

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 741 849**

51 Int. Cl.:

G10L 19/00 (2013.01)

G10L 19/24 (2013.01)

G10L 21/02 (2013.01)

G10L 19/083 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.03.2013 E 16187948 (1)**

97 Fecha y número de publicación de la concesión europea: **15.05.2019 EP 3193331**

54 Título: **Procedimiento y aparato de procesamiento de señales de voz/audio**

30 Prioridad:

01.03.2012 CN 201210051672

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

12.02.2020

73 Titular/es:

**HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian,
Longgang District
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:

**LIU, ZEXIN y
MIAO, LEI**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 741 849 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

Procedimiento y aparato de procesamiento de señales de voz/audio

5 Campo técnico

La presente invención se refiere al campo de las tecnologías de procesamiento de señales digitales y, en particular, a un procedimiento y aparato de procesamiento de señales de voz/audio.

10 Antecedentes

En el campo de las comunicaciones digitales, la transmisión de voz, imágenes, audio y vídeos es necesaria en una gran variedad de aplicaciones tales como llamadas de teléfono móvil, conferencias de audio/vídeo, televisión mediante radiodifusión y el entretenimiento multimedia. El audio se digitaliza y se transmite desde un terminal a otro usando una red de comunicaciones de audio. En el presente documento, el terminal puede ser un teléfono móvil, un terminal de teléfono digital o un terminal de audio de cualquier otro tipo, donde el terminal de teléfono digital es, por ejemplo, un teléfono VOIP, un teléfono ISDN, un ordenador o un teléfono de comunicaciones por cable. Para reducir los recursos ocupados por una señal de voz/audio durante el almacenamiento o la transmisión, la señal de voz/audio se comprime en un extremo de transmisión y después se transmite a un extremo de recepción, y en el extremo de recepción, la señal de voz/audio se restaura mediante un procesamiento de descompresión y se reproduce.

En la codificación de voz/audio de velocidad múltiple, debido a los diferentes estados de red, una red trunca flujos de bits a diferentes velocidades binarias, donde los flujos de bits se transmiten desde un codificador a la red y, en un decodificador, los flujos de bits truncados se descodifican en señales de voz/audio de diferentes anchos de banda. Como resultado, las señales de voz/audio de salida conmutan entre diferentes anchos de banda.

Una conmutación repentina entre señales de diferentes anchos de banda genera molestias auditivas perceptibles en el oído humano. Además, debido a que la actualización de los estados de los filtros durante la transformación tiempo-frecuencia o la transformación frecuencia-tiempo requiere generalmente el uso de un parámetro entre tramas consecutivas, cuando no se lleva a cabo un procesamiento apropiado durante la conmutación de ancho de banda, puede producirse un error durante la actualización de estos estados, lo que provoca ciertos fenómenos de cambios de energía abruptos y el deterioro de la calidad auditiva.

Un ejemplo de un codificador de voz y audio de banda ancha escalable conocido se da a conocer en el documento US 2011/0270614 A1.

Resumen

Un objetivo de la presente invención es proporcionar un procedimiento y aparato de procesamiento de señales de voz/audio como reivindicados en las reivindicaciones 1 y 9 respectivamente con el fin de mejorar el confort auditivo durante la conmutación de ancho de banda de señales de voz/audio. Formas de realización específicas se definen en las reivindicaciones subordinadas.

45 Breve descripción de los dibujos

Para describir con mayor claridad las soluciones técnicas de las formas de realización de la presente invención o de la técnica anterior, a continuación se introducen brevemente los dibujos adjuntos requeridos para describir las formas de realización o la técnica anterior. Evidentemente, los dibujos adjuntos de la siguiente descripción muestran simplemente algunas formas de realización de la presente invención, y un experto en la técnica puede obtener otros dibujos a partir de estos dibujos adjuntos sin realizar investigaciones adicionales.

La FIG. 1 es un diagrama de flujo esquemático de una forma de realización de un procedimiento de procesamiento de señales de voz/audio.

La FIG. 2 es un diagrama de flujo esquemático de otra forma de realización de un procedimiento de procesamiento de señales de voz/audio según la presente invención.

