante la siguiente ecuación 1 usando la primera señal decodificada de canal  $sd\_ch1(n)$ .

$$sp\_ch2(n) = g \cdot sd\_ch1(n - D) \dots \text{(Ecuación 1)}$$

20 Además, la sección 21 de análisis de parámetro de predicción calcula los parámetros de predicción (diferencia de retardo D y relación de amplitud g) que minimizan la distorsión "Dist" expresada mediante la ecuación 2, es decir, la distorsión Dist entre la segunda señal de voz de canal  $s\_ch2(n)$  y la segunda señal predicha de canal  $sp\_ch2(n)$ . La sección 21 de análisis de parámetro de predicción puede calcular como los parámetros de predicción, la diferencia de retardo D que maximiza la correlación entre la segunda señal de voz de canal y la primera señal decodificada de canal y relación de amplitud media g en unidades de trama.

$$Dist = \sum_{n=0}^{NF-1} \{ s\_ch2(n) - sp\_ch2(n) \}^2 \dots \text{(Ecuación 2)}$$

25 A continuación, la sección 22 de cuantificación de parámetro de predicción se describirá en detalle.

30 Entre la diferencia de retardo D y la relación de amplitud g calculadas en la sección 21 de análisis de parámetro de predicción, hay una relación (correlación) que resulta de las características espaciales (por ejemplo, distancia) del origen de una señal al punto de recepción. Es decir, hay una relación que cuando la diferencia de retardo D (>0) se hace mayor (mayor en la dirección positiva (dirección de retardo)), la relación de amplitud g se hace menor (<1,0), y, por otra parte, cuando la diferencia de retardo D (<0) se hace menor (mayor en la dirección negativa (dirección hacia delante)), la relación de amplitud g (>1,0) se hace mayor. Utilizando esta relación, la sección 22 de cuantificación de parámetro de predicción usa menos bits de cuantificación de modo que se realiza igual distorsión de cuantificación, para codificar eficazmente los parámetros de predicción inter-canal (diferencia de retardo D y relación de amplitud g).

35 La configuración de la sección 22 de cuantificación de parámetro de predicción de acuerdo con la presente realización es como se muestra en <ejemplo de configuración 1> de la Figura 3 o <ejemplo de configuración 2> de la Figura 5.

**<Ejemplo de configuración 1>**

40 En ejemplo de configuración 1 (Figura 3), la diferencia de retardo D y la relación de amplitud g se expresan mediante un vector bidimensional, y la cuantificación de vector se realiza en el vector bidimensional. La Figura 4 muestra características de vectores de código mostrados mediante el símbolo circular ("s") como el vector bidimensional.

45 En la Figura 3, la sección 31 de cálculo de distorsión calcula la distorsión entre los parámetros de predicción expresados mediante el vector bidimensional (D y g) formado con la diferencia de retardo D y la relación de amplitud g, y vectores de código del libro de códigos 33 de parámetro de predicción.

La sección 32 de búsqueda de distorsión mínima busca el vector de código que tiene la mínima distorsión de todos los vectores de código, transmite el resultado de búsqueda al libro de códigos 33 de parámetro de predicción y emite el índice que corresponde al vector de código como los segundos datos codificados de parámetro de predicción de canal.

50

Basándose en el resultado de búsqueda, el libro de códigos 33 de parámetro de predicción emite el vector de código que tiene la mínima distorsión como parámetros de predicción cuantificados.

En este punto, si el k-ésimo del libro de códigos 33 de parámetro de predicción es (Dc(k), gc(k)) (donde k es entre 0 y Ncb-1 y Ncb es el tamaño de libro de códigos), la distorsión Dst(k) del k-ésimo vector de código calculada mediante la sección 31 de cálculo de distorsión se expresa mediante la siguiente ecuación 3. En la ecuación 3, wd y wg son constantes de ponderación para ajustar la ponderación entre la distorsión de cuantificación de la diferencia de retardo y la distorsión de cuantificación de la relación de amplitud tras el cálculo de distorsión.

$$D_{st}(k) = w_d \cdot (D - D_c(k))^2 + w_g \cdot (g - g_c(k))^2 \dots \text{(Ecuación 3)}$$

El libro de códigos 33 de parámetro de predicción se prepara con antelación mediante aprendizaje, basándose en la correspondencia entre la diferencia de retardo D y la relación de amplitud g. Además, se obtiene una pluralidad de datos (datos de aprendizaje) que indican la correspondencia entre la diferencia de retardo D y la relación de amplitud g con antelación a partir de una señal de voz estéreo para uso de aprendizaje. Existe la relación anterior entre los parámetros de predicción de la diferencia de retardo y la relación de amplitud y se obtienen los datos de aprendizaje basándose en esta relación. Por lo tanto, en el libro de códigos 33 de parámetro de predicción obtenido mediante aprendizaje, como se muestra en la Figura 4, la distribución de vectores de código alrededor del centro del símbolo circular donde la diferencia de retardo D y la relación de amplitud g es (D, g) = (0, 1,0) en proporción negativa es densa y la otra distribución es dispersa. Usando un libro de códigos de parámetro de predicción que tiene características como se muestra en la Figura 4, es posible realizar pocos errores de cuantificación de parámetros predichos que tienen lugar frecuentemente entre los parámetros de predicción que indican la correspondencia entre diferencias de retardo y relaciones de amplitud. Como resultado, es posible mejorar la eficacia de cuantificación.

**<Ejemplo de configuración 2>**

En el ejemplo de configuración 2 (Figura 5), la función para estimar la amplitud g a partir de la diferencia de retardo D se determina con antelación, y, después de que se cuantifica la diferencia de retardo D, se cuantifica la predicción residual de la relación de amplitud estimada a partir del valor de cuantificación usando la función.

En la Figura 5, la sección 51 de cuantificación de diferencia de retardo cuantifica diferencia de retardo D de los parámetros predichos, emite esta diferencia de retardo cuantificada Dq a la sección 52 de estimación de relación de amplitud y emite el parámetro de predicción cuantificado. La sección 51 de cuantificación de diferencia de retardo emite el índice de diferencia de retardo cuantificado obtenido cuantificando la diferencia de retardo D como segundos datos codificados de parámetro de predicción de canal.

