

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 372 133**

51 Int. Cl.:
G10L 21/02 (2006.01)
H03G 3/32 (2006.01)
H04R 3/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **09004246 .6**
96 Fecha de presentación: **25.03.2009**
97 Número de publicación de la solicitud: **2141695**
97 Fecha de publicación de la solicitud: **06.01.2010**

54 Título: **DISPOSITIVO DE MEJORA DE SONIDO DE VOZ.**

30 Prioridad:
02.07.2008 JP 2008174016

45 Fecha de publicación de la mención BOPI:
16.01.2012

45 Fecha de la publicación del folleto de la patente:
16.01.2012

73 Titular/es:
FUJITSU LIMITED
1-1, KAMIKODANAKA 4-CHOME NAKAHARA-KU
KAWASAKI-SHI, KANAGAWA 211-8588, JP

72 Inventor/es:
Endo, Kaori;
Ota, Yasuji;
Otani, Takeshi y
Togawa, Taro

74 Agente: **Ungría López, Javier**

ES 2 372 133 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 mejora de sonido de voz

5 La presente descripción se refiere a un esquema de mejora de señal audio y un dispositivo de mejora de señal audio capaces de evitar que el sonido de voz recibido sea indiscernible en el caso en que un usuario esté hablando por un sistema de transmisión de voz en un entorno donde el ruido ambiente sea alto. Específicamente, el esquema y el dispositivo de mejora de señal audio mejoran el sonido de voz recibido de modo que el usuario sea capaz de oír fácilmente el sonido de voz recibido.

10

Antecedentes

15 Los procesos de mejora de sonido de voz existentes conocidos incluyen una tecnología para usar características de señales audio y mejorar formantes, que son múltiples picos espectrales del sonido de voz recibido e influyen en la identificación de vocales, según el nivel de ruido ambiente (véase el documento de Patente 1, por ejemplo), y una tecnología que mejora el sonido de voz recibido según la SNR del sonido de voz recibido y el ruido ambiente (por ejemplo, el documento de Patente 2).

20 [Documento de Patente 1] Publicación de Patente japonesa número 4018571

[Documento de Patente 2] Publicación de la Solicitud de Patente japonesa número 2000-349893

25 Sin embargo, en el caso donde el sonido de voz recibido se mejora según la tecnología convencional anterior (documento de Patente 1 o 2), el equilibrio entre el rango de frecuencias bajas y el rango de frecuencias altas se pierde, y el sonido mejorado es apagado (una componente de frecuencia baja es excesivamente dominante) o es agudo (una componente de frecuencia alta es excesivamente dominante). Así, las tecnologías convencionales tienen el problema de que su mejora está sujeta a una reducción de la calidad del sonido de voz y no es suficiente para aclarar el sonido de voz recibido.

30 En US 2005/290401 A1 se describe un circuito de supresión de ruido, en el que una señal de entrada es convertida al dominio de frecuencia y dividida en bandas Bark. Un factor de supresión de ruido se ajusta inversamente proporcional a la relación de señal a ruido de la señal de entrada. Se resta una estimación de ruido de la señal en cada banda y las señales son convertidas de nuevo al dominio de tiempo.

Resumen

35 En vista de los problemas convencionales antes descritos, la presente descripción tiene la finalidad de proporcionar un dispositivo de mejora de sonido de voz capaz de mejorar el sonido de voz recibido de modo que el usuario en un entorno donde el ruido ambiente sea alto, sea capaz de oír fácilmente el sonido de voz recibido.

40 Este objeto se logra con la materia de la reivindicación independiente 1. Otras realizaciones ventajosas se describen en las reivindicaciones dependientes.

45 Un aspecto de la presente descripción puede ser un dispositivo de mejora de sonido de voz incluyendo una unidad de cálculo SNR configurada para calcular una SNR que es una relación de sonido de voz recibido a ruido ambiente; una unidad de cálculo de magnitud de mejora de primer rango de frecuencias configurada para calcular, en base a la SNR e información de división de rango de frecuencias que indica un primer rango de frecuencias y un segundo rango de frecuencias, la magnitud de mejora del primer rango de frecuencias en el caso de que el sonido de voz recibido sea enmascarado por el ruido ambiente, obteniéndose preliminarmente el primer rango de frecuencias y contribuyendo a una mejora de la inteligibilidad subjetiva del sonido de voz recibido, obteniéndose preliminarmente el segundo rango de frecuencias y contribuyendo a una mejora de la articulación subjetiva del sonido de voz recibido; una unidad de cálculo de magnitud de mejora de segundo rango de frecuencias configurada para calcular la magnitud de mejora del segundo rango de frecuencias en base a la magnitud de mejora del primer rango de frecuencias; y una unidad de procesamiento de espectro configurada para procesar espectros del sonido de voz recibido usando la magnitud de mejora del primer rango de frecuencias, la magnitud de mejora del segundo rango de frecuencias y la información de división de rango de frecuencias.

50 Según la estructura anterior se realizan las operaciones siguientes: la unidad de cálculo SNR calcula el grado de la indiscernibilidad del sonido de voz recibido en base al sonido de voz recibido y el ruido ambiente; la unidad de cálculo de magnitud de mejora de primer rango de frecuencias calcula la magnitud de mejora de primer rango de frecuencias usado para mejorar suficientemente la inteligibilidad subjetiva del sonido de voz recibido; la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias calcula la magnitud de mejora de segundo rango de frecuencias usada para asegurar que se mantenga suficientemente la articulación de sonido de voz después de que el primer rango de frecuencias del sonido de voz recibido sea mejorado; y la unidad de procesamiento de espectro calcula espectros de sonido de voz mejorado realizando un proceso de mejora en cada rango de frecuencias usando la magnitud de mejora de primer rango de frecuencias y la magnitud de mejora de segundo rango de frecuencias.

65

El objeto y las ventajas de la descripción se realizarán y lograrán por medio de los elementos y combinaciones expuestos en particular en las reivindicaciones.

- 5 Se ha de entender que tanto la descripción general anterior como la descripción detallada siguiente son ejemplares y explicativas y no son restrictivas de la presente descripción reivindicada.

Breve descripción de los dibujos

- 10 La figura 1 es una ilustración de la primera realización de la presente descripción.
La figura 2 es una ilustración de la segunda realización de la presente descripción.
La figura 3 es una ilustración de la tercera realización de la presente descripción.
15 La figura 4 es una ilustración de la cuarta realización de la presente descripción.
La figura 5 es una ilustración de la quinta realización de la presente descripción.
20 La figura 6 es un gráfico para calcular una magnitud de mejora de primer rango de frecuencias.
La figura 7 es un gráfico para calcular una magnitud de mejora de segundo rango de frecuencias.
La figura 8 es un gráfico para corregir la magnitud de mejora de segundo rango de frecuencias.
25 La figura 9 es un ejemplo específico 1 de un gráfico usado para calcular la magnitud de mejora de cada rango de frecuencias.
Y la figura 10 es un ejemplo específico 2 de un gráfico usado para calcular la magnitud de mejora de cada rango de frecuencias.
30

Descripción de la realización

- 35 A continuación se explican con referencia a los dibujos realizaciones que describen el mejor modo de llevar a la práctica la presente invención.

[Primera realización]

- 40 La figura 1 es una ilustración que representa una estructura de la primera realización de la presente descripción. Los números de referencia 1 y 2 denotan unidades de análisis de frecuencia; el número de referencia 3 denota una unidad de cálculo SNR; el número de referencia 4 denota una unidad de cálculo de magnitud de mejora de primer rango de frecuencias; el número de referencia 5 denota una unidad de cálculo de magnitud de mejora de segundo rango de frecuencias; el número de referencia 6 denota una unidad de procesamiento de espectro de sonido de voz; y el número de referencia 7 es una unidad sintetizadora.
45 La estructura ilustrada en la figura 1 se explica a continuación según un proceso de señal audio. La operación detallada de cada componente estructural se describe más adelante.
50 En primer lugar, la unidad de análisis de frecuencia 1 introduce ruido ambiente, y envía espectros del ruido ambiente.
La unidad de análisis de frecuencia 2 introduce sonido de voz recibido, y envía espectros del sonido de voz recibido.
La unidad de cálculo SNR 3 introduce los espectros del ruido ambiente y los espectros del sonido de voz recibido de las unidades de análisis de frecuencia 1 y 2, respectivamente, y calcula y envía una SNR.
55 La unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 recibe la SNR salida de la unidad de cálculo SNR e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido. Entonces, la unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la inteligibilidad subjetiva del sonido de voz recibido y la envía como magnitud de mejora de primer rango de frecuencias.
60 La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5 introduce la magnitud de mejora de primer rango de frecuencias, y calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la articulación subjetiva del sonido de voz recibido y la envía como magnitud de mejora de segundo rango
65

de frecuencias.

5 La unidad de procesado de espectro de sonido de voz 6 introduce la información de división de rango de frecuencias, la magnitud de mejora de primer rango de frecuencias y la magnitud de mejora de segundo rango de frecuencias, y envía espectros de sonido de voz mejorados.

La unidad sintetizadora 7 introduce los espectros de sonido de voz mejorados salidos de la unidad de procesado de espectro de sonido de voz 6, y envía sonido de voz mejorado.

10 El proceso para mejorar el sonido de voz recibido se realiza según el flujo de procesado anterior.

A continuación se describe una operación de procesado detallada de cada componente estructural.

Unidades de análisis de frecuencia 1 y 2

15 La unidad de análisis de frecuencia 1 usa ruido ambiente como una señal de entrada y calcula espectros del ruido ambiente usando un proceso de transformación tiempo-frecuencia, tal como el método de transformada de Fourier. La unidad de análisis de frecuencia 2 usa sonido de voz recibido como una señal de entrada y calcula espectros del sonido de voz recibido usando un proceso de transformación tiempo-frecuencia, tal como el método de transformada de Fourier.

Unidad de cálculo SNR 3

25 La unidad de cálculo SNR 3 calcula una SNR por la ecuación siguiente usando los espectros introducidos del ruido ambiente (componentes de ruido ambiente (N) para respectivas frecuencias) y los espectros introducidos del sonido de voz (componentes de sonido de voz recibido (S) para respectivas frecuencias).

$$\text{SNR}(i) = S(i) - N(i),$$

30 donde i es el índice de frecuencia (que indica un número de un rango de frecuencias (es decir, componente de frecuencia) usado en la transformación tiempo-frecuencia);

SNR(i) es la SNR de la i-ésima componente de frecuencia (dB);

35 S(i) es la magnitud de una componente de sonido de voz recibido en la i-ésima componente de frecuencia (dB); y

N(i) es la magnitud de una componente de ruido ambiente en la i-ésima componente de frecuencia (dB).

40 Obsérvese que, en el cálculo de SNR, se puede introducir un efecto de la magnitud subjetiva de sonido que cambia con respecto a cada frecuencia.

Específicamente, la SNR calculada se corrige multiplicándola por un coeficiente que representa la magnitud subjetiva de una frecuencia específica. El coeficiente que representa la magnitud subjetiva de una frecuencia específica se calcula usando información existente, tal como una curva isofónica.

Unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4

50 La unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 recibe la SNR salida de la unidad de cálculo SNR e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido, y calcula la magnitud de mejora de primer rango de frecuencias según el proceso siguiente.

55 (1) Se calcula una SNR media del primer rango de frecuencias.

[Ecuación 1]

$$\text{LSNR} = \frac{1}{le - ls + 1} \sum_{i=ls}^{le} \text{SNR}(i)$$

60 LSNR: la SNR media del primer rango de frecuencias (dB)

i: el índice de frecuencia

SNR(i): la SNR de la i-ésima componente de frecuencia (dB)

Is: el índice de frecuencia de límite inferior del primer rango de frecuencias

Ie: el índice de frecuencia de límite superior del primer rango de frecuencias

“Is” e “Ie” se calculan como sigue a continuación usando la información de división de rango de frecuencias (un índice de frecuencia límite I_b entre el rango de frecuencias bajas y el rango de frecuencias altas) y la anchura de rango de frecuencias de la señal de entrada. Es = el índice de frecuencia de límite inferior de la anchura de rango de frecuencias de la señal de entrada

$I_e = I_b$

(2) La SNR media del primer rango de frecuencias se aplica a un gráfico de magnitud de mejora, obteniendo por ello la magnitud de mejora de primer rango de frecuencias.

El gráfico de magnitud de mejora del primer rango de frecuencias se ilustra en la figura 6.

LE: la magnitud de mejora de primer rango de frecuencias (dB)

LSNR: la SNR media del primer rango de frecuencias (dB)

LE MAX: el límite superior de la magnitud de mejora de primer rango de frecuencias (dB)

LE MIN: el límite inferior de la magnitud de mejora de primer rango de frecuencias (dB)

LSNR MAX: la SNR media del primer rango de frecuencias en el que se obtiene el límite inferior de la magnitud de mejora de primer rango de frecuencias (dB)

LSNR MIN: la SNR media del primer rango de frecuencias en el que se obtiene el límite superior de la magnitud de mejora de primer rango de frecuencias (dB)

El rango posible de cada parámetro del gráfico es el siguiente.

LE MAX: de 0 a 50 dB, el límite superior proporcionado de modo que la mejora del primer rango de frecuencias no exceda del nivel máximo si la SNR media del primer rango de frecuencias es igual o menor que un valor predeterminado.

LE MIN: de 0 a 10 dB, el límite inferior proporcionado de modo que la mejora del primer rango de frecuencias no se quede corta del nivel mínimo si la SNR media del primer rango de frecuencias es igual o mayor que un valor predeterminado. Obsérvese que LE MAX debe ser mayor que LE MIN ($LE\ MIN < LE\ MAX$).

LSNR MAX: de 0 a 50 dB, la SNR media máxima del primer rango de frecuencias con que se considera que es apropiado poner la magnitud de mejora de primer rango de frecuencias al nivel mínimo dado que el sonido de voz es mayor que el ruido de extremo próximo en el primer rango de frecuencias.

LSNR MIN: de -50 a 50 dB, la SNR media mínima del primer rango de frecuencias con que se considera que es apropiado poner la magnitud de mejora de primer rango de frecuencias al nivel máximo dado que el sonido de voz es menor que el ruido de extremo próximo en el primer rango de frecuencias.

Obsérvese que LSNR MAX debe ser mayor que LSNR MIN ($LSNR\ MIN < LSNR\ MAX$).

Unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5

La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5 introduce la magnitud de mejora de primer rango de frecuencias, y calcula la magnitud de mejora de segundo rango de frecuencias usando un gráfico ilustrado en la figura 7 de la siguiente manera.

[Ecuación 2]

$$HE(dB) = HE_{MIN} + \frac{HE_{MAX} - HE_{MIN}}{LE_{MAX} - LE_{MIN}} * (LE - LE_{MIN})$$

HE: la magnitud de mejora de segundo rango de frecuencias (dB)

LE: la magnitud de mejora de primer rango de frecuencias (dB)

5

HE MAX: el límite superior de la magnitud de mejora del segundo rango de frecuencias (dB)

HE MIN: el límite inferior de la magnitud de mejora del segundo rango de frecuencias (dB)

10 LE MAX: la magnitud de mejora de primer rango de frecuencias en que se obtiene el límite superior de la magnitud de mejora de segundo rango de frecuencias (dB)

LE MIN: la magnitud de mejora de primer rango de frecuencias en que se obtiene el límite inferior de la magnitud de mejora de segundo rango de frecuencias (dB)

15

El rango posible de cada parámetro del gráfico es el siguiente.

LE MAX: de 0 a 50 dB, el límite superior proporcionado de modo que la mejora del primer rango de frecuencias no exceda del nivel máximo si la SNR media del primer rango de frecuencias es igual o menor que un valor predeterminado.

20

LE MIN: de 0 a 10 dB, el límite inferior proporcionado de modo que la mejora del primer rango de frecuencias no se quede corta del nivel mínimo si la SNR media del primer rango de frecuencias es igual o mayor que un valor predeterminado. Obsérvese que LE MAX debe ser mayor que LE MIN (LE MIN < LE MAX). (Estos dos parámetros son los mismos que los de la figura 6).

25

HE MAX: de 0 a 50 dB, el valor límite superior de la magnitud de mejora de segundo rango de frecuencias que se usa cuando la magnitud de mejora de segundo rango de frecuencias se determina en base a la magnitud de mejora de primer rango de frecuencias y se facilita de modo que la mejora del segundo rango de frecuencias no exceda del nivel máximo si la magnitud de mejora de primer rango de frecuencias es igual o mayor que un valor predeterminado.

30

HE MIN: de 0 a 50 dB, el valor límite inferior de la magnitud de mejora de segundo rango de frecuencias que se usa cuando la magnitud de mejora de segundo rango de frecuencias se determina en base a la magnitud de mejora de primer rango de frecuencias y se facilita de modo que la mejora del segundo rango de frecuencias no se quede corta del nivel mínimo si la magnitud de mejora de primer rango de frecuencias es igual o menor que un valor predeterminado.

35

Obsérvese que HE MAX debe ser mayor que HE MIN (HE MIN < HE MAX).

40 **Unidad de procesamiento de espectro de sonido de voz 6**

La unidad de procesamiento de espectro de sonido de voz 6 introduce los espectros del sonido de voz salidos de la unidad de análisis de frecuencia 2, la magnitud de mejora de primer rango de frecuencias, la magnitud de mejora de segundo rango de frecuencias y la información de división de rango de frecuencias, y calcula la magnitud de mejora para los espectros del sonido de voz según el proceso siguiente y realiza un proceso de mejora.

45

A continuación se describen ejemplos específicos del cálculo de magnitud de mejora de cada componente de frecuencia realizado por la unidad de procesamiento de espectro de sonido de voz 6. «Ejemplo específico 1 del cálculo de magnitud de mejora de cada componente de frecuencia realizado por la unidad de procesamiento de espectro de sonido de voz 6»

50

La magnitud de mejora de cada componente de frecuencia se calcula usando un gráfico ilustrado en la figura 9.

[Ecuación 3]

55

$$EMP(i) = \begin{cases} LE & i \leq l1e \\ LE + \frac{HE}{h1s - l1e} * (i - l1e) & l1e < i < h1s \\ HE & i \geq h1s \end{cases}$$

EMP(i): la magnitud de mejora de cada componente de frecuencia (dB)

5 i: el índice de frecuencia

LE: la magnitud de mejora de primer rango de frecuencias (dB)

HE: la magnitud de mejora de segundo rango de frecuencias (dB)

10

l1s: el índice de frecuencia de límite inferior del primer rango de frecuencias

l1e: el índice de frecuencia de límite superior del primer rango de frecuencias

15

h1s: el índice de frecuencia de límite inferior del segundo rango de frecuencias

h1e: el índice de frecuencia de límite superior del segundo rango de frecuencias

20

“l1e”, “l1e”, “h1s” y “h1e” se calculan de la siguiente manera usando la información de división de rango de frecuencias (el índice de frecuencia límite I_b entre el primer rango de frecuencias y el segundo rango de frecuencias) y la anchura de rango de frecuencias de la señal de entrada.

l1s = el índice de frecuencia de límite inferior de la anchura de rango de frecuencias de la señal de entrada

$$l1e = I_b - \psi$$

25

$$h1s = I_b + \zeta$$

h1e = el índice de frecuencia de límite superior de la anchura de rango de frecuencias de la señal de entrada

30

ψ, ζ : valores positivos ((valor entre 1 a 0,5) x ind max)

ind max: el valor máximo del índice de frecuencia

35

ψ y ζ son valores para determinar una anchura de rango de frecuencias, que se facilita para evitar la aparición de discontinuidad en la magnitud de mejora en el límite de los rangos de frecuencia. «Ejemplo específico 2 del cálculo de magnitud de mejora de cada componente de frecuencia realizado por la unidad de procesado de espectro de sonido de voz»

La magnitud de mejora de cada componente de frecuencia se calcula usando un gráfico ilustrado en la figura 10.

40

[Ecuación 4]

$$EMP(i) = \begin{cases} LE & i \leq l2e \\ LE + (i - l2e) * HE & l2e < i \leq h2e \cap EMP(i) < EMP \text{ MAX} \\ EMP \text{ MAX} & l2e < i \leq h2e \cap EMP(i) \geq EMP \text{ MAX} \end{cases}$$

EMP(i): la magnitud de mejora de cada componente de frecuencia (dB)

i: el índice de frecuencia

LE: la magnitud de mejora de primer rango de frecuencias (dB)

HE: la magnitud de mejora de segundo rango de frecuencias (dB)

EMP MAX: el límite superior de la magnitud de mejora (dB)

l2s: el índice de frecuencia de límite inferior del primer rango de frecuencias

l2e: el índice de frecuencia de límite superior del primer rango de frecuencias

h2s: el índice de frecuencia de límite inferior del segundo rango de frecuencias, h2s = l2e

h2e: el índice de frecuencia de límite superior del segundo rango de frecuencias

“h2s”, “h2e”, “l2s” y “l2e” se calculan de la siguiente manera usando la información de división de rango de frecuencias (el índice de frecuencia límite l_b entre el primer rango de frecuencias y el segundo rango de frecuencias) y la anchura de rango de frecuencias de la señal de entrada.

l2s = el índice de frecuencia de límite inferior de la anchura de rango de frecuencias de la señal de entrada

h2e = l_b

l2s = l_b

h2e = el índice de frecuencia de límite superior de la anchura de rango de frecuencias de la señal de entrada

A continuación se describe un ejemplo específico de procesado de espectro para cada componente de frecuencia realizado por la unidad de procesado de espectro de sonido de voz 6. «Ejemplo específico de procesado de espectro para cada componente de frecuencia realizado por la unidad de procesado de espectro de sonido de voz »

En primer lugar, usando la magnitud de mejora de cada componente de frecuencia EMP(i), la unidad de procesado de espectro de sonido de voz 6 calcula un coeficiente emp coef(i) a aplicar a los espectros de cada componente de frecuencia.