La FIG. 3 es un diagrama de flujo esquemático de otra forma de realización de un procedimiento de procesamiento de señales de voz/audio.

La FIG. 4 es un diagrama de flujo esquemático de otra forma de realización de un procedimiento de procesamiento de señales de voz/audio.

La FIG. 5 es un diagrama estructural esquemático de una forma de realización de un aparato de procesamiento de señales de voz/audio según la presente invención.

La FIG. 6 es un diagrama estructural esquemático de una forma de realización de un aparato de procesamiento de señales de voz/audio según la presente invención.

La FIG. 7 es un diagrama estructural esquemático de una forma de realización de una unidad de obtención de parámetros según la presente invención.

La FIG. 8 es un diagrama estructural esquemático de una forma de realización de una unidad de obtención de parámetros de ganancia global según la presente invención.

La FIG. 9 es un diagrama estructural esquemático de una forma de realización de una unidad de adquisición según la presente invención.

5 La FIG. 10 es un diagrama estructural esquemático de otra forma de realización de un aparato de procesamiento de señales de voz/audio según la presente invención.

Descripción de formas de realización

10 A continuación se describe de manera clara y completa las soluciones técnicas en las formas de realización de la presente invención con referencia a los dibujos adjuntos en las formas de realización de la presente invención. Evidentemente, las formas de realización descritas son simplemente una parte en lugar de todas las formas de realización de la presente invención. Todas las demás formas de realización obtenidas por un experto en la técnica tomando como base las formas de realización de la presente invención sin realizar investigaciones adicionales
15 estarán dentro del alcance de protección de la presente invención.

En el campo del procesamiento de señales digitales, los códecs de audio y los códecs de vídeo se aplican de manera generalizada en varios dispositivos electrónicos, por ejemplo un teléfono móvil, un aparato inalámbrico, un asistente de datos personal (PDA), un ordenador manual o portátil, un receptor/navegador GPS, una cámara, un
20 reproductor de audio/vídeo, una cámara de vídeo, una grabadora de vídeo y un dispositivo de supervisión. Normalmente, este tipo de dispositivo electrónico incluye un codificador de audio o un descodificador de audio, donde el codificador o descodificador de audio pueden implementarse directamente mediante un circuito o chip digital, por ejemplo un DSP (procesador de señales digitales) o implementarse mediante un código de software que hace que un procesador ejecute un proceso del código de software.

25 En la técnica anterior, puesto que los anchos de banda de señales de voz/audio transmitidas en una red son diferentes, en un proceso de transmisión de señales de voz/audio, los anchos de banda de las señales de voz/audio cambian con frecuencia y se produce el fenómeno de conmutar desde una señal de voz/audio de frecuencia estrecha a una señal de voz/audio de frecuencia ancha y de conmutar desde una señal de voz/audio de frecuencia
30 ancha a una señal de voz/audio de frecuencia estrecha. Este proceso de conmutar una señal de voz/audio entre bandas de alta y baja frecuencia se denomina conmutación de ancho de banda. La conmutación de ancho de banda incluye conmutar desde una señal de frecuencia estrecha a una señal de frecuencia ancha y conmutar desde una señal de frecuencia ancha a una señal de frecuencia estrecha. La señal de frecuencia estrecha mencionada en la presente invención es una señal de voz que solo tiene una componente de baja frecuencia, y una componente de
35 alta frecuencia está vacía después de un muestreo ascendente y un filtrado paso bajo, mientras que la señal de voz/audio de frecuencia ancha tiene tanto una componente de señal de baja frecuencia como una componente de señal de alta frecuencia. La señal de frecuencia estrecha y la señal de frecuencia ancha son relativas. Por ejemplo, para una señal de banda estrecha, una señal de banda ancha es una señal de frecuencia ancha; y para una señal de banda ancha, una señal de banda superancha es una señal de frecuencia ancha. Generalmente, una señal de
40 banda estrecha es una señal de voz/audio cuya velocidad de muestreo es de 8 kHz; una señal de banda ancha es una señal de voz/audio cuya frecuencia de muestreo es de 16 kHz; y una señal de banda superancha es una señal de voz/audio cuya frecuencia de muestreo es de 32 kHz.