La sección 52 de estimación de relación de amplitud obtiene el valor de estimación (relación de amplitud estimada) gp de la relación de amplitud a partir de la diferencia de retardo cuantificada Dq, y emite el resultado a la sección 53 de cuantificación residual de estimación de relación de amplitud. La estimación de relación de amplitud usa una función preparada con antelación para estimar la amplitud a partir de la diferencia de retardo cuantificada. Esta función se prepara con antelación mediante aprendizaje basándose en la correspondencia entre la diferencia de retardo cuantificada Dq y la relación de amplitud estimada gp. Además, se obtiene una pluralidad de datos que indican la correspondencia entre la diferencia de retardo cuantificada Dq y la relación de amplitud estimada gp a partir de señales estéreo para uso de aprendizaje.

La sección 53 de cuantificación residual de estimación de relación de amplitud calcula la estimación residual δg de la relación de amplitud g con respecto a la relación de amplitud estimada gp usando la ecuación 4.

$$\delta g = g - g_p \dots \text{(Ecuación 4)}$$

La sección 53 de cuantificación residual de estimación de relación de amplitud cuantifica la estimación residual δg obtenida a partir de la ecuación 4, y emite la estimación cuantificada residual como un parámetro de predicción cuantificado. La sección 53 de cuantificación residual de estimación de relación de amplitud emite el índice residual de estimación cuantificada obtenido cuantificando la estimación residual δg como segundos datos codificados de parámetro de predicción de canal.

La Figura 6 muestra un ejemplo de la función usada en la sección 52 de estimación de relación de amplitud. Los parámetros de predicción introducidos (D,g) se indican como un vector bidimensional mediante símbolos circulares en el plano de coordenadas mostrado en la Figura 6. Como se muestra en la Figura 6, la función 61 para estimar la relación de amplitud a partir de la diferencia de retardo es en proporción negativa de manera que la función 61 para el punto (D,g) = (0,1,0) o su cercanía. Además, la sección 52 de estimación de relación de amplitud obtiene la relación de amplitud estimada gp a partir de la diferencia de retardo cuantificada Dq usando esta función. Además, la sección 53 de cuantificación residual de estimación de relación de amplitud calcula la estimación residual δg de la relación de amplitud g del parámetro de predicción de entrada con respecto a la relación de amplitud estimada gp, y cuantifica esta estimación residual δg. De esta manera, cuantificando la estimación residual, es posible reducir

adicionalmente el error de cuantificación que cuantificar directamente la relación de amplitud, y, como resultado, mejorar la eficacia de cuantificación.

5 Se ha descrito una configuración en la descripción anterior donde la relación de amplitud estimada  $g_p$  se calcula a partir de la diferencia de retardo cuantificada  $D_q$  usando la función para estimar la relación de amplitud a partir de la diferencia de retardo cuantificada, y se cuantifica la estimación residual  $\delta g$  de la relación de amplitud de entrada  $g$  con respecto a esta relación de amplitud estimada  $g_p$ . Sin embargo, puede ser posible una configuración que cuantifica la relación de amplitud de entrada  $g$ , calcula la diferencia de retardo estimada  $D_p$  a partir de la relación de amplitud cuantificada  $g_q$  usando la función para estimar la diferencia de retardo a partir de la relación de amplitud cuantificada y cuantifica la estimación residual  $\delta D$  de la diferencia de retardo de entrada  $D$  con respecto a la diferencia de retardo estimada  $D_p$ .

(Realización 2)

15 La configuración de la sección 22 de cuantificación de parámetro de predicción (Figura 2, Figura 3 y Figura 5) del aparato de codificación de voz de acuerdo con la presente realización se diferencia de la sección 22 de cuantificación de parámetro de predicción de la realización 1. Al cuantificar parámetros de predicción en la presente realización, se cuantifica una diferencia de retardo y una relación de amplitud de manera que los errores de cuantificación de parámetros tanto de la diferencia de retardo como la relación de amplitud se anulan perceptualmente entre sí. Es decir, cuando tiene lugar un error de cuantificación de una diferencia de retardo en la dirección positiva, la cuantificación se lleva a cabo de manera que el error de cuantificación de una relación de amplitud se hace mayor. Por otra parte, cuando tiene lugar el error de cuantificación de una diferencia de retardo en la dirección negativa, la cuantificación se lleva a cabo de manera que el error de cuantificación de una relación de amplitud se hace más pequeño.

20 En este punto, las características perceptuales humanas hacen posible ajustar la diferencia de retardo y la relación de amplitud mutuamente para conseguir la localización del mismo sonido estéreo. Es decir, cuando la diferencia de retardo se hace más significativa que la diferencia de retardo real, puede conseguirse localización igual aumentando la relación de amplitud. En la presente realización, basándose en la característica perceptual anterior, la diferencia de retardo y la relación de amplitud se cuantifican ajustando el error de cuantificación de la diferencia de retardo y el error de cuantificación de la relación de amplitud, de manera que la localización del sonido estéreo no cambia. Como resultado, es posible la codificación eficaz de parámetros predichos. Es decir, es posible realizar igual calidad de sonido a tasas de bits de codificación inferiores y calidad de sonido superior a tasas de bits de codificación iguales.

30 La configuración de la sección 22 de cuantificación de parámetro de predicción de acuerdo con la presente realización es como se muestra en <ejemplo de configuración 3> de la Figura 7 o <ejemplo de configuración 4> de la Figura 9.

**<Ejemplo de configuración 3>**

35 El cálculo de la distorsión en ejemplo de configuración 3 (Figura 7) es diferente de la configuración 1 (Figura 3). En la Figura 7, se asignan los mismos componentes que en la Figura 3 a los mismos números de referencia y se omitirá la descripción de los mismos.

En la Figura 7, la sección 71 de cálculo de distorsión calcula la distorsión entre los parámetros de predicción expresados mediante el vector bidimensional  $(D, g)$  formado con la diferencia de retardo  $D$  y la relación de amplitud  $g$ , y los vectores de código del libro de códigos 33 de parámetro de predicción.