[Ecuación 5]

$$\text{empcoef}(i) = 10^{\frac{\text{EMP}(i)}{20}}$$

A continuación, la parte real y la parte imaginaria de los espectros de cada componente de frecuencia se multiplican por el coeficiente correspondiente emp coef(i), procesando por ello los espectros de cada componente de frecuencia.

[Ecuación 6]

$$\text{ESPE re}(i) = \text{SPE re}(i) * \text{empcoef}(i)$$

$$\text{ESPE im}(i) = \text{SPE im}(i) * \text{empcoef}(i)$$

$$\text{empcoef}(i) = 10^{\frac{\text{EMP}(i)}{20}}$$

EMP(i): la magnitud de mejora de cada componente de frecuencia (dB)

i: el índice de frecuencia

SPE re(i): la parte real de los espectros de cada componente de frecuencia

SPE im(i): la parte imaginaria de los espectros de cada componente de frecuencia

ESPE re(i): la parte real de los espectros mejorados de cada componente de frecuencia

5 ESPE im(i): la parte imaginaria de los espectros mejorados de cada componente de frecuencia

emp coef (i): el factor de multiplicación de los espectros de cada componente de frecuencia

Unidad sintetizadora 7

10 La unidad sintetizadora 7 introduce los espectros de sonido de voz mejorado salidos de la unidad de procesamiento de espectro de sonido de voz 6, y convierte los espectros de sonido de voz mejorado a una forma de onda de dominio de tiempo usando un proceso de transformación de frecuencia-tiempo, tal como el método de transformada de Fourier inversa. De esta forma, la unidad sintetizadora 7 genera y envía sonido de voz mejorado.

15 Así, según la presente realización, con la operación de procesamiento de cada componente estructural, es posible generar sonido de voz mejorado que tiene una calidad suficiente tanto en inteligibilidad subjetiva como en articulación subjetiva en el caso donde el sonido de voz recibido es difícil de oír debido a ruido ambiente.

[Segunda realización]

20 La figura 2 es una ilustración que representa una estructura de la segunda realización de la presente descripción, y los componentes que son comunes a los de la primera realización reciben los mismos números de referencia. Los números de referencia 1 y 2 denotan las unidades de análisis de frecuencia; el número de referencia 3 denota la
 25 unidad de cálculo SNR; el número de referencia 4 denota la unidad de cálculo de magnitud de mejora de primer rango de frecuencias; el número de referencia 5' denota una unidad de cálculo de magnitud de mejora de segundo rango de frecuencia; el número de referencia 6 denota la unidad de procesamiento de espectro de sonido de voz; el número de referencia 7 es la unidad sintetizadora; y el número de referencia 8 es una unidad de cálculo de articulación de sonido de voz.

30 La estructura ilustrada en la figura 2 se explica a continuación según un proceso de señal audio. La operación detallada de cada componente estructural se describe más adelante.

35 En primer lugar, la unidad de análisis de frecuencia 1 introduce ruido ambiente, y envía espectros del ruido ambiente.

La unidad de análisis de frecuencia 2 introduce sonido de voz recibido, y envía espectros del sonido de voz recibido.

40 La unidad de cálculo de articulación de sonido de voz 8 introduce los espectros del sonido de voz recibido enviados desde la unidad de análisis de frecuencia 2 e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido, y después envía información que indica articulación del sonido de voz.

45 La unidad de cálculo SNR 3 introduce los espectros del ruido ambiente y los espectros del sonido de voz recibido de las unidades de análisis de frecuencia 1 y 2, respectivamente, y calcula y envía una SNR.

50 La unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 introduce la SNR salida de la unidad de cálculo SNR y la información de división de rango de frecuencias, y después calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la inteligibilidad subjetiva del sonido de voz recibido y la envía como magnitud de mejora de primer rango de frecuencias.

55 La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5' introduce la magnitud de mejora de primer rango de frecuencias y la información que indica articulación del sonido de voz enviada desde la unidad de cálculo de articulación de sonido de voz 8, y calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la articulación subjetiva del sonido de voz recibido y la envía como magnitud de mejora de segundo rango de frecuencias.

60 La unidad de procesamiento de espectro de sonido de voz 6 introduce la información de división de rango de frecuencias, la magnitud de mejora de primer rango de frecuencias y la magnitud de mejora de segundo rango de frecuencias, y envía espectros de sonido de voz mejorado.

La unidad sintetizadora 7 introduce los espectros de sonido de voz mejorado enviados desde la unidad de procesamiento de espectro de sonido de voz 6, y envía sonido de voz mejorado.

65 El proceso para mejorar el sonido de voz recibido se realiza según el flujo de procesamiento anterior.

A continuación se describen las operaciones de procesado detalladas de los componentes estructurales que son diferentes de los de la primera realización.

Unidad de cálculo de articulación de sonido de voz 8

5 La unidad de cálculo de articulación de sonido de voz 8 introduce los espectros del sonido de voz recibido y la información de división de rango de frecuencias, y genera y envía la información que indica articulación del sonido de voz según el proceso siguiente.

10 [Ecuación 7]

$$C = H - L$$

$$H = \sum_{i=h3s}^{h3e} S(i)$$

$$L = \sum_{i=l3s}^{l3e} S(i)$$

C: la articulación del sonido de voz (dB)

15 H: la potencia del segundo rango de frecuencias del sonido de voz recibido (dB)

L: la potencia del primer rango de frecuencias del sonido de voz recibido (dB)

i: el índice de frecuencia

20 S(i): la magnitud de una componente de sonido de voz recibido en la i-ésima componente de frecuencia (dB)

l3s: el índice de frecuencia de límite inferior del primer rango de frecuencias

25 l3e: el índice de frecuencia de límite superior del primer rango de frecuencias

h3s: el índice de frecuencia de límite inferior del segundo rango de frecuencias

30 h3e: el índice de frecuencia de límite superior del segundo rango de frecuencias

“h3s”, “h3e”, “l3s” y “l3e” se calculan de la siguiente manera usando la información de división de rango de frecuencias (el índice de frecuencia límite l_b entre el primer rango de frecuencias y el segundo rango de frecuencias) y la anchura de rango de frecuencias de la señal de entrada.

35 $h3s = l_b$

h3e = el índice de frecuencia de límite superior de la anchura de rango de frecuencias de la señal de entrada

l3s = el índice de frecuencia de límite inferior de la anchura de rango de frecuencias de la señal de entrada

40 $l3e = l_b - 1$

Obsérvese que, en el cálculo de articulación, se puede introducir un efecto donde la magnitud subjetiva de sonido cambia con respecto a cada frecuencia.

45 Específicamente, cuando se calculan H y L, se realiza corrección multiplicando S(i) por un coeficiente que representa la magnitud subjetiva de una frecuencia específica. El coeficiente que representa la magnitud subjetiva de una frecuencia específica se calcula usando la información existente, tal como una curva isofónica.

Unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5'

La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5' introduce la magnitud de mejora de

primer rango de frecuencias y la información que indica articulación del sonido de voz, y calcula la magnitud de mejora de segundo rango de frecuencias usando el gráfico ilustrado en la figura 7 de la siguiente manera.

5 La magnitud de mejora de segundo rango de frecuencias HE se calcula de manera similar a la calculada por la unidad de cálculo de ampliación de mejora de segundo rango de frecuencias 5 de la primera realización.

La HE calculada se corrige usando un gráfico ilustrado en la figura 8.

10 [Ecuación 8]

$$\text{COEF} = \text{COEF MAX} + \frac{\text{COEF MIN} - \text{COEF MAX}}{\text{LE MAX} - \text{LE MIN}} * (\text{LE} - \text{LE MIN})$$

$$\text{HE}' = \text{HE} * \text{COEF}$$

El HE' corregido es asignado nuevamente como la magnitud de mejora de segundo rango de frecuencias HE.

15 HE: la magnitud de mejora de segundo rango de frecuencias (dB)

HE': la magnitud de mejora de segundo rango de frecuencias corregida obtenida usando la articulación del sonido de voz (dB)

20 COEF: el coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias

COEF MAX: el límite superior del coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias

25 COEF MIN: el límite inferior del coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias

C MAX: la articulación del sonido de voz en la que se obtiene el límite inferior del coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias (dB)

30 C MIN: la articulación del sonido de voz en la que se obtiene el límite superior del coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias (dB)

El rango posible de cada parámetro de la tabla es el siguiente.

35 COEF MAX: de 0 a 1, el límite superior proporcionado de modo que el coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias HE no exceda del nivel máximo si la articulación del sonido de voz es igual o menor que un valor predeterminado.

40 COEF MIN: de 0 a 1, el límite inferior proporcionado de modo que el coeficiente usado para corregir la magnitud de mejora de segundo rango de frecuencias HE no se quede corto del nivel mínimo si la articulación del sonido de voz es igual o mayor que un valor predeterminado.

Obsérvese que COEF MAX debe ser mayor que COEF MIN (COEF MIN < COEF MAX).

45 C MAX: de -50 a 50 dB, el valor límite superior de la articulación del sonido de voz usado para determinar el coeficiente para corregir la magnitud de mejora de segundo rango de frecuencias.

50 C MIN: de -90 a 0 dB, el valor límite inferior de la articulación del sonido de voz usado para determinar el coeficiente para corregir la magnitud de mejora de segundo rango de frecuencias.

Obsérvese que C MAX debe ser mayor que C MIN (C MIN < C MAX).

55 En la segunda realización, la unidad de cálculo de articulación de sonido de voz 8 se añade a la estructura de la primera realización con el fin de calcular la articulación del sonido de voz. Consiguientemente, la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5' calcula la magnitud de mejora de segundo rango de frecuencias, tomando en consideración la articulación del sonido de voz. Como resultado, es posible regular la

5 magnitud de mejora de segundo rango de frecuencias de modo que sea pequeña en el caso donde la articulación del sonido de voz es excesiva y, por lo tanto, es probable que el sonido mejorado sea agudo si la magnitud de mejora de segundo rango de frecuencias es grande. Por otra parte, la magnitud de mejora de segundo rango de frecuencias se regula de manera que sea grande en el caso donde el nivel de la articulación del sonido de voz sea bajo y, por lo tanto, es probable que el sonido mejorado se apagado si la magnitud de mejora de segundo rango de frecuencias es pequeña. De esta manera, es posible evitar que el sonido de voz mejorado sea apagado o agudo.

[Tercera realización]

10 La figura 3 es una ilustración que representa una estructura de la tercera realización de la presente descripción, y los componentes que son comunes a los de la primera realización reciben los mismos números de referencia. El número de referencia 1 denota la unidad de análisis de frecuencia; el número de referencia 9 denota una unidad de estimación de espectro de sonido de voz; el número de referencia 10 denota una unidad de estimación de espectro de ruido; el número de referencia 11 denota una unidad de cálculo NNR; el número de referencia 3 denota la unidad de cálculo SNR; el número de referencia 4 es la unidad de cálculo de magnitud de mejora de primer rango de frecuencias; el número de referencia 5 denota la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias; el número de referencia 6' denota la unidad de procesado de espectro de sonido de voz; y el número de referencia 7 es la unidad sintetizadora.