45 Cuando un algoritmo de codificación/descodificación de una señal de alta frecuencia antes de la conmutación se selecciona entre algoritmos de codificación/descodificación de dominio de tiempo y de dominio de frecuencia según diferentes tipos de señal, o cuando un algoritmo de codificación de la señal de alta frecuencia antes de la conmutación es un algoritmo de codificación de dominio de tiempo, con el fin de garantizar la continuidad de las señales de salida durante la conmutación, un algoritmo de conmutación se mantiene en un dominio de señal para el procesamiento, donde el dominio de señal es el mismo que el del algoritmo de codificación/descodificación de alta
50 frecuencia antes de la conmutación. Es decir, cuando el algoritmo de codificación/descodificación de dominio de tiempo se usa para la señal de alta frecuencia antes de la conmutación, un algoritmo de conmutación de dominio de tiempo se usa como un algoritmo de conmutación que va a usarse; cuando el algoritmo de codificación/descodificación de dominio de frecuencia se usa para la señal de alta frecuencia antes de la conmutación, un algoritmo de conmutación de dominio de frecuencia se usa como un algoritmo de conmutación que va a usarse. En la técnica anterior, cuando un algoritmo de extensión de banda de frecuencia de dominio de tiempo se usa antes de la conmutación, una tecnología de conmutación de dominio de tiempo similar no se usa después de la conmutación.

55 En la codificación de voz/audio, el procesamiento se lleva a cabo generalmente usando una trama como unidad. Una trama de audio de entrada actual que necesita procesarse es una trama actual de señal de voz/audio. La trama actual de señal de voz/audio incluye una señal de frecuencia estrecha y una señal de alta frecuencia, es decir, una señal de frecuencia estrecha de la trama actual y una señal de alta frecuencia de la trama actual. Cualquier trama de señal de voz/audio antes de la trama actual de señal de alta frecuencia es una trama histórica de señal de voz/audio, que también incluye una trama histórica de señal de frecuencia estrecha y una trama histórica de señal de alta
60 frecuencia. Una trama de señal de voz/audio antes de la trama actual de señal de voz/audio es una trama anterior de señal de voz/audio.

Con referencia a la FIG. 1, una forma de realización de un procedimiento de procesamiento de señales de voz/audio incluye:

5 S101: Cuando una señal de voz/audio conmuta el ancho de banda, obtener una señal inicial de alta frecuencia correspondiente a una trama actual de señal de voz/audio.

10 La trama actual de señal de voz/audio incluye una señal de frecuencia estrecha de la trama actual y una señal de dominio de tiempo de alta frecuencia de la trama actual. La conmutación de ancho de banda incluye conmutar desde una señal de frecuencia estrecha a una señal de frecuencia ancha y conmutar desde una señal de frecuencia ancha a una señal de frecuencia estrecha. En caso de conmutar desde una señal de frecuencia estrecha a una señal de frecuencia ancha, la trama actual de señal de voz/audio es la señal de frecuencia ancha de la trama actual, que incluye una señal de frecuencia estrecha y una señal de alta frecuencia, y la señal inicial de alta frecuencia de la trama actual de señal de voz/audio es una señal real y puede obtenerse directamente a partir de la trama actual de la señal de voz/audio. En caso de conmutar desde una señal de frecuencia ancha a una señal de frecuencia estrecha, la trama actual de señal de voz/audio es la señal de frecuencia estrecha de la trama actual cuya señal de dominio de tiempo de alta frecuencia de la trama actual está vacía, la señal inicial de alta frecuencia de la trama actual de señal de voz/audio es una señal real predicha, y una señal de alta frecuencia correspondiente a la señal de frecuencia estrecha de la trama actual tiene que predecirse y usarse como la señal inicial de alta frecuencia.

20 S102: Obtener un parámetro de ganancia global de dominio de tiempo correspondiente a la señal inicial de alta frecuencia.