40 El  $k$ -ésimo vector del libro de códigos 33 de parámetro de predicción se establece como  $(D_c(k), g_c(k))$  (donde  $k$  es entre 0 y  $N_{cb}$  y  $N_{cb}$  es el tamaño de libro de códigos). La sección 71 de cálculo de distorsión mueve el vector bidimensional  $(D, g)$  para los parámetros de predicción introducidos al punto equivalente perceptualmente más cercano  $(D_c'(k), g_c'(k))$  a los vectores de código  $(D_c(k), g_c(k))$ , y calcula la distorsión  $D_{st}(k)$  de acuerdo con la ecuación 5. En la ecuación 5,  $w_d$  y  $w_g$  son constantes de ponderación para ajustar la ponderación entre la distorsión de cuantificación de la diferencia de retardo y la distorsión de cuantificación de la relación de amplitud tras el cálculo de distorsión.

$$D_{st}(k) = w_d \cdot ((D_c'(k) - D_c(k))^2 + w_g \cdot (g_c'(k) - g_c(k))^2) \dots \text{(Ecuación 5)}$$

50 Como se muestra en la Figura 8, el punto equivalente perceptualmente más cercano a los vectores de código  $(D_c(k), g_c(k))$  corresponde al punto al cual una perpendicular va desde los vectores de código verticalmente hasta la función 81 que tiene el conjunto de localización de sonido estéreo perceptualmente equivalente al vector de parámetro de predicción de entrada  $(D, g)$ . Esta función 81 coloca la diferencia de retardo  $D$  y la relación de amplitud  $g$  en proporción entre sí en la dirección positiva. Es decir, esta función 81 tiene una característica perceptual de conseguir la localización perceptualmente equivalente haciendo la relación de amplitud mayor cuando la diferencia de retardo se hace mayor y haciendo la relación de amplitud menor cuando la diferencia de retardo se hace menor.

Cuando el vector de parámetro de predicción de entrada (D,g) se mueve al punto equivalente perceptualmente más cercano a los vectores de código (Dc(k),gc(k)) en la función 81, se impone una penalización haciendo la distorsión mayor con respecto al movimiento al punto a través del más lejano de la distancia predeterminada.

5 Cuando se lleva a cabo la cuantificación de vector usando la distorsión obtenida de esta manera, por ejemplo, en la Figura 8, en lugar del vector de código A (distorsión de cuantificación A) que es el más cercano al vector de parámetro de predicción de entrada o localización de sonido estéreo del vector de código B (distorsión de cuantificación B), vector de código C (distorsión de cuantificación C) que es perceptualmente más cercano al vector de parámetro de predicción de entrada se hace el valor de cuantificación. Por lo tanto, es posible llevar a cabo la cuantificación con menor distorsión perceptual.

10 **<Ejemplo de configuración 4>**

El ejemplo de configuración 4 (Figura 9) se diferencia del ejemplo de configuración 2 (Figura 5) en cuantificar la estimación residual de la relación de amplitud que está corregida a un valor perceptualmente equivalente (relación de amplitud corregida) teniendo en cuenta el error de cuantificación de la diferencia de retardo. En la Figura 9, se asignan los mismos componentes que en la Figura 5 a los mismos números de referencia y se omitirá una descripción de los mismos.

15 En la Figura 9, la sección 51 de cuantificación de diferencia de retardo emite la diferencia de retardo cuantificada Dq a la sección 91 de corrección de relación de amplitud.

La sección 91 de corrección de relación de amplitud corrige la relación de amplitud g a un valor perceptualmente equivalente teniendo en cuenta el error de cuantificación de la diferencia de retardo, y obtiene la relación de amplitud corregida g'. Esta relación de amplitud corregida g' se introduce a la sección 92 de cuantificación residual de estimación de relación de amplitud.

La sección 92 de cuantificación residual de estimación de relación de amplitud obtiene la estimación residual  $\delta g$  de la relación de amplitud corregida g' con respecto a relación de amplitud estimada gp de acuerdo con la ecuación 6.

$$\delta g = g' - g_p \dots \text{(Ecuación 6)}$$

25 La sección 92 de cuantificación residual de estimación de relación de amplitud cuantifica el residuo estimado  $\delta g$  obtenido de acuerdo con la ecuación 6, y emite la estimación cuantificada residual como los parámetros de predicción cuantificados. La sección 92 de cuantificación residual de estimación de relación de amplitud emite el índice residual de estimación cuantificado obtenido cuantificando la estimación residual  $\delta g$  como los segundos datos codificados de parámetro de predicción de canal.

30 La Figura 10 muestra ejemplos de las funciones usadas en la sección 91 de corrección de relación de amplitud y la sección 52 de estimación de relación de amplitud. La función 81 usada en la sección 91 de corrección de relación de amplitud es la misma que la función 81 usada en el ejemplo de configuración 3. La función 61 usada en la sección 52 de estimación de relación de amplitud es la misma que la función 61 usada en ejemplo de configuración 2.

35 Como se ha descrito anteriormente, la función 81 coloca la diferencia de retardo D y la relación de amplitud g en proporción en la dirección positiva. La sección 91 de corrección de relación de amplitud usa esta función 81 y obtiene la relación de amplitud corregida g' que es perceptualmente equivalente a la relación de amplitud g teniendo en cuenta el error de cuantificación de la diferencia de retardo, de la diferencia de retardo cuantificada. Como se ha descrito anteriormente, la función 61 es una función que incluye un punto (D,g) = (0,1,0) o sus cercanías y tiene proporción inversa. La sección 52 de estimación de relación de amplitud usa esta función 61 y obtiene la relación de amplitud estimada gp de la diferencia de retardo cuantificada Dq. La sección 92 de cuantificación residual de estimación de relación de amplitud calcula la estimación residual  $\delta g$  de la relación de amplitud corregida g' con respecto a la relación de amplitud estimada gp, y cuantifica esta estimación residual  $\delta g$ .

40 Por lo tanto, la estimación residual se calcula a partir de la relación de amplitud que se corrige a un valor perceptualmente equivalente (relación de amplitud corregida) teniendo en cuenta el error de cuantificación de la diferencia de retardo, y se cuantifica la estimación residual, de modo que es posible llevar a cabo la cuantificación con perceptualmente pequeña distorsión y pequeño error de cuantificación.

**<Ejemplo de configuración 5>**

50 Cuando la diferencia de retardo D y la relación de amplitud g se cuantifican por separado, las características perceptuales con respecto a la diferencia de retardo y la relación de amplitud pueden usarse como en la presente realización. La Figura 11 muestra la configuración de la sección 22 de cuantificación de parámetro de predicción en este caso. En la Figura 11, se asignan a los mismos componentes que en el ejemplo de configuración 4 (Figura 9) los mismos números de referencia.