20 La estructura ilustrada en la figura 3 se explica a continuación según un proceso de señal audio. La operación detallada de cada componente estructural se describe más adelante.

En primer lugar, la unidad de análisis de frecuencia 1 introduce ruido ambiente, y envía espectros del ruido ambiente.

25 La unidad de estimación de espectro de sonido de voz 9 introduce sonido de voz recibido, y genera y envía espectros del sonido de voz recibido del que se quitó ruido de fondo superpuesto en el sonido de voz recibido.

La unidad de estimación de espectro de ruido 10 introduce el sonido de voz recibido, y genera y envía espectros del ruido de fondo superpuesto en el sonido de voz recibido.

30 La unidad de cálculo NNR 11 introduce los espectros del ruido ambiente salidos de la unidad de análisis de frecuencia 1 y los espectros del ruido de fondo del sonido de voz recibido salidos de la unidad de estimación de espectro de ruido 10, y calcula y envía una NNR, que es una relación de ruido ambiente a ruido de fondo superpuesto en el sonido de voz.

35 La unidad de cálculo SNR 3 introduce los espectros del ruido ambiente salidos de la unidad de análisis de frecuencia 1 y los espectros del sonido de voz, del que se ha quitado ruido de fondo, salidos de la unidad de estimación de espectro de sonido de voz 9, y calcula y envía una SNR.

40 La unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 introduce la SNR salida de la unidad de cálculo SNR e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido. Entonces, la unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la inteligibilidad subjetiva del sonido de voz recibido y la envía como magnitud de mejora de primer rango de frecuencias.

50 La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5 introduce la magnitud de mejora de primer rango de frecuencias, y calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la articulación subjetiva del sonido de voz recibido y la envía como magnitud de mejora de segundo rango de frecuencias.

55 La unidad de procesado de espectro de sonido de voz 6' introduce la información de división de rango de frecuencias, la magnitud de mejora de primer rango de frecuencias, la magnitud de mejora de segundo rango de frecuencias, la NNR y los espectros del sonido de voz, y envía espectros de sonido de voz mejorado.

La unidad sintetizadora 7 introduce los espectros de sonido de voz mejorado salidos de la unidad de procesado de espectro de sonido de voz 6, y envía sonido de voz mejorado.

60 El proceso para mejorar el sonido de voz recibido se realiza según el flujo de procesado anterior.

A continuación se describen las operaciones de procesado detalladas de los componentes estructurales que son diferentes de los de las realizaciones primera y segunda.

65 **Unidad de estimación de espectro de sonido de voz 9 y unidad de estimación de espectro de ruido 10**

La unidad de estimación de espectro de sonido de voz 9 y la unidad de estimación de espectro de ruido 10 reciben el sonido de voz recibido, y calculan espectros del sonido de voz recibido y espectros del ruido de fondo superpuesto en el sonido de voz recibido, respectivamente.

5 Específicamente, los cálculos pueden ser realizados usando una tecnología de supresión de ruido de conocimiento público (consúltase la publicación de la Solicitud de Patente japonesa número 2005-165021).

10 Por ejemplo, se analiza una señal audio de entrada, en la que se superpone ruido, y se calcula la amplitud espectral (espectros de amplitud) de la señal audio. Usando la amplitud espectral calculada, se estiman los espectros del sonido de voz. Aquí, los espectros del sonido de voz corresponden, entre componentes de la señal audio de entrada en la que se superpone ruido, a componentes sin ruido, es decir, información que representa una señal audio pura.

Los espectros de ruido también son estimados de manera similar.

15 **Unidad de cálculo NNR 11**

La unidad de cálculo NNR 11 introduce los espectros del ruido ambiente y los espectros del ruido de fondo del sonido de voz recibido, y calcula una NNR, que es una relación de ruido ambiente a ruido de fondo superpuesto en el sonido de voz, usando la ecuación siguiente y envía la NNR.

$$20 \text{ NNR}(i) = N(i) - N'(i)$$

donde i es el índice de frecuencia;

25 $\text{NNR}(i)$ es la NNR de la i -ésima componente de frecuencia (dB);

$N(i)$ es el espectro de potencia de la i -ésima componente de frecuencia (dB); y

30 $N'(i)$ es el espectro de potencia del ruido de fondo del sonido de voz recibido en la i -ésima componente de frecuencia (dB).

El índice de frecuencia anterior indica un número de un rango de frecuencias (es decir, componente de frecuencia) usado en el cálculo espectral con la transformada de Fourier.

35 Obsérvese que, en el cálculo de NNR, se puede introducir un efecto donde la magnitud subjetiva de sonido cambia con respecto a cada frecuencia.

Específicamente, la $\text{NNR}(i)$ calculada se corrige multiplicándola por un coeficiente que representa la magnitud subjetiva de una frecuencia específica. El coeficiente que representa la magnitud subjetiva de una frecuencia específica se calcula usando información existente, tal como una curva isofónica.

40 **Unidad de procesamiento de espectro de sonido de voz 6'**

45 La unidad de procesamiento de espectro de sonido de voz 6' introduce los espectros del sonido de voz salidos de la unidad de análisis de frecuencia 2, la magnitud de mejora de primer rango de frecuencias, la magnitud de mejora de segundo rango de frecuencias, la NNR y la información de división de rango de frecuencias, y calcula la magnitud de mejora para los espectros del sonido de voz según el proceso siguiente y realiza un proceso de mejora.

50 A continuación se describe un ejemplo específico del cálculo de magnitud de mejora de cada componente de frecuencia realizado por la unidad de procesamiento de espectro de sonido de voz 6'.

En primer lugar, la unidad de procesamiento de espectro de sonido de voz 6' calcula la magnitud de mejora de cada componente de frecuencia $\text{EMP}(i)$ de la misma manera que la unidad de procesamiento de espectro de sonido de voz 6.

55 A continuación, la magnitud de mejora de cada componente de frecuencia se corrige usando la NNR.

La corrección de la magnitud de mejora se realiza de tal manera que, después de mejorar el sonido de voz recibido, la NNR, que es una relación de ruido ambiente y ruido de fondo del sonido de voz recibido, no se quede corta de un umbral.

$$60 \text{ MOD EMP}(i) = \text{EMP}(i), \text{ cuando } \text{EMP} \text{ NNR}(i) \leq \text{TH NNR}$$

$$\text{MOD EMP}(i) = \text{NNR}(i) - \text{TH NNR}, \text{ cuando } \text{EMP} \text{ NNR}(i) > \text{TH NNR}$$

65 donde $\text{EMP}(i)$ es la magnitud de mejora de cada componente de frecuencia (dB);

MOD EMP (i) es la magnitud de mejora corregida de cada componente de frecuencia (dB);

NNR(i) es la NNR de cada componente de frecuencia (dB); y

5 TH NNR es el límite inferior de la NNR (dB).

El rango posible de cada parámetro de las ecuaciones es el siguiente.

10 TH NNR: de -50 a 50 dB, el límite inferior de la NNR proporcionado para evitar que el ruido superpuesto en el sonido de voz recibido sea desagradable al oído debido a la mejora.

A continuación, usando la magnitud de mejora corregida de cada componente de frecuencia MOD EMP (i), la unidad de procesado de espectro de sonido de voz 6' calcula un coeficiente emp coef(i) a aplicar a los espectros de cada componente de frecuencia.

15 [Ecuación 9]

$$\text{emp coef}(i) = 10^{\frac{\text{MODEMP}(i)}{20}}$$

20 A continuación, la parte real y la parte imaginaria de los espectros de cada componente de frecuencia son multiplicados por el coeficiente correspondiente emp coef(i), procesando por ello los espectros de cada componente de frecuencia.

[Ecuación 10]

25

$$\text{ESPE re}(i) = \text{SPE re}(i) * \text{emp coef}(i)$$

$$\text{ESPE im}(i) = \text{SPE im}(i) * \text{emp coef}(i)$$

$$\text{emp coef}(i) = 10^{\frac{\text{EMP}(i)}{20}}$$

MOD EMP (i): la magnitud de mejora corregida de cada componente de frecuencia (dB)

i: el índice de frecuencia

30

SPE re(i): la parte real de los espectros de cada componente de frecuencia

SPE im(i): la parte imaginaria de los espectros de cada componente de frecuencia

35 ESPE re(i): la parte real de los espectros mejorados de cada componente de frecuencia

ESPE im(i): la parte imaginaria de los espectros mejorados de cada componente de frecuencia

40 emp coef(i): el factor de multiplicación de los espectros de cada componente de frecuencia

En la tercera realización, la unidad de cálculo NNR 11 se añade a la estructura de la primera realización con el fin de calcular una NNR, que es una relación de ruido ambiente a ruido de fondo superpuesto en el sonido de voz. Consiguientemente, la unidad de procesado de espectro de sonido de voz 6' calcula la magnitud de mejora de segundo rango de frecuencias, tomando en consideración la NNR. Como resultado, es posible controlar la magnitud de mejora de modo que el ruido de fondo del sonido de voz no sea excesivo después de la mejora.

45

[Cuarta realización]

La figura 4 es una ilustración que representa una estructura de la cuarta realización de la presente descripción, y los componentes que son comunes a los de la primera realización reciben los mismos números de referencia. El número de referencia 1 denota la unidad de análisis de frecuencia; el número de referencia 9 denota la unidad de estimación de espectro de sonido de voz; el número de referencia 10 denota la unidad de estimación de espectro de ruido; el número de referencia 11 denota la unidad de cálculo NNR; el número de referencia 8 denota la unidad de cálculo de articulación de sonido de voz; el número de referencia 3 denota la unidad de cálculo SNR; el número de referencia 4

50

es la unidad de cálculo de magnitud de mejora de primer rango de frecuencias; el número de referencia 5' denota la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias; el número de referencia 6' denota la unidad de procesado de espectro de sonido de voz; y el número de referencia 7 es la unidad sintetizadora.

5 La estructura ilustrada en la figura 4 se explica a continuación según un proceso de señal audio. La operación detallada de cada componente estructural se describe más adelante.

En primer lugar, la unidad de análisis de frecuencia 1 introduce ruido ambiente, y envía espectros del ruido ambiente.

10 La unidad de estimación de espectro de sonido de voz 9 introduce sonido de voz recibido, y genera y envía espectros del sonido de voz recibido del que se ha quitado ruido de fondo superpuesto en el sonido de voz recibido.

La unidad de estimación de espectro de ruido 10 introduce el sonido de voz recibido, y genera y envía espectros del ruido de fondo superpuesto en el sonido de voz recibido.

15 La unidad de cálculo NNR 11 introduce los espectros del ruido ambiente salidos de la unidad de análisis de frecuencia 1 y los espectros del ruido de fondo del sonido de voz recibido salidos de la unidad de estimación de espectro de ruido 10, y calcula y envía una NNR, que es una relación de ruido ambiente a ruido de fondo superpuesto en el sonido de voz.

20 La unidad de cálculo de articulación de sonido de voz 8 introduce los espectros del sonido de voz recibido salidos de la unidad de análisis de frecuencia 2 e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido, y después envía información que indica articulación del sonido de voz.