25 En caso de conmutar desde una señal de frecuencia estrecha a una señal de frecuencia ancha, el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia puede obtenerse mediante descodificación. En caso de conmutar desde una señal de frecuencia ancha a una señal de frecuencia estrecha, el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia puede obtenerse según la trama actual de señal: el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia se obtiene según un parámetro de inclinación de espectro de la señal de frecuencia estrecha y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica.

35 S103: Llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica de señal de voz/audio y la energía de la señal inicial de alta frecuencia de la trama actual de señal de voz/audio.

40 Una trama histórica de señal de salida final de voz/audio se usa como la trama histórica de señal de voz/audio, y la señal inicial de alta frecuencia se usa como la trama actual de señal de voz/audio. La relación de energía es $Relación = Esyn(-1) / Esyn_tmp$, donde $Esyn(-1)$ representa la energía de la señal de salida de dominio de tiempo de alta frecuencia syn de la trama histórica, y $Esyn_tmp$ representa la energía de la señal inicial de dominio de tiempo de alta frecuencia syn correspondiente a la trama actual.

45 El parámetro predicho de ganancia global es $ganancia = \alpha * Relación + \beta * ganancia'$, donde $ganancia'$ es el parámetro de ganancia global de dominio de tiempo, $\alpha + \beta = 1$ y los valores de α y β son diferentes según diferentes tipos de señal.

50 S104: Corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia.

55 La corrección se refiere a que la señal se multiplica, es decir, la señal inicial de alta frecuencia se multiplica por el parámetro predicho de ganancia global. En otra forma de realización, en la etapa S102, se obtiene un parámetro de envolvente de dominio de tiempo y el parámetro de ganancia global de dominio de tiempo correspondientes a la señal inicial de alta frecuencia; por lo tanto, en la etapa S104, la señal inicial de alta frecuencia se corrige usando el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global, para obtener la señal corregida de dominio de tiempo de alta frecuencia; es decir, la señal predicha de alta frecuencia se multiplica por el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global de dominio de tiempo con el fin de obtener la señal corregida de dominio de tiempo de alta frecuencia.

60 En caso de conmutar desde una señal de frecuencia estrecha a una señal de frecuencia ancha, el parámetro de envolvente de dominio de tiempo de la señal de alta frecuencia puede obtenerse mediante descodificación. En caso de conmutar desde una señal de frecuencia ancha a una señal de frecuencia estrecha, el parámetro de envolvente de dominio de tiempo de la señal de alta frecuencia puede obtenerse según la trama actual de señal: una serie de valores predeterminados o un parámetro de envolvente de dominio de tiempo de alta frecuencia de la trama histórica puede usarse como el parámetro de envolvente de dominio de tiempo de alta frecuencia de la trama actual de señal de voz/audio.

S105: Sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

5 En la forma de realización anterior, durante la conmutación entre una banda de frecuencia ancha y una banda de frecuencia estrecha, se corrige una señal de alta frecuencia con el fin de implementar una transición gradual de la señal de alta frecuencia entre la banda de frecuencia ancha y la banda de frecuencia estrecha, por lo que se elimina de este modo la molestia auditiva producida por la conmutación entre la banda de frecuencia ancha y la banda de frecuencia estrecha; además, puesto que un algoritmo de conmutación de ancho de banda y un algoritmo de codificación/descodificación de la señal de alta frecuencia antes de la conmutación están en un mismo dominio de
10 señal, no solo se garantiza que no se añada ningún retardo adicional y que el algoritmo sea sencillo, sino que además se garantiza el rendimiento de una señal de salida.

15 Con referencia a la FIG. 2, una forma de realización de un procedimiento de procesamiento de señales de voz/audio de la presente invención incluye:

S201: Cuando una señal de frecuencia ancha conmuta a una señal de frecuencia estrecha, predecir una señal predicha de alta frecuencia correspondiente a una señal de frecuencia estrecha de la trama actual.