En la Figura 11, como en el ejemplo de configuración 4, la sección 91 de corrección de relación de amplitud corrige la relación de amplitud g a un valor perceptualmente equivalente teniendo en cuenta el error de cuantificación de la

diferencia de retardo, y obtiene la relación de amplitud corregida  $g'$ . Esta relación de amplitud corregida  $g'$  se introduce a la sección 1101 de cuantificación de relación de amplitud.

5 La sección 1101 de cuantificación de relación de amplitud cuantifica la relación de amplitud corregida  $g'$  y emite la relación de amplitud cuantificada como un parámetro de predicción cuantificado. Además, la sección 1101 de cuantificación de relación de amplitud emite el índice de relación de amplitud cuantificado obtenido cuantificando la relación de amplitud corregida  $g'$  como segundos datos codificados de parámetro de predicción de canal.

10 En las realizaciones anteriores, los parámetros de predicción (diferencia de retardo  $D$  y relación de amplitud  $g$ ) se describen como valores escalares (valores unidimensionales). Sin embargo, una pluralidad de parámetros de predicción obtenidos a través de una pluralidad de unidades de tiempo (tramas) pueden expresarse mediante vectores de dos o más dimensiones, y a continuación someterse a la cuantificación anterior.

15 Además, las realizaciones anteriores pueden aplicarse a un aparato de codificación de voz que tiene una configuración de monoaural a estéreo escalable. En este caso, en una capa de núcleo monoaural, una señal monoaural se genera a partir de una señal de estéreo de entrada (señales de voz del primer canal y segundo canal) y se codifica. Además, en una capa de mejora estéreo, la señal de voz del primer canal (o segundo canal) se predice a partir de la señal monoaural usando predicción inter-canal, y se codifica una señal residual de predicción de esta señal predicha y la señal de voz del primer canal (o segundo canal). Además, puede usarse codificación CELP en la codificación en la capa de núcleo monoaural y la capa de mejora estéreo. En este caso, en la capa de mejora estéreo, la señal de excitación monoaural obtenida en la capa de núcleo monoaural se somete a predicción inter-canal, y la predicción residual se codifica mediante codificación por excitación CELP. En una configuración  
20 escalable, los parámetros de predicción inter-canal se refieren a parámetros para predicción del primer canal (o segundo canal) a partir de la señal monoaural.

25 Cuando se aplican las realizaciones anteriores al aparato de codificación de voz que tiene las configuraciones de monoaural a estéreo escalables, las diferencias de retardo ( $Dm1$  y  $Dm2$ ) y las relaciones de amplitud ( $gm1$  y  $gm2$ ) del primer canal y la segunda señal de voz de canal de la señal monoaural pueden cuantificarse de manera colectiva como en la realización 2. En este caso, existe la correlación entre diferencias de retardo (entre  $Dm1$  y  $Dm2$ ) y las relaciones de amplitud (entre  $gm1$  y  $gm2$ ) de los canales, de modo que es posible mejorar la eficacia de codificación de parámetros predichos en la configuración de monoaural a estéreo escalable utilizando la correlación.

30 El aparato de codificación de voz y el aparato de decodificación de voz de las realizaciones anteriores pueden montarse también en un aparato de comunicación de radio tal como el aparato de estación móvil de comunicación inalámbrica y el aparato de estación base de comunicación de radio usados en los sistemas de comunicación móviles.

También, se han descrito casos con las realizaciones anteriores donde la presente invención está configurada mediante hardware. Sin embargo, la presente invención puede realizarse también mediante software.

35 Cada bloque de función empleado en la descripción de cada una de las realizaciones anteriormente mencionadas puede implementarse típicamente como un LSI constituido por un circuito integrado. Estos pueden ser chips individuales o estar parcial o totalmente contenidos en un único chip.

Se adopta "LSI" en este punto pero puede denominarse también como "CI", sistema LSI, "súper LSI" o "ultra LSI" dependiendo de diferentes puntos de integración.

40 Además, el procedimiento de integración de circuito no está limitado a LSI, y es también posible la implementación usando circuitería especializada o procesadores de fin general. Después de la fabricación de LSI, es también posible la utilización de un FPGA (Campo de Matriz de Puertas Programables) o un procesador reconfigurable donde las conexiones y ajustes de las células de circuitos en un LSI pueden reconfigurarse.

45 Además, si la tecnología de circuitos integrados reemplaza a LSI como resultado del avance de la tecnología de semiconductores u otra tecnología derivada, es también evidentemente posible llevar a cabo la integración del bloque de función usando esta tecnología. La aplicación de biotecnología es también posible.

### **Aplicabilidad industrial**

La presente invención es aplicable a usos en el aparato de comunicación de los sistemas de comunicación móvil y sistemas de comunicación de paquetes que emplean el protocolo de internet.



**REIVINDICACIONES**

1. Un aparato de codificación de voz que comprende:
  - una sección de análisis de parámetro de predicción que calcula una diferencia de retardo y una relación de amplitud entre una primera señal de canal y una segunda señal de canal como parámetros de predicción; y
  - 5 una sección de cuantificación que calcula parámetros de predicción cuantificados a partir de los parámetros de predicción basándose en una correlación entre la diferencia de retardo y la relación de amplitud.
2. El aparato de codificación de voz de acuerdo con la reivindicación 1, en el que la sección de cuantificación calcula los parámetros de predicción cuantificados cuantificando un residuo de la relación de amplitud con respecto a una relación de amplitud estimada a partir de la diferencia de retardo.
- 10 3. El aparato de codificación de voz de acuerdo con la reivindicación 1, en el que la sección de cuantificación calcula los parámetros de predicción cuantificados cuantificando un residuo de la diferencia de retardo con respecto a una diferencia de retardo estimada a partir de la relación de amplitud.
4. El aparato de codificación de voz de acuerdo con la reivindicación 1, en el que la sección de cuantificación calcula los parámetros de predicción cuantificados llevando a cabo la cuantificación de manera que tiene lugar un error de cuantificación de la diferencia de retardo y un error de cuantificación de la relación de amplitud en una dirección donde el error de cuantificación de la diferencia de retardo y el error de cuantificación de la relación de amplitud perceptualmente se anulan entre sí.
- 15 5. El aparato de codificación de voz de acuerdo con la reivindicación 1, en el que la sección de cuantificación calcula los parámetros de predicción cuantificados usando un vector bidimensional comprendido por la diferencia de retardo y la relación de amplitud.
- 20 6. Un aparato de estación móvil de comunicación inalámbrica que comprende el aparato de codificación de voz de acuerdo con la reivindicación 1.
7. Un aparato de estación base de comunicación inalámbrica que comprende el aparato de codificación de voz de acuerdo con la reivindicación 1.
- 25 8. Un procedimiento de codificación de voz que comprende las etapas de:
  - calcular una diferencia de retardo y una relación de amplitud entre una primera señal de canal y una segunda señal de canal como un parámetro de predicción; y
  - calcular parámetros de predicción cuantificados a partir de los parámetros de predicción basándose en una correlación entre la diferencia de retardo y la relación de amplitud.
- 30