25 La unidad de cálculo SNR 3 introduce los espectros del ruido ambiente salidos de la unidad de análisis de frecuencia 1 y los espectros del sonido de voz, del que se ha quitado ruido de fondo, salidos de la unidad de estimación de espectro de sonido de voz 9, y calcula y envía una SNR.

30 La unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 introduce la SNR salida de la unidad de cálculo SNR e información de división de rango de frecuencias, que indica un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias obtenido preliminarmente que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido. Posteriormente, la unidad de cálculo de magnitud de mejora de primer rango de frecuencias 4 calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la inteligibilidad subjetiva del sonido de voz recibido y la envía como magnitud de mejora de primer rango de frecuencias.

35 La unidad de cálculo de magnitud de mejora de segundo rango de frecuencias 5' introduce la magnitud de mejora de primer rango de frecuencias y la información que indica articulación del sonido de voz, y calcula la magnitud de mejora aplicada al rango de frecuencias que contribuye a la mejora de la articulación subjetiva del sonido de voz recibido y la envía como magnitud de mejora de segundo rango de frecuencias.

40 La unidad de procesado de espectro de sonido de voz 6' introduce la información de división de rango de frecuencias, la magnitud de mejora de primer rango de frecuencias, la magnitud de mejora de segundo rango de frecuencias, la NNR y los espectros del sonido de voz, y envía espectros de sonido de voz mejorado.

45 La unidad sintetizadora 7 introduce los espectros de sonido de voz mejorado salidos de la unidad de procesado de espectro de sonido de voz 6, y envía sonido de voz mejorado.

50 El proceso para mejorar el sonido de voz recibido se realiza según el flujo de procesado anterior.

55 La estructura de la presente realización se logra combinando componentes estructurales de las realizaciones segunda y tercera que se han modificado con respecto a los de la primera realización. Las operaciones de procesado de los respectivos componentes estructurales de las presentes realizaciones son las descritas en las realizaciones primera a tercera.

[Quinta realización]

60 La quinta realización es un ejemplo donde una estructura para calcular información de división de rango de frecuencias, que indica un rango de frecuencias que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido, se añade al dispositivo de mejora de sonido de voz de la cuarta realización. Los componentes estructurales añadidos son una unidad de cálculo de frecuencia tonal 12 configurada para introducir sonido de voz recibido y calcular y enviar una frecuencia tonal; y una unidad de cálculo de información de división de rango de frecuencias 13 configurada para introducir la frecuencia tonal e información que indica articulación del sonido de voz y generar la

información de división de rango de frecuencias.

5 El proceso de mejora de sonido de voz de la presente realización es el mismo que el de la cuarta realización, a excepción de generar la información de división de rango de frecuencias en el dispositivo de mejora de sonido de voz.

A continuación se describen en detalle estructuras de los componentes estructurales añadidos de la quinta realización.

10 La unidad de cálculo de frecuencia tonal 12 introduce el sonido de voz recibido y calcula una frecuencia tonal según las ecuaciones siguientes.

[Ecuación 11]

$$\text{corr}(a) = \frac{\sum_{i=0}^{M-1} x(i-a)x(i)}{\sqrt{\sum_{i=0}^{M-1} x(i-a)^2} \sqrt{\sum_{i=0}^{M-1} x(i)^2}}$$

15

Tono = $\text{frec}/a \text{ max}$

20 x: la señal de entrada

M: la longitud de un período durante el que se calcula un coeficiente de correlación (muestra)

a: la posición de inicio de una señal para la que se calcula el coeficiente de correlación

25 tono: la frecuencia tonal (Hz)

corr(a): el coeficiente de correlación cuando la posición desplazada es a

30 a max: a correspondiente al coeficiente de correlación máxima

i: el índice de frecuencia de la señal (muestra)

frec: la frecuencia de muestreo (Hz)

35 La unidad de cálculo de información de división de rango de frecuencias 13 introduce la frecuencia tonal y la información que indica articulación del sonido de voz, y genera información de división de rango de frecuencias según uno de los métodos siguientes. «Ejemplo específico 1 del cálculo de información de división de rango de frecuencias realizado por unidad de cálculo de información de división de rango de frecuencias 13»

40 En el ejemplo específico 1, la información de división de rango de frecuencias se calcula en el orden siguiente.

(a) Se calcula la frecuencia tonal del sonido de voz recibido.

45 (b) Se calcula la información de división de rango de frecuencias.

El método de cálculo se describe a continuación en detalle.

50 El índice de frecuencia límite l_b entre los rangos de frecuencias primero y segundo se calcula con la ecuación siguiente. Se dan valores de referencia para l_s y P_s . Cuanto más alta es la frecuencia tonal de la voz recibida, más probable es que el sonido de voz mejorado sea agudo si se mejora el segundo rango de frecuencias de la voz recibida. Consiguientemente, cuando la frecuencia tonal de la voz recibida es más alta que su valor de referencia, se asigna un valor mayor a l_b . De esta forma, cuanto más alta es la frecuencia tonal de la voz recibida, más se eleva el índice de frecuencia de límite inferior del segundo rango de frecuencias. Como resultado, la anchura del segundo rango de frecuencias es estrecha, por lo que es posible evitar que el sonido de voz mejorado sea agudo cuando se realice la mejora.

55

$$I_b = I_s + \alpha (p - p_s)$$

I_s : el valor de referencia del índice de frecuencia límite entre los rangos de frecuencias primero y segundo

5 I_b : el índice de frecuencia límite entre los rangos de frecuencias primero y segundo

p : la frecuencia tonal del sonido de voz recibido

10 p_s : el valor de referencia de la frecuencia tonal del sonido de voz recibido (50 a 500 Hz)

α : valor positivo

Los parámetros p_s y α deberán estar en el rango siguiente.

15 p_s : de 50 a 500 Hz, el tono de voz normal

α : de 0 a 1, la tasa de contribución de la frecuencia tonal usada para determinar I_b en base a un desplazamiento del tono del sonido de voz recibido con respecto al tono de voz normal

20 «Ejemplo específico 2 del cálculo de información de división de rango de frecuencias realizado por la unidad de cálculo de información de división de rango de frecuencias 13»

En el ejemplo específico 2, la información de división de rango de frecuencias se calcula en el orden siguiente.

25 (a) Se calcula el espectro de potencia del sonido de voz recibido.

(b) Se calcula la información de división de rango de frecuencias.

El método de cálculo se describe a continuación en detalle.

30 El índice de frecuencia límite I_b entre los rangos de frecuencias primero y segundo se calcula con la ecuación siguiente. Se dan valores de referencia para I_s y p_s . Cuanto menor es la pendiente del espectro de potencia del sonido de voz recibido, más probable es que el sonido de voz mejorado sea agudo si se mejora el segundo rango de frecuencias de la voz recibida. Consiguientemente, cuando la pendiente del espectro de potencia del sonido de voz recibido es menor que su valor de referencia, se asigna un valor mayor a I_b . De esta forma, cuanto menor es la pendiente del espectro de potencia del sonido de voz recibido, más se eleva el índice de frecuencia de límite inferior del segundo rango de frecuencias. Como resultado, la anchura del segundo rango de frecuencias es estrecha, por lo que es posible evitar que el sonido de voz mejorado sea agudo cuando se realice la mejora.

$$I_b = I_s + \beta (q - q_s)$$

I_s : el valor de referencia del índice de frecuencia límite entre los rangos de frecuencias primero y segundo

45 I_b : el índice de frecuencia límite entre los rangos de frecuencias primero y segundo

q : la pendiente del espectro de potencia del sonido de voz recibido

q_s : el valor de referencia de la pendiente del espectro de potencia del sonido de voz recibido

50 β : valor negativo

Los parámetros q_s y β deberán estar en el rango siguiente.

55 q_s : de 50 a 500 Hz, el rango de tono de voz normal

β : de -100 a 0, la tasa de contribución de la pendiente del espectro de potencia del sonido de voz recibido usado para determinar I_b en base a un desplazamiento de la pendiente del espectro de potencia del sonido de voz recibido con respecto a la pendiente del espectro de potencia de voz normal

60 «Ejemplo específico 3 del cálculo de información de división de rango de frecuencias realizado por la unidad de cálculo de información de división de rango de frecuencias 13»

En el ejemplo específico 3, la información de división de rango de frecuencias se calcula en el orden siguiente.

65 (a) Se calcula la frecuencia tonal del sonido de voz recibido.

(b) Se calcula el espectro de potencia del sonido de voz recibido.

(c) Se calcula la información de división de rango de frecuencias.

5 El método de cálculo se describe a continuación en detalle.

10 El índice de frecuencia límite l_b entre los rangos de frecuencias primero y segundo se calcula con la ecuación siguiente. Se dan valores de referencia para l_s , p_s y q_s . Cuanto más alta es la frecuencia tonal del sonido de voz recibido y cuanto menor es la pendiente del espectro de potencia del sonido de voz recibido, más probable es que el sonido de voz mejorado sea agudo si se mejora el segundo rango de frecuencias de la voz recibida. Consiguientemente, cuando la frecuencia tonal del sonido de voz recibido es más alta que el valor de referencia de frecuencia tonal y cuando la pendiente del espectro de potencia del sonido de voz recibido es menor que el valor de referencia de pendiente de espectro de potencia, se asigna un valor mayor a l_b . De esta forma, cuanto más alta es la frecuencia tonal del sonido de voz recibido y cuanto menor es la pendiente del espectro de potencia del sonido de voz recibido, más se eleva el índice de frecuencia de límite inferior del segundo rango de frecuencias. Como resultado, la anchura del segundo rango de frecuencias es estrecha, por lo que es posible evitar que el sonido de voz mejorado sea agudo cuando se realice la mejora.

20
$$l_b = l_s + \alpha (p - p_s) + \beta (q - q_s)$$

l_s : el valor de referencia del índice de frecuencia límite entre los rangos de frecuencias primero y segundo

l_b : el índice de frecuencia límite entre los rangos de frecuencias primero y segundo

25 p : la frecuencia tonal del sonido de voz recibido

p_s : el valor de referencia de la frecuencia tonal (el valor de referencia de frecuencia tonal) del sonido de voz recibido

30 α : valor positivo

q : la pendiente del espectro de potencia del sonido de voz recibido

q_s : el valor de referencia de la pendiente del espectro de potencia (el valor de referencia de pendiente de espectro de potencia) del sonido de voz recibido

35 β : valor negativo

40 Estos rangos posibles de los parámetros son los mismos que los de los ejemplos específicos 1 y 2 de la presente realización.

La "pendiente del espectro de potencia" de los ejemplos específicos 1 a 3 anteriores se calcula de la siguiente manera.

45 (1) El espectro de potencia se calcula con la ecuación siguiente usando los espectros.

$$PSP(i) = 10\log_{10}(SP_{re}(i) * SP_{re}(i) + SP_{im}(i) * SP_{im}(i))$$

$PSP(i)$: el espectro de potencia

50 i : el índice de frecuencia

$SP_{re}(i)$: la parte real de los espectros

55 $SP_{im}(i)$: la parte imaginaria de los espectros

(2) Se calcula la pendiente del espectro de potencia.