20 Cuando una señal de frecuencia ancha conmuta a una señal de frecuencia estrecha, una trama previa es la señal de frecuencia ancha, y una trama actual es la señal de frecuencia estrecha. La etapa de predecir una señal predicha de alta frecuencia correspondiente a la señal de frecuencia estrecha de la trama actual incluye: predecir una señal de excitación de la señal de alta frecuencia de la trama actual de señal de voz/audio según la señal de frecuencia estrecha de la trama actual; predecir un coeficiente LPC (codificación predictiva lineal) de la señal de alta frecuencia
25 de la trama actual de señal de voz/audio; y sintetizar la señal predicha de excitación de alta frecuencia y el coeficiente LPC para obtener la señal predicha de alta frecuencia *syn_tmp*.

30 En una forma de realización, parámetros tales como un periodo de tono, un libro de códigos algebraico y una ganancia pueden extraerse a partir de la señal de frecuencia estrecha, y la señal de excitación de alta frecuencia se predice mediante un nuevo muestreo y mediante filtrado.

35 En otra forma de realización, operaciones tales como un muestreo ascendente, paso bajo y obtener un valor absoluto o un valor elevado al cuadrado, pueden llevarse a cabo en la señal de dominio de tiempo de frecuencia estrecha o una señal de excitación de dominio de tiempo de frecuencia estrecha con el fin de predecir la señal de excitación de alta frecuencia.

40 Para predecir el coeficiente LPC de la señal de alta frecuencia, un coeficiente LPC de alta frecuencia de una trama histórica o una serie de valores prefijados puede usarse como el coeficiente LPC de la trama actual; o diferentes maneras de predicción pueden usarse para diferentes tipos de señal.

S202: Obtener un parámetro de envolvente de dominio de tiempo y un parámetro de ganancia global de dominio de tiempo correspondientes a la señal predicha de alta frecuencia.

45 Una serie de valores predeterminados puede usarse como el parámetro de envolvente de dominio de tiempo de alta frecuencia de la trama actual. Las señales de banda estrecha pueden clasificarse generalmente en varios tipos, una serie de valores puede prefijarse para cada tipo, y un grupo de parámetros prefijados de envolvente de dominio de tiempo puede seleccionarse según los tipos de trama actual de señales de banda estrecha; o puede fijarse un grupo de valores de envolvente de dominio de tiempo; por ejemplo, cuando el número de envolventes de dominio de tiempo es M, los valores prefijados pueden ser M 0,3536s. En esta forma de realización, la obtención de un
50 parámetro de envolvente de dominio de tiempo es una etapa opcional, no una etapa necesaria.

55 El parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia se obtiene según un parámetro de inclinación de espectro de la señal de frecuencia estrecha y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica, que incluye las siguientes etapas en una forma de realización:

60 S2021: Clasificar la trama actual de señal de voz/audio como un primer tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama actual de señal de voz/audio y la correlación entre la señal de frecuencia estrecha de la trama actual y la señal de frecuencia estrecha de la trama histórica, donde en una forma de realización, el primer tipo de señal es una señal fricativa, y el segundo tipo de señal es una señal no fricativa; y cuando el parámetro de inclinación de espectro *inclinación* es mayor que 5 y un parámetro de correlación *cor* es inferior a un valor dado, clasificar la señal de frecuencia estrecha como fricativa y el resto como no fricativas.

65 El parámetro *cor* que muestra la correlación entre la señal de frecuencia estrecha de la trama actual y la de señal de frecuencia estrecha de la trama histórica puede determinarse según una relación de magnitud de energía entre

señales de una misma banda de frecuencia, o puede determinarse según una relación de energía entre varias bandas de la misma frecuencia, o puede calcularse según una fórmula que muestre una autocorrelación o una correlación cruzada entre señales de dominio de tiempo o que muestre una autocorrelación o una correlación cruzada entre señales de excitación de dominio de tiempo.

5 S2022: Cuando la trama actual de señal de voz/audio es un primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia. Es decir, cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es inferior o igual al primer valor predeterminado, un valor original del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es superior al primer valor predeterminado, el primer valor predeterminado se usa como el valor de límite de parámetro de inclinación de espectro.

15 El parámetro de ganancia global de dominio de tiempo *ganancia'* se obtiene según la siguiente fórmula:

$$ganancia' = \begin{cases} inclinación, & inclinación \leq \delta 1 \\ \delta 1, & inclinación > \delta 1 \end{cases}$$

20 donde *inclinación* es el parámetro de inclinación de espectro, y $\delta 1$ es el primer valor predeterminado.