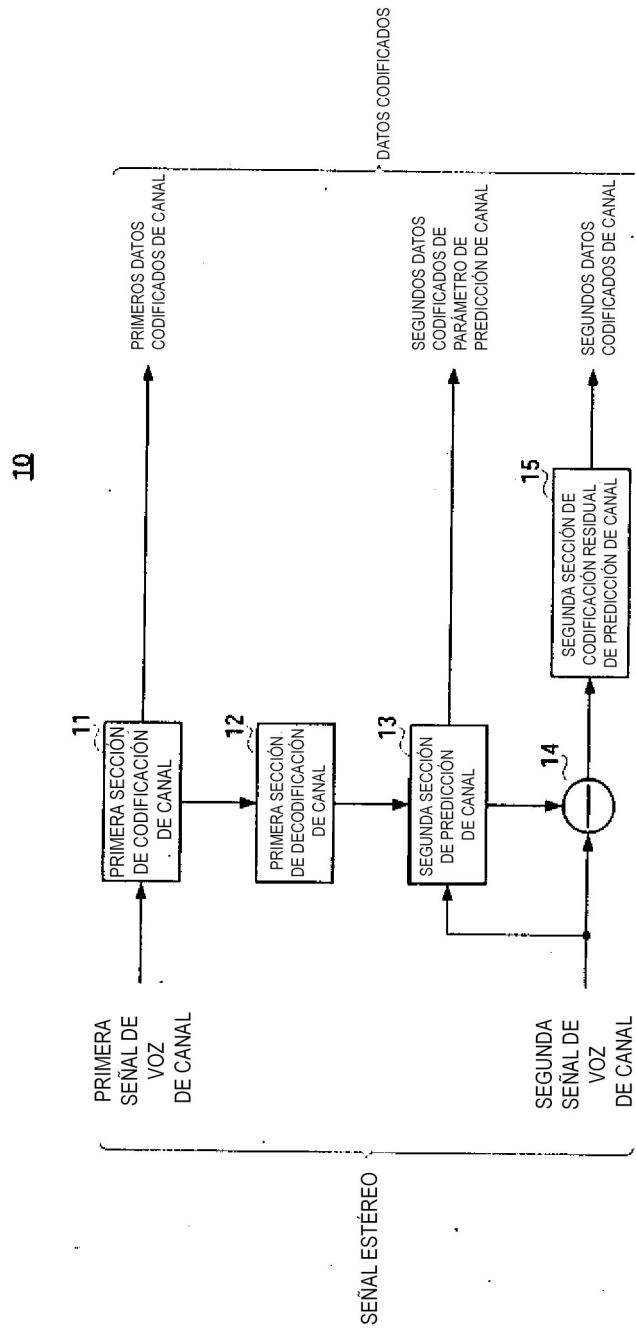


FIG.1

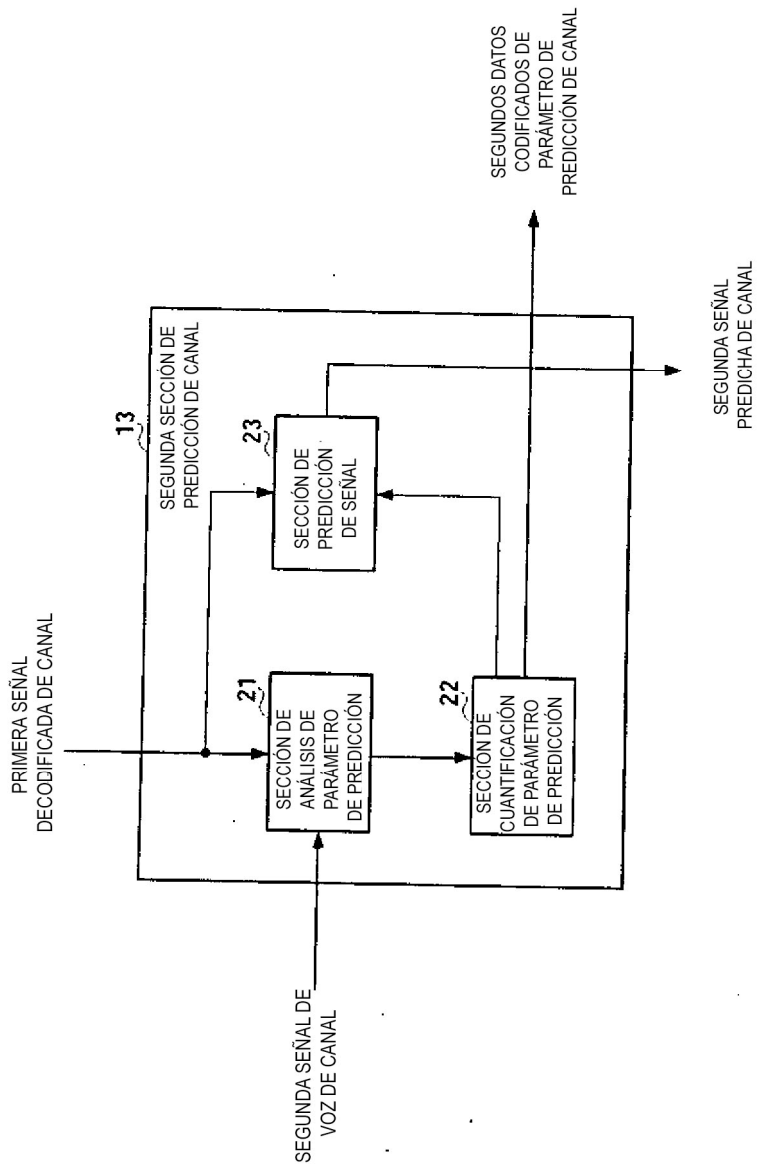


FIG. 2

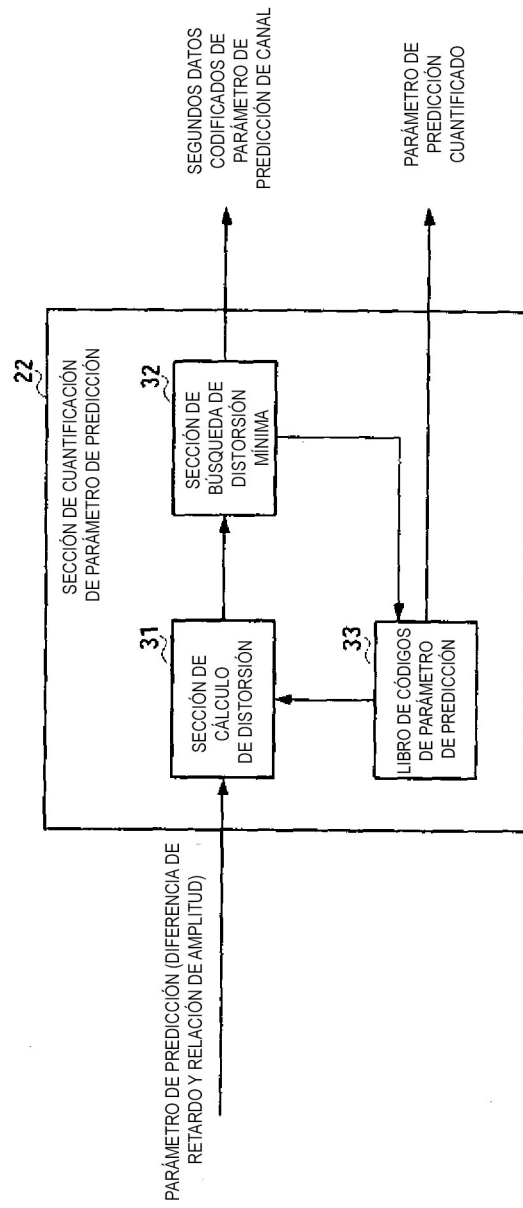