60 El espectro de potencia para cada índice de frecuencia se expresa como $(l, PSP(i))$, y la pendiente del espectro de potencia se calcula aplicando una función lineal usando el método de los cuadrados mínimos.

La quinta realización anterior se describe como un ejemplo donde los componentes estructurales únicos de la quinta realización se añaden a la estructura de la cuarta realización. Sin embargo, añadiendo los componentes estructurales únicos de la quinta realización a la estructura de cualquiera de las realizaciones primera a tercera, la información de división de rango de frecuencias puede ser generada a partir del sonido de voz recibido de forma similar.

65

5 Así, además de los efectos logrados con las realizaciones primera a cuarta, la quinta realización es capaz de controlar los rangos de frecuencias primero y segundo según el sonido de voz haciendo que la unidad de cálculo de información de división de rango de frecuencias 13 calcule información de división de rango de frecuencias usando el sonido de voz recibido. Consiguientemente, es posible lograr un control adecuado de la mejora de sonido de voz según el sonido de voz.

10 Obsérvese que la información de división de rango de frecuencias usada en las realizaciones primera a cuarta se puede obtener estadísticamente como un rango de frecuencias que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido y un rango de frecuencias que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido. Así, los resultados obtenidos pueden ser usados como la información de división de rango de frecuencias.

15 Además, se pueden facilitar una unidad de mejora de primer rango de frecuencias y una unidad de mejora de segundo rango de frecuencias con el fin de realizar la mejora de sonido de voz solamente en rangos de frecuencia preliminarmente especificados por la información de división de rango de frecuencias. En este caso, el recorrido de entrada de la información de división de rango de frecuencias es innecesario, y se pueden lograr los mismos efectos que los de las realizaciones primera a cuarta anteriores.

20 En resumen, según un aspecto de la presente realización, en el caso donde el sonido de voz recibido es difícil de oír debido a ruido ambiente, se genera sonido de voz mejorado que tiene una calidad suficiente tanto en inteligibilidad subjetiva como en articulación, y, por lo tanto, es capaz de resolver los problemas de las tecnologías convencionales de mejora de sonido de voz -sonido de voz apagado (es decir, como resultado de la mejora de sonido de voz, la magnitud de un rango de frecuencias que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido es demasiado pequeña en comparación con la magnitud de un rango de frecuencias que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido, y por lo tanto, el sonido de voz mejorado carece de articulación) y sonido de voz agudo (como resultado de la mejora de sonido de voz, la magnitud del rango de frecuencias que contribuye a una mejora de la articulación subjetiva del sonido de voz recibido es demasiado grande en comparación con la magnitud del rango de frecuencias que contribuye a una mejora de la inteligibilidad subjetiva del sonido de voz recibido, y por lo tanto, la articulación es excesiva en el sonido de voz mejorado).
 25
 30 Consiguientemente, es posible generar el sonido de voz mejorado fácil de escuchar.

35 Todos los ejemplos y el lenguaje condicional aquí usados tienen fines pedagógicos para ayudar al lector a comprender la presente descripción y los conceptos aportados por el inventor para hacer progresar la técnica, y han de ser interpretados sin limitación a dichos ejemplos y condiciones específicamente expuestos, tampoco la organización de tales ejemplos en la memoria descriptiva se refiere a una representación de la superioridad o inferioridad de la presente descripción. Aunque las realizaciones de la presente descripción se han descrito en detalle, se deberá entender que se podrían hacer en ella varios cambios, sustituciones y alteraciones a condición de que caigan dentro del alcance de las reivindicaciones anexas.

40

REIVINDICACIONES

1. Un dispositivo de mejora de sonido de voz incluyendo:

5 una unidad de cálculo SNR (3) configurada para calcular una SNR que es una relación de la magnitud de un sonido de voz recibido a la magnitud de un ruido ambiente;

10 una unidad de cálculo de información de división de rango de frecuencias (13) configurada para calcular, en base a una frecuencia tonal del sonido de voz recibido, una información de división de rango de frecuencias, indicando un primer rango de frecuencias y un segundo rango de frecuencias, donde cada uno de los rangos de frecuencias primero y segundo cubre diferentes valores de frecuencia;

15 una unidad de cálculo de magnitud de mejora de primer rango de frecuencias (4) configurada para calcular, en base a la SNR y la información de división de rango de frecuencias, la magnitud de mejora del primer rango de frecuencias,

20 una unidad de cálculo de magnitud de mejora de segundo rango de frecuencias (5, 5') configurada para calcular, en base a la magnitud de mejora del primer rango de frecuencias, la magnitud de mejora del segundo rango de frecuencias; y

una unidad de procesamiento de espectro (6, 6') configurada para procesar espectros del sonido de voz recibido usando la magnitud de mejora del primer rango de frecuencias y la magnitud de mejora del segundo rango de frecuencias.

25 2. El dispositivo de mejora de sonido de voz según la reivindicación 1, incluyendo además una unidad de cálculo de articulación de sonido de voz (8) configurada para calcular la articulación del sonido de voz recibido en base a al menos uno de una frecuencia tonal del sonido de voz recibido y un espectro de potencia del sonido de voz recibido, donde la articulación corresponde a la relación entre la potencia del segundo rango de frecuencias y la potencia del primer rango de frecuencias, donde la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias (5') calcula la magnitud de mejora del segundo rango de frecuencias en base a la magnitud de mejora del primer rango de frecuencias y la articulación calculada del sonido de voz recibido, y la unidad de procesamiento de espectro (6) procesa los espectros del sonido de voz recibido usando la magnitud de mejora del primer rango de frecuencias y la magnitud de mejora del segundo rango de frecuencias.

35 3. El dispositivo de mejora de sonido de voz según la reivindicación 1, incluyendo además una unidad de cálculo NNR (11) configurada para calcular una NNR que es una relación del ruido ambiente a ruido de fondo superpuesto en el sonido de voz recibido, donde la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias (5; 5') calcula la magnitud de mejora del segundo rango de frecuencias en base a la magnitud de mejora del primer rango de frecuencias y la NNR, y la unidad de procesamiento de espectro (6) procesa los espectros del sonido de voz recibido usando la magnitud de mejora del primer rango de frecuencias y la magnitud de mejora del segundo rango de frecuencias.

45 4. El dispositivo de mejora de sonido de voz según la reivindicación 1, incluyendo además: una unidad de cálculo de articulación de sonido de voz (8) configurada para calcular la articulación del sonido de voz recibido en base a al menos uno de una frecuencia tonal del sonido de voz recibido y un espectro de potencia del sonido de voz recibido, donde la articulación corresponde a la relación entre la potencia del segundo rango de frecuencias y la potencia del primer rango de frecuencias; y una unidad de cálculo NNR (11) configurada para calcular una NNR que es una relación del ruido ambiente a ruido de fondo superpuesto en el sonido de voz recibido; donde la unidad de cálculo de magnitud de mejora de primer rango de frecuencias (4) calcula la magnitud de mejora del primer rango de frecuencias en base a la SNR, la unidad de cálculo de magnitud de mejora de segundo rango de frecuencias (5') calcula la magnitud de mejora del segundo rango de frecuencias en base a la magnitud de mejora del primer rango de frecuencias, la NNR y la articulación calculada de la voz recibida, y la unidad de procesamiento de espectro (6') procesa los espectros del sonido de voz recibido usando la magnitud de mejora del primer rango de frecuencias y la magnitud de mejora del segundo rango de frecuencias.

55 5. El dispositivo de mejora de sonido de voz según la reivindicación 1, donde la unidad de cálculo de información de división de rango de frecuencias (13) está configurada para calcular la información de división de rango de frecuencias en el caso de que el sonido de voz recibido esté enmascarado por el ruido ambiente.

60 6. El dispositivo de mejora de sonido de voz según la reivindicación 2, donde, en el cálculo de la articulación del sonido de voz, una magnitud de una componente de sonido de voz recibido es multiplicada por un coeficiente calculado usando una curva isofónica.

65 7. El dispositivo de mejora de sonido de voz según la reivindicación 4, donde, en el cálculo de la articulación del sonido de voz, una magnitud de una componente de sonido de voz recibido es multiplicada por un coeficiente calculado usando una curva isofónica.

8. El dispositivo de mejora de sonido de voz según la reivindicación 3, donde, en el cálculo de la NNR, una magnitud de una componente de sonido de voz recibido es multiplicada por un coeficiente calculado usando una curva isofónica.
- 5 9. El dispositivo de mejora de sonido de voz según la reivindicación 4, donde, en el cálculo de la NNR, una magnitud de una componente de sonido de voz recibido es multiplicada por un coeficiente calculado usando una curva isofónica.
- 10 10. El dispositivo de mejora de sonido de voz según la reivindicación 1, donde, en el cálculo de la SNR, una magnitud de una componente de sonido de voz recibido es multiplicada por un coeficiente calculado usando una curva isofónica.
- 15 11. El dispositivo de mejora de sonido de voz según la reivindicación 1, donde un límite superior del primer rango de frecuencias es igual o menor que 3000 Hz.
12. El dispositivo de mejora de sonido de voz según la reivindicación 1, donde un límite inferior del segundo rango de frecuencias es igual o mayor que 500 Hz.