S2023: Cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia. Es decir, cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio pertenece al primer intervalo, un valor original del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es mayor que un límite superior del primer intervalo, el límite superior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de la señal de voz/audio es inferior a un límite inferior del primer intervalo, el límite inferior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro.

35 El parámetro de ganancia global de dominio de tiempo *ganancia'* se obtiene según la siguiente fórmula:

$$ganancia' = \begin{cases} inclinación, & inclinación \in [a, b] \\ a, & inclinación < a \\ b, & inclinación > b \end{cases}$$

donde *inclinación* es el parámetro de inclinación de espectro, y $[a, b]$ es el primer intervalo.

40 En una forma de realización se obtiene un parámetro de inclinación de espectro *inclinación* de una señal de frecuencia estrecha y un parámetro *cor* que muestra una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica; la trama actual de señales se clasifica en dos tipos, fricativa y no fricativa, según los parámetros *inclinación* y *cor*; cuando el parámetro de inclinación de espectro *inclinación* es mayor que 5 y el parámetro de correlación *cor* es inferior a un valor dado, la señal de frecuencia estrecha se clasifica como fricativa, donde el resto son no fricativas; el parámetro *inclinación* está limitado dentro del intervalo de valores $0,5 \leq inclinación \leq 1,0$ y se usa como un parámetro de ganancia global de dominio de tiempo de una señal no fricativa, y el parámetro *inclinación* está limitado al intervalo de valores $inclinación \leq 8,0$ y se usa como un parámetro de ganancia global de dominio de tiempo de una señal fricativa. En lo que respecta a una señal fricativa, un parámetro de inclinación de espectro puede ser cualquier valor superior a 5, y en lo que respecta a una señal no fricativa, un parámetro de inclinación de espectro puede ser cualquier valor inferior o igual a 5, o puede ser superior a 5. Para garantizar que un parámetro de inclinación de espectro *inclinación* pueda usarse como un parámetro estimado de ganancia global de dominio de tiempo, el parámetro *inclinación* se limita dentro de un intervalo de valores y después se usa como un parámetro de ganancia global de dominio de tiempo. Es decir, cuando el parámetro *inclinación* es superior a 8, se determina que $inclinación=8$ se usa como un parámetro de ganancia global de dominio de tiempo de una señal fricativa; cuando $inclinación < 0,5$ se determina que $inclinación=0,5$; o cuando $inclinación > 1,0$ se determina que $inclinación=1,0$, y 0,5 o 1,0 se usa como un parámetro de ganancia global de dominio de tiempo de una señal no fricativa.

S203: Llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica de señal de voz/audio y la energía de la señal inicial de alta frecuencia de la trama actual de señal de voz/audio.

Se calcula la relación de energía $Relación = E_{syn}(-1) / E_{syn_tmp}$, y el valor ponderado de *inclinación* y *Relación* se usa como el parámetro predicho de ganancia global *ganancia* de la trama actual, es decir, $ganancia = \alpha * Relación + \beta * ganancia'$, donde *ganancia'* es el parámetro de ganancia global de dominio de tiempo, $\alpha + \beta = 1$, los valores α y β son diferentes según diferentes tipos de señal, $E_{syn}(-1)$ representa la energía de la señal final de salida de dominio de tiempo de alta frecuencia *syn* de la trama histórica, y E_{syn_tmp} representa la energía de la señal predicha de dominio de tiempo de alta frecuencia *syn* de la trama actual.

S204: Corregir la señal predicha de alta frecuencia usando el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia.

La señal predicha de alta frecuencia se multiplica por el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global de dominio de tiempo para obtener la señal de dominio de tiempo de alta frecuencia.

En esta forma de realización, el parámetro de envolvente de dominio de tiempo es opcional. Solamente cuando se incluye el parámetro de ganancia global de dominio de tiempo, la señal predicha de alta frecuencia puede corregirse usando el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia. Es decir, la señal predicha de alta frecuencia se multiplica por el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia.