FIG. 3

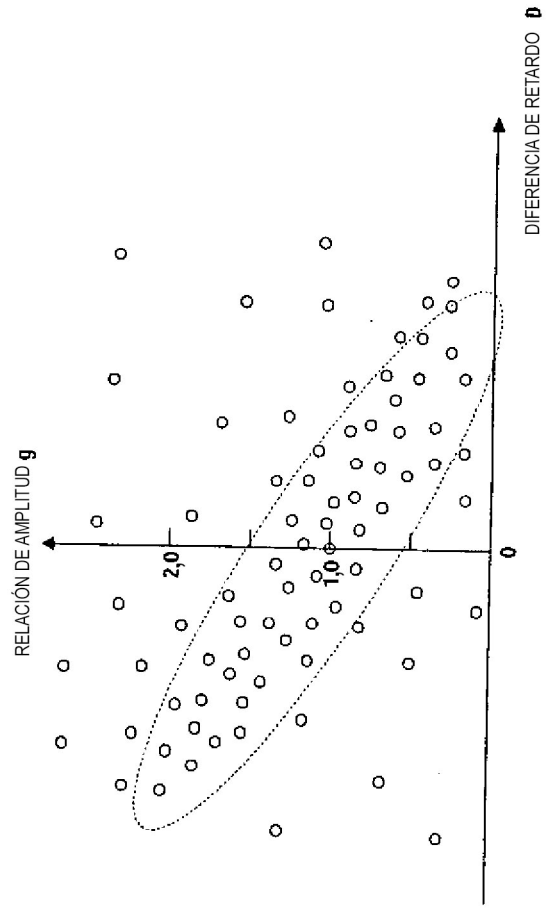


FIG. 4

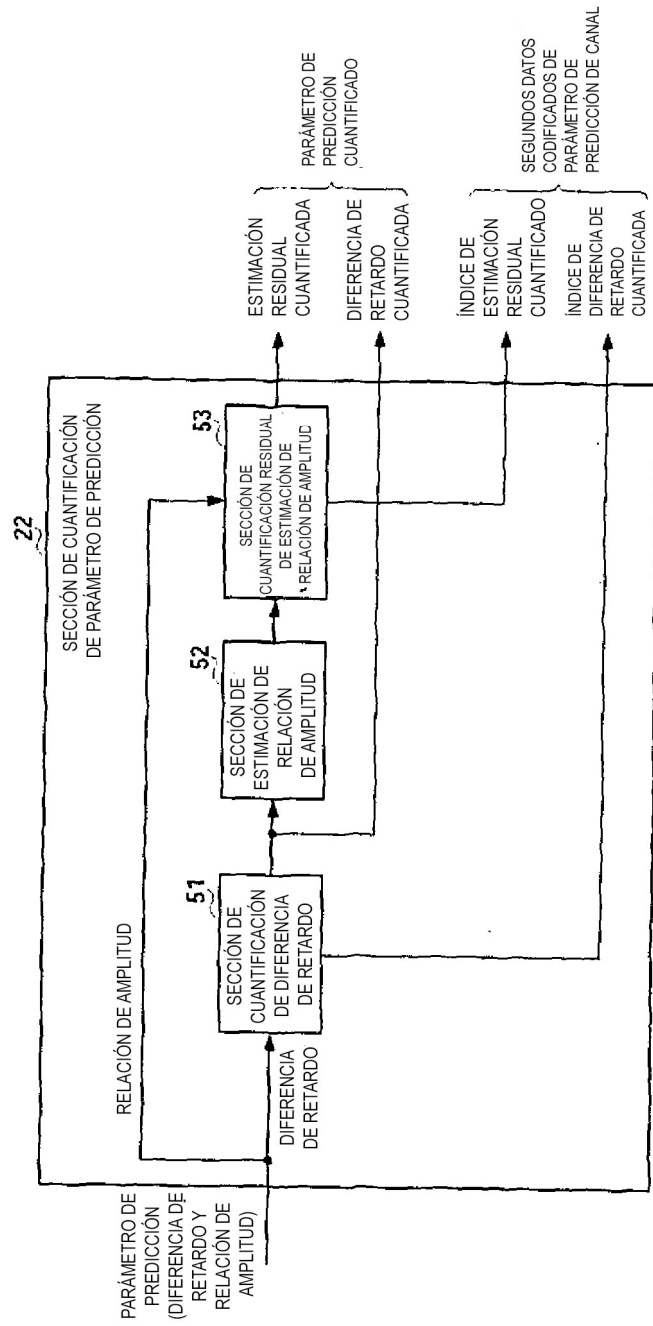


FIG. 5

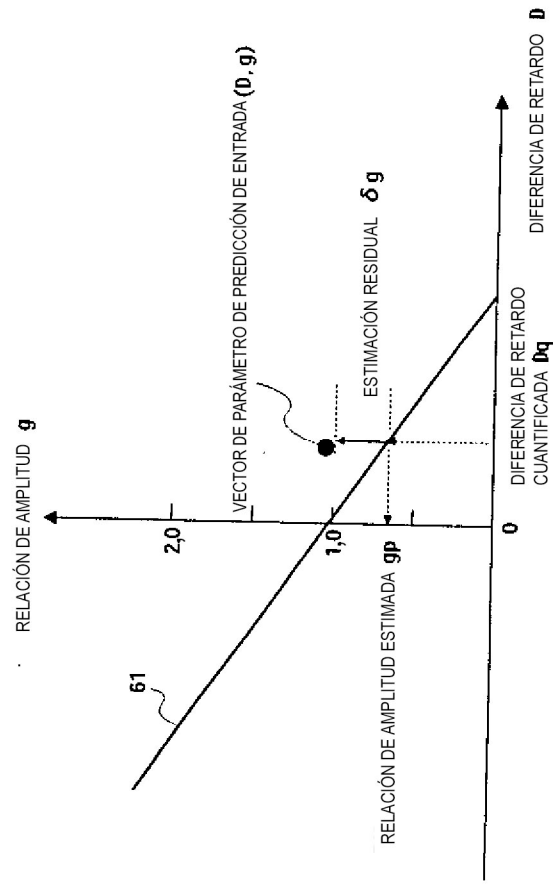