FIG.1

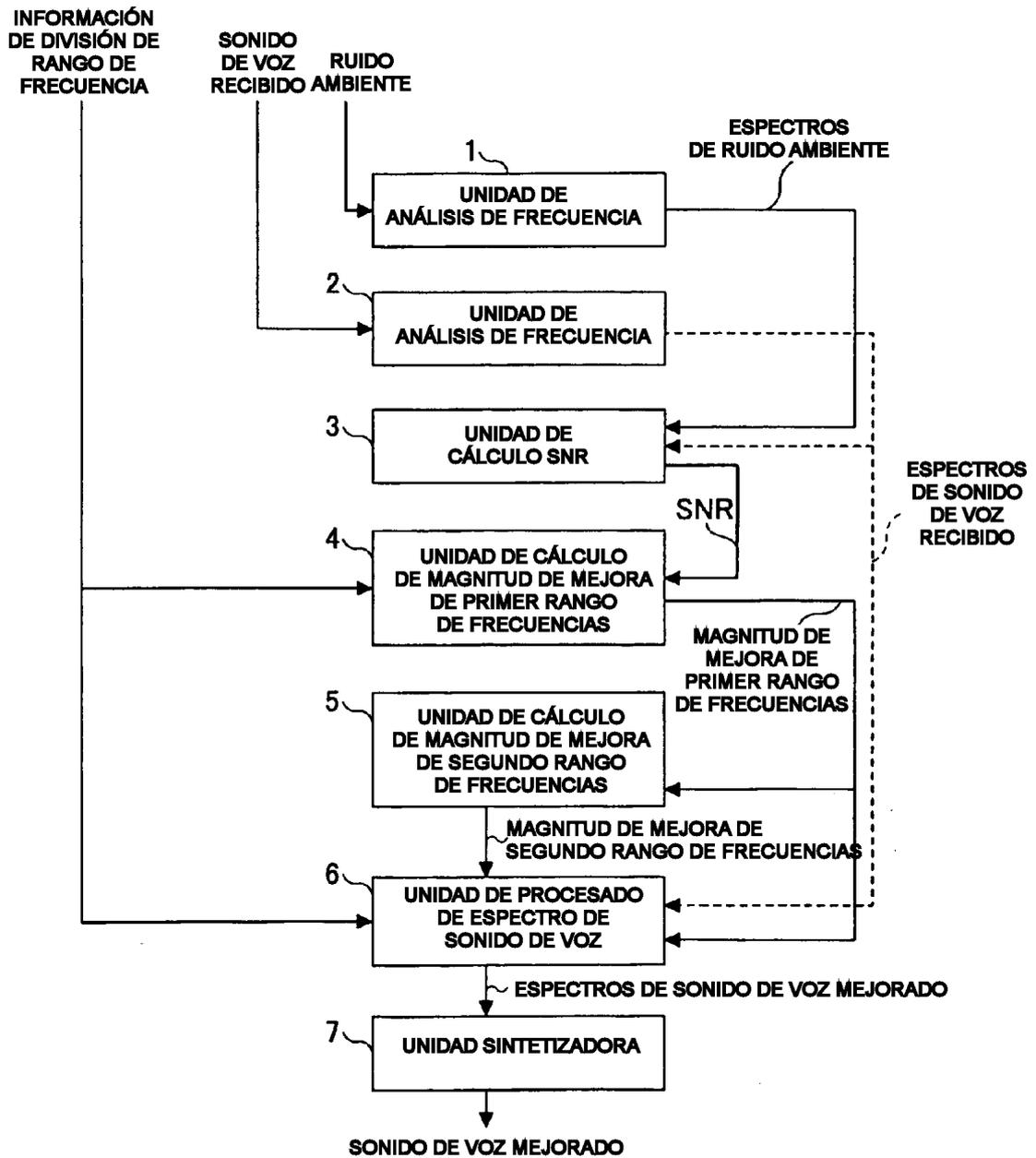


FIG.2

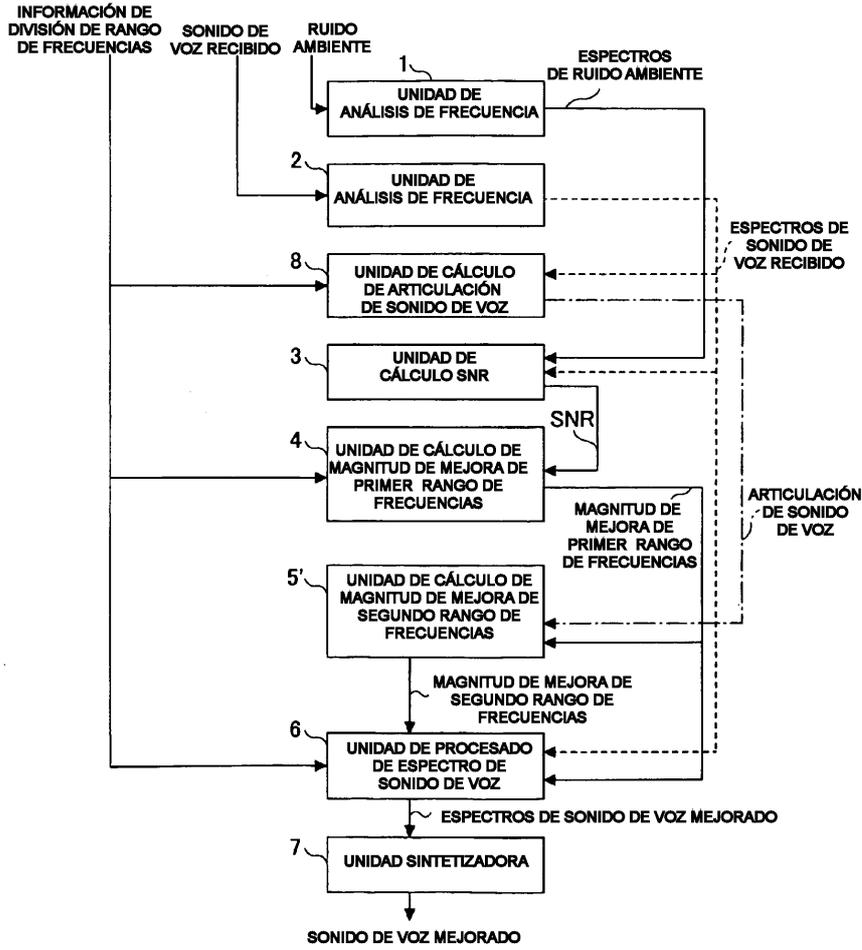


FIG.3

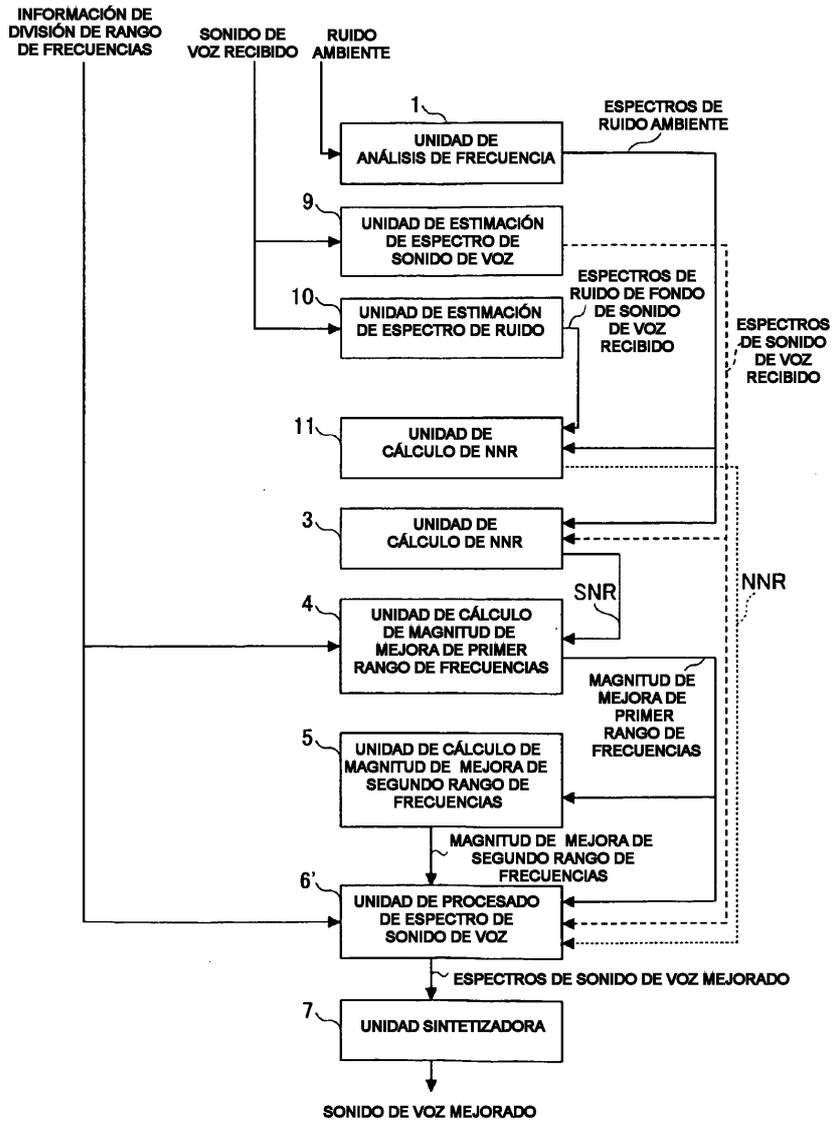


FIG.4

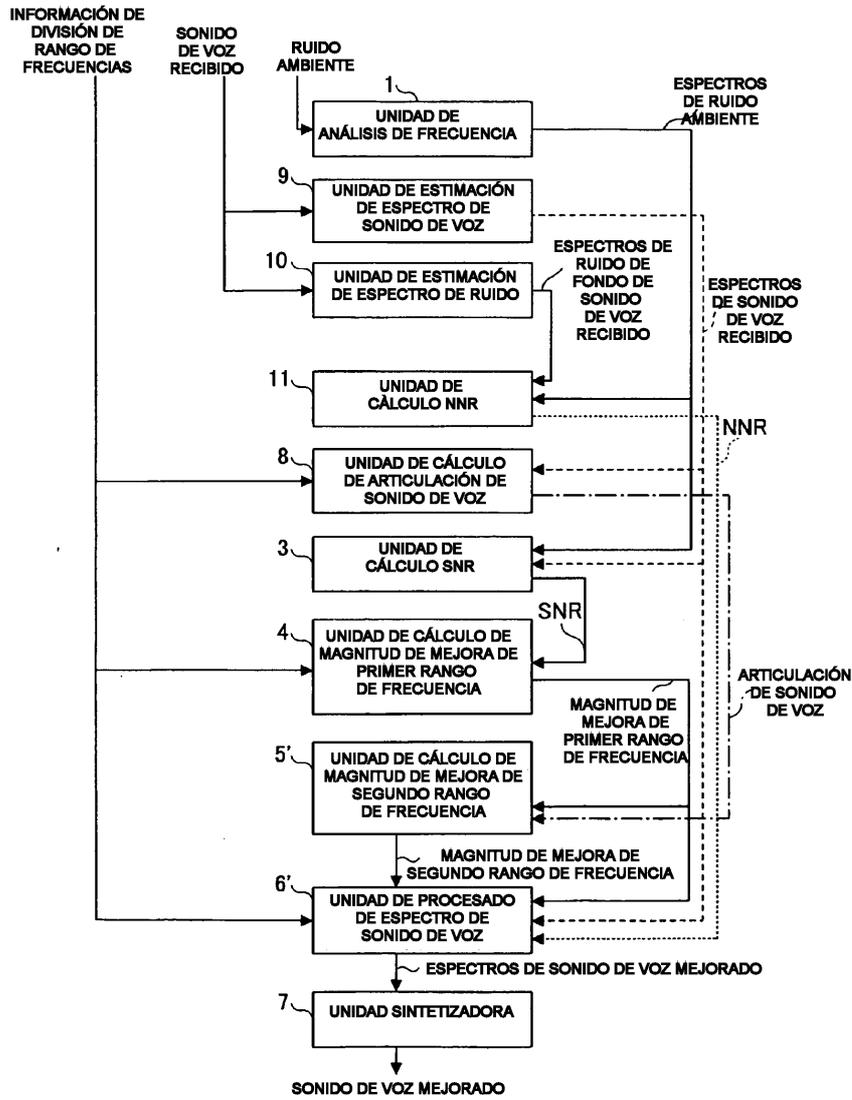


FIG.5

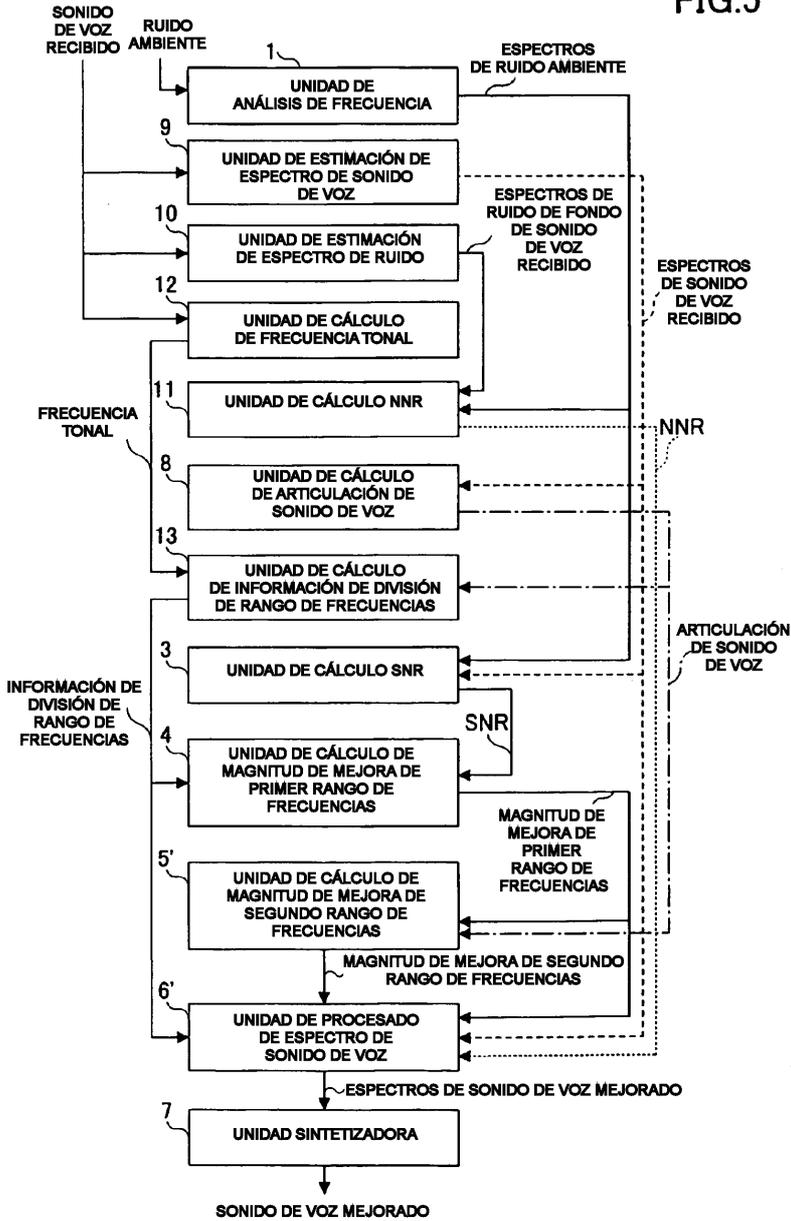