S205: Sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

La energía E_{syn} de la señal de dominio de tiempo de alta frecuencia *syn* se usa para predecir un parámetro de ganancia global de dominio de tiempo de una trama siguiente. Es decir, el valor de E_{syn} se asigna a $E_{syn}(-1)$.

En la forma de realización anterior se corrige una banda de alta frecuencia de una señal de frecuencia estrecha que sigue una señal de frecuencia ancha con el fin de implementar una transición gradual de la parte de alta frecuencia entre una banda de frecuencia ancha y una banda de frecuencia estrecha, por lo que se elimina de manera eficaz la molestia auditiva producida por la conmutación entre la banda de frecuencia ancha y la banda de frecuencia estrecha; además, puesto que el procesamiento correspondiente se lleva a cabo en la trama durante la conmutación, el problema que se produce durante la actualización de los parámetros y del estado se elimina indirectamente. Mantener en un mismo dominio de señal un algoritmo de conmutación de ancho de banda y un algoritmo de codificación/descodificación de la señal de alta frecuencia antes de la conmutación no solo garantiza que no se añada ningún retardo adicional y que el algoritmo sea sencillo, sino que también garantiza el rendimiento de una señal de salida.

Con referencia a la FIG. 3, otra forma de realización de un procedimiento de procesamiento de señales de voz/audio incluye:

S301: Cuando una señal de frecuencia estrecha conmuta a una señal de frecuencia ancha, obtener una trama actual de señal de alta frecuencia.

Cuando una señal de frecuencia estrecha conmuta a una señal de frecuencia ancha, una trama previa es una señal de frecuencia estrecha, y una trama actual es una señal de frecuencia ancha.

S302: Obtener un parámetro de envolvente de dominio de tiempo y un parámetro de ganancia global de dominio de tiempo correspondientes a la señal predicha de alta frecuencia.

El parámetro de envolvente de dominio de tiempo y el parámetro de ganancia global de dominio de tiempo pueden obtenerse directamente a partir de la trama actual de señal de alta frecuencia. Obtener un parámetro de envolvente de dominio de tiempo es una etapa opcional.

S303: Llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica de señal de voz/audio y la energía de una señal inicial de alta frecuencia de una trama actual de señal de voz/audio.

Puesto que la trama actual es una señal de frecuencia ancha, todos los parámetros de la señal de alta frecuencia pueden obtenerse mediante descodificación. Para garantizar una transición gradual durante la conmutación, el parámetro de ganancia global de dominio de tiempo se suaviza de la siguiente manera:

5 Se calcula la relación de energía $Relación = E_{syn(-1)} / E_{syn_tmp}$, donde $E_{syn(-1)}$ representa la energía de una señal final de salida de dominio de tiempo de alta frecuencia $sync$ de una trama histórica, y E_{syn_tmp} representa la energía de una señal de dominio de tiempo de alta frecuencia $sync$ de la trama actual.

10 El valor ponderado del parámetro de ganancia global de dominio de tiempo $ganancia$ y del parámetro $Relación$ que se obtienen mediante descodificación se usa como el parámetro predicho de ganancia global $ganancia$ de la trama actual, es decir, $ganancia = \alpha * Relación + \beta * ganancia'$, donde $ganancia'$ es el parámetro de ganancia global de dominio de tiempo, $\alpha + \beta = 1$, y los valores de α y β son diferentes según diferentes tipos de señal.

15 Cuando señales de banca estrecha de la trama de audio actual y una trama anterior de señal de voz/audio tienen una correlación predeterminada, un valor obtenido atenuando, según un valor diferencial específico, un valor α de factor de ponderación de la relación de energía correspondiente a la trama anterior de señal de voz/audio se usa como un factor de ponderación de la relación de energía correspondiente a la trama de audio actual, donde la atenuación se realiza trama a trama hasta que el valor α sea 0.

20 Cuando las señales de frecuencia estrecha de tramas consecutivas son de un mismo tipo de señal, o una correlación entre señales de frecuencia estrecha de tramas consecutivas satisface una condición específica, es decir, las tramas consecutivas tienen una correlación específica o los tipos de señal de las tramas consecutivas son similares, el valor α se reduce trama a trama según un valor diferencial específico hasta que el valor α valga 0; cuando las señales de frecuencia estrecha de las tramas consecutivas no están correlacionadas, el valor α se reduce directamente a 0, es decir, un resultado de descodificación actual se mantiene sin llevar a cabo una ponderación o una corrección.