FIG. 6

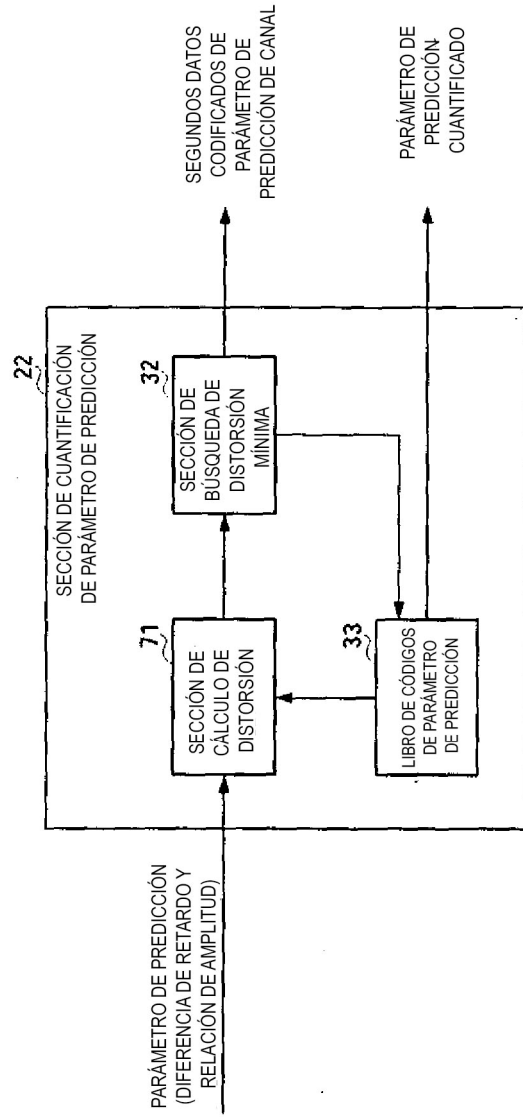


FIG. 7



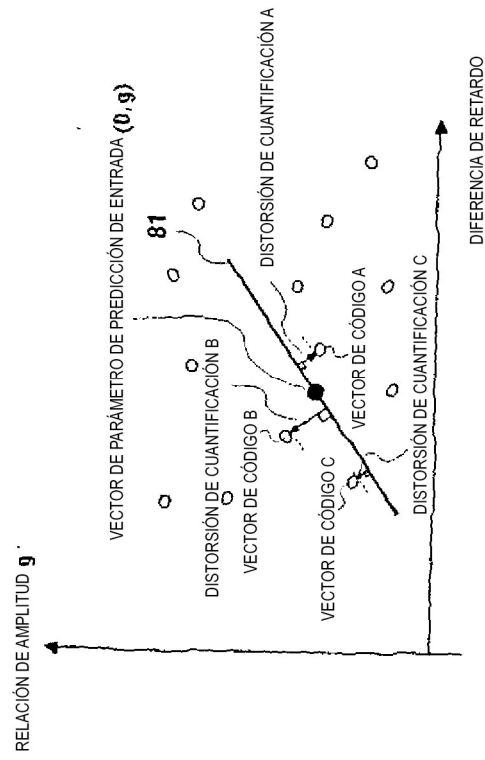


FIG. 8