FIG.6

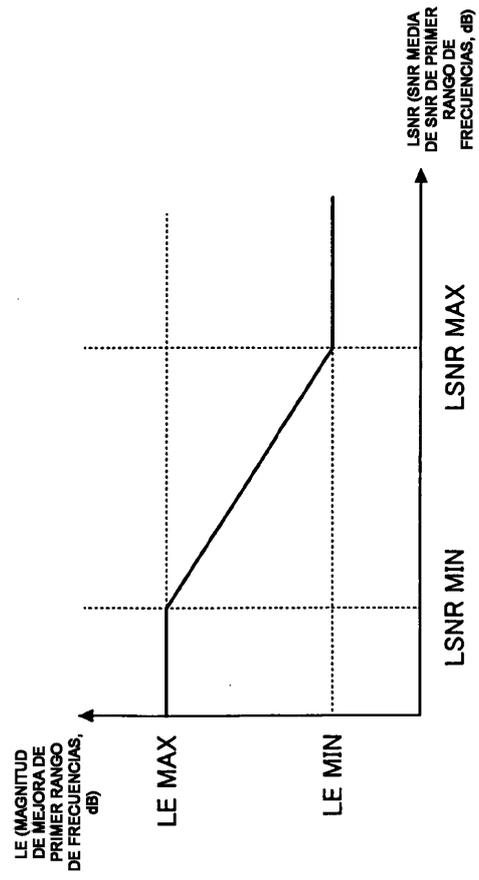


FIG.7

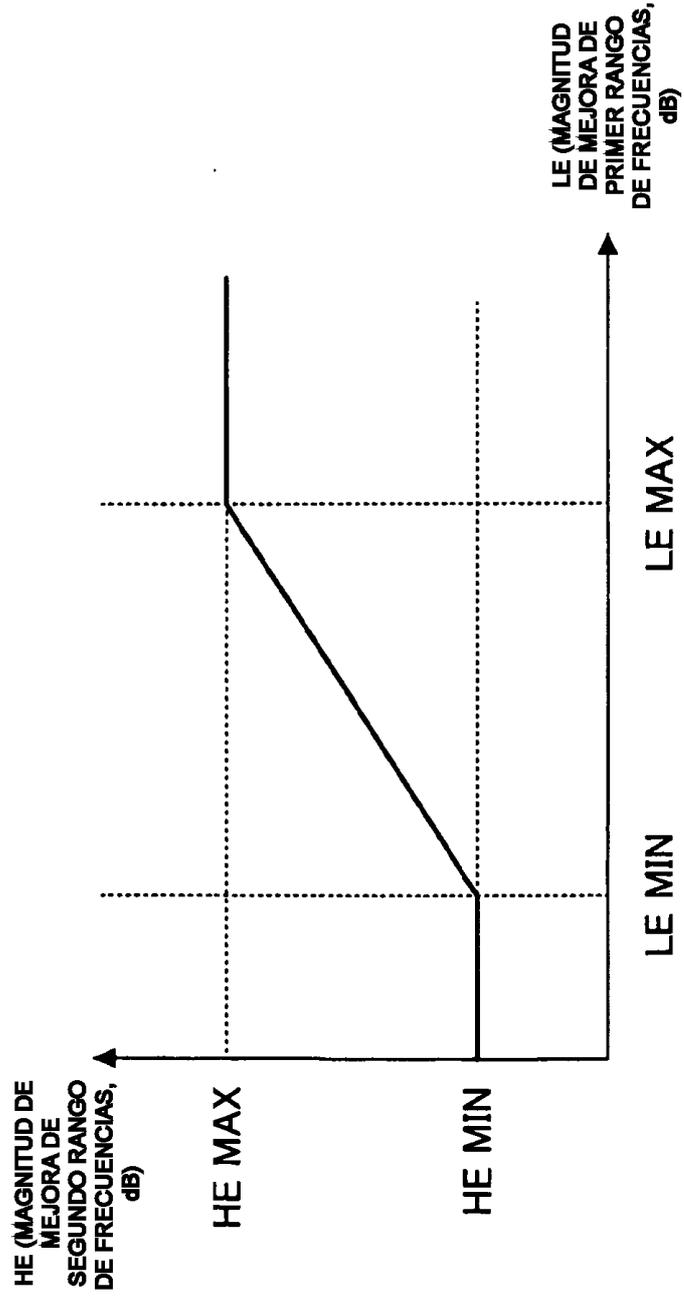


FIG.8

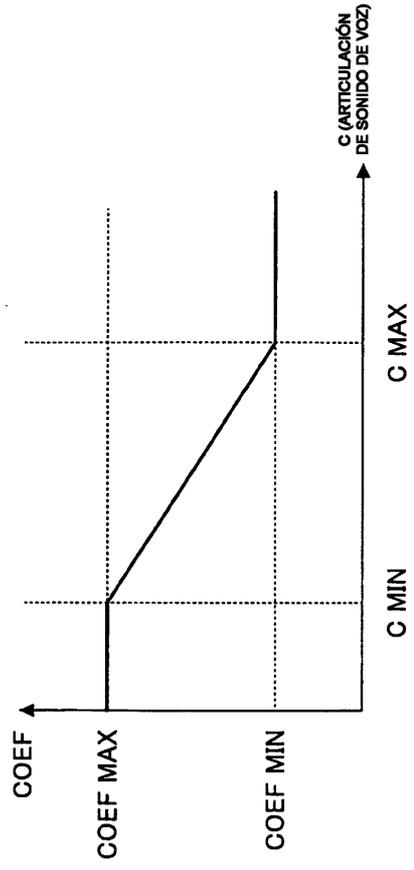


FIG.9

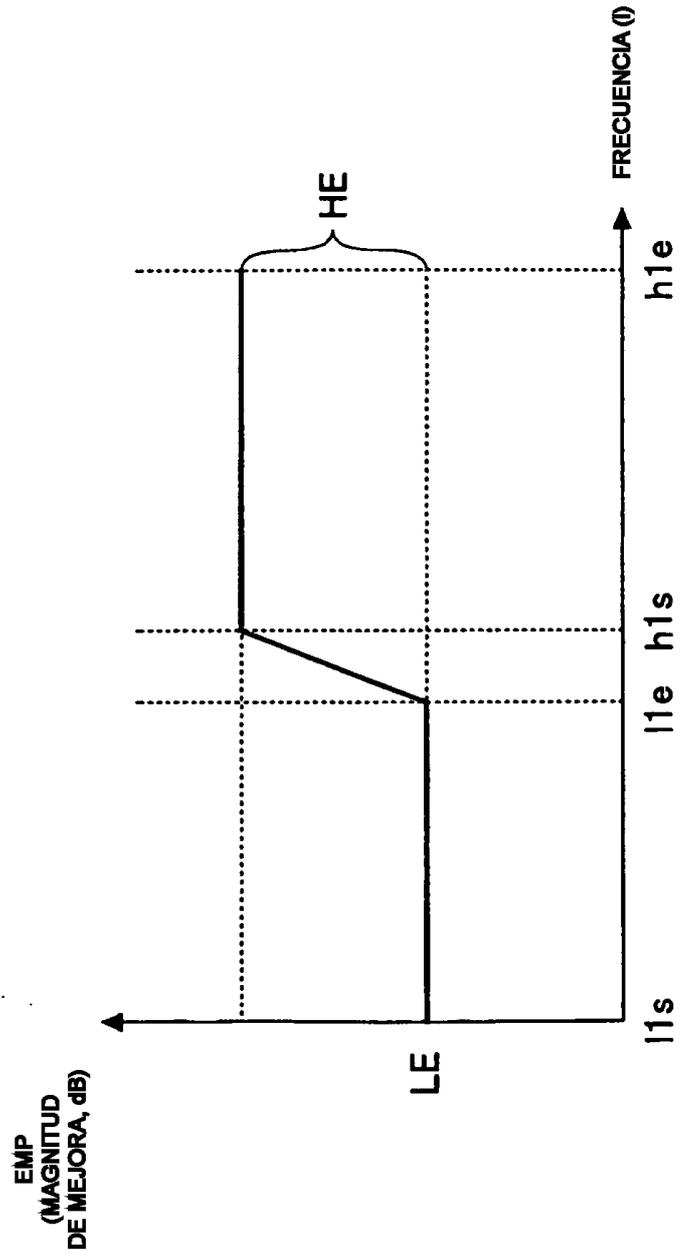


FIG.10