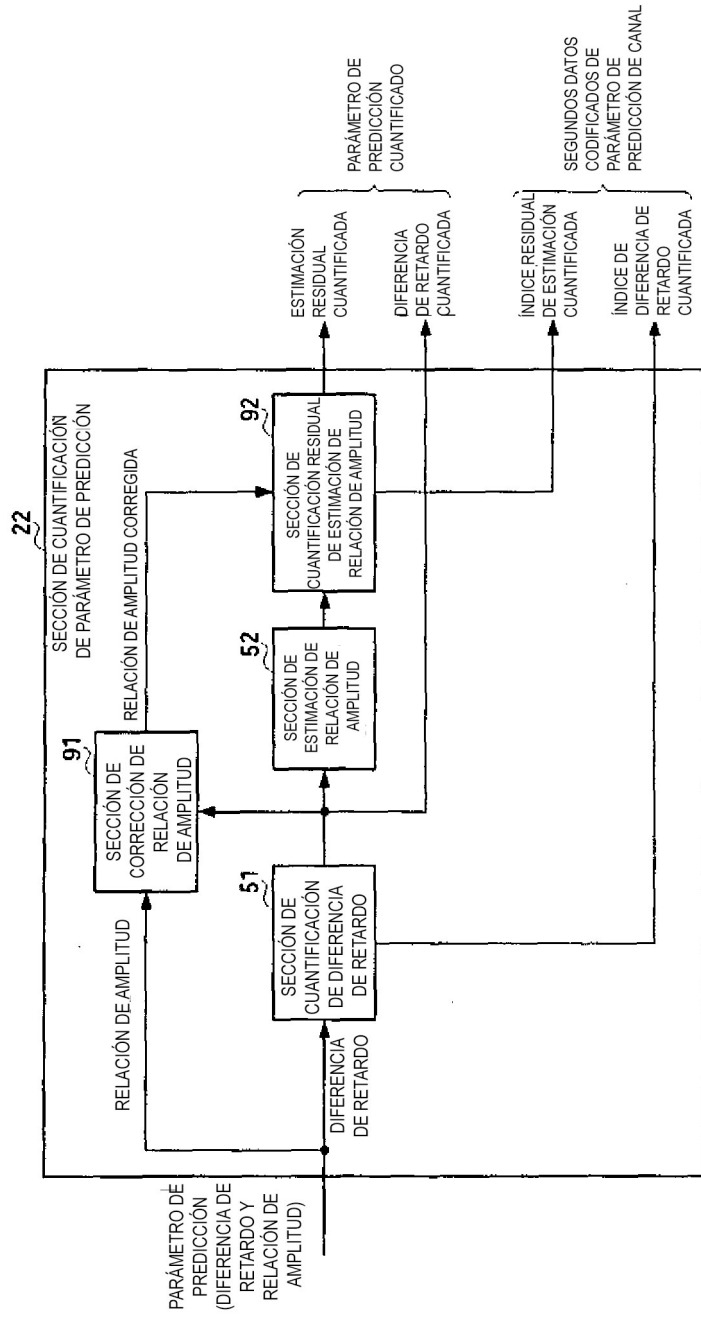


FIG. 9

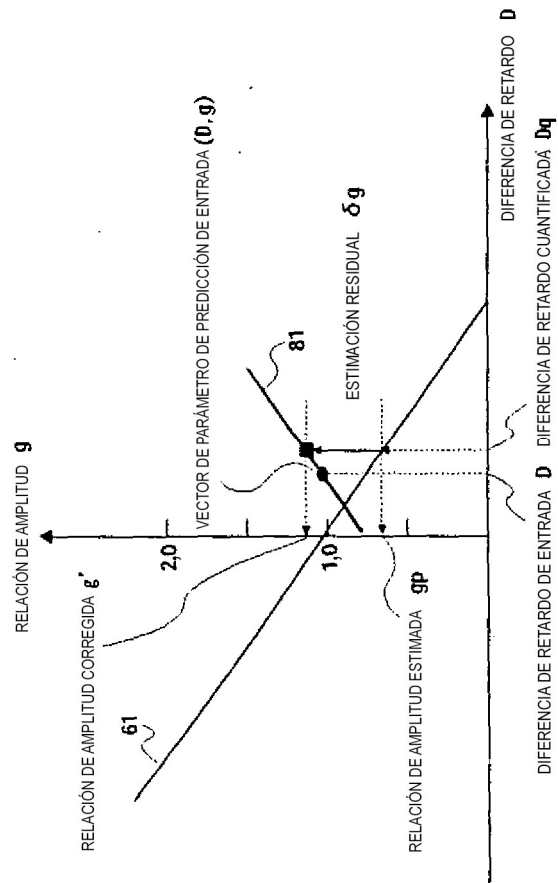


FIG.10

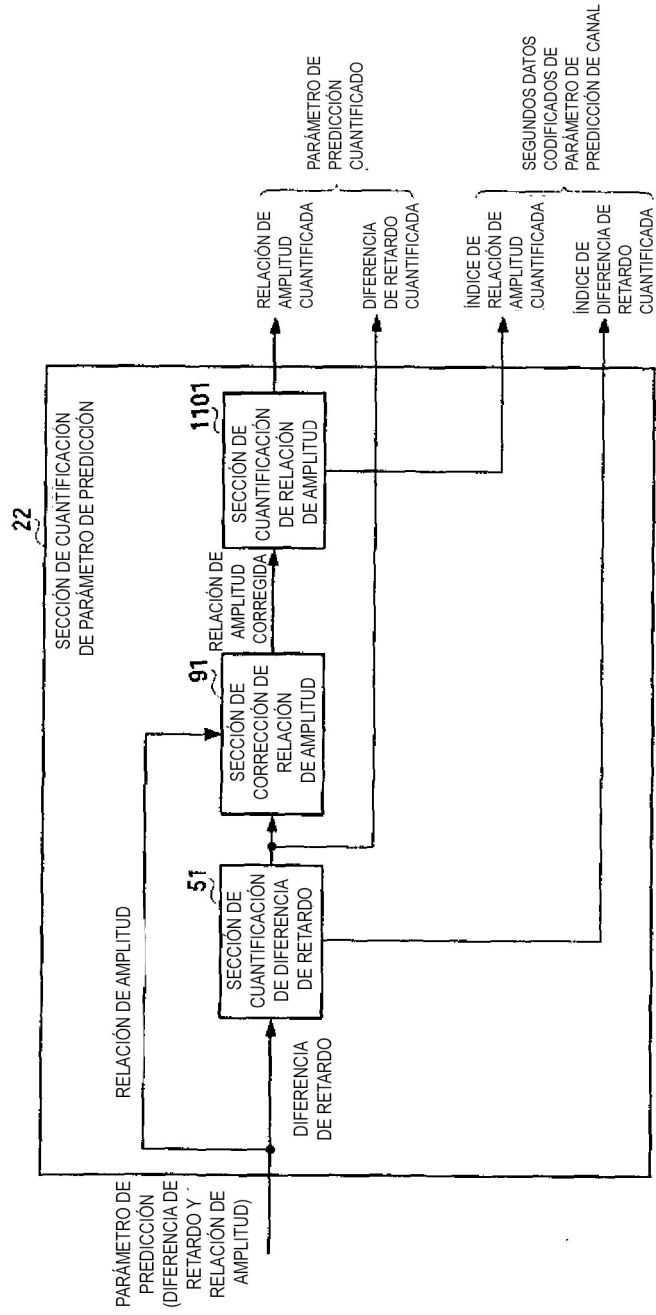


FIG.11