30 S304: Corregir la señal de alta frecuencia usando el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia.

La corrección se refiere a que la señal de alta frecuencia se multiplica por el parámetro de envolvente de dominio de tiempo y por el parámetro predicho de ganancia global de dominio de tiempo para obtener la señal corregida de dominio de tiempo de alta frecuencia.

35 En esta forma de realización, el parámetro de envolvente de dominio de tiempo es opcional. Solamente cuando se incluye el parámetro de ganancia global de dominio de tiempo, la señal de alta frecuencia puede corregirse usando el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia. Es decir, la señal de alta frecuencia se multiplica por el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia.

40 S305: Sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

45 En la forma de realización anterior se corrige una banda de alta frecuencia de una señal de frecuencia ancha que sigue a una señal de frecuencia estrecha con el fin de implementar una transición gradual de la parte de alta frecuencia entre una banda de frecuencia ancha y una banda de frecuencias estrecha, por lo que se elimina de manera eficaz la molestia auditiva producida por la conmutación entre la banda de frecuencia ancha y la banda de frecuencia estrecha; además, puesto que el procesamiento correspondiente se lleva a cabo en la trama durante la conmutación, el problema que se produce durante la actualización de los parámetros y del estado se elimina indirectamente. Mantener en un mismo dominio de señal un algoritmo de conmutación de ancho de banda y un algoritmo de codificación/descodificación de la señal de alta frecuencia antes de la conmutación no solo garantiza que no se añada ningún retardo adicional y que el algoritmo sea sencillo, sino que también garantiza el rendimiento de una señal de salida.

55 Con referencia a la FIG. 4, otra forma de realización de un procedimiento de procesamiento de señales de voz/audio incluye:

60 S401: Cuando una señal de voz/audio conmuta desde una señal de frecuencia ancha a una señal de frecuencia estrecha, obtener una señal inicial de alta frecuencia correspondiente a una trama actual de la señal de voz/audio.

65 Cuando una señal de frecuencia ancha conmuta a una señal de frecuencia estrecha, una trama previa es la señal de frecuencia ancha, y una trama actual es la señal de frecuencia estrecha. La etapa de predecir una señal inicial de alta frecuencia correspondiente a una señal de frecuencia estrecha de la trama actual incluye: predecir una señal de excitación de la señal de alta frecuencia de la trama actual de señal de voz/audio según la señal de frecuencia estrecha de la trama actual; predecir un coeficiente LPC de la señal de alta frecuencia de la trama actual de señal de

voz/audio; y sintetizar la señal predicha de excitación de alta frecuencia y el coeficiente LPC para obtener la señal predicha de alta frecuencia *syn_tmp*.

5 En una forma de realización, parámetros tales como un periodo de tono, un libro de códigos algebraico y una ganancia pueden extraerse a partir de la señal de frecuencia estrecha, y la señal de excitación de alta frecuencia se predice mediante un nuevo muestreo y mediante filtrado.

10 En otra forma de realización, operaciones tales como un muestreo ascendente, paso bajo y obtener un valor absoluto o un valor elevado al cuadrado, pueden llevarse a cabo en la señal de dominio de tiempo de frecuencia estrecha o una señal de excitación de dominio de tiempo de frecuencia estrecha con el fin de predecir la señal de excitación de alta frecuencia.

15 Para predecir el coeficiente LPC de la señal de alta frecuencia, un coeficiente LPC de alta frecuencia de una trama histórica o una serie de valores prefijados puede usarse como el coeficiente LPC de la trama actual; o diferentes maneras de predicción pueden usarse para diferentes tipos de señal.

20 S402: Obtener un parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia según un parámetro de inclinación de espectro de la trama actual de señal de voz/audio y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica.

En una forma de realización se incluyen las siguientes etapas:

25 S2021: Clasificar la trama actual de señal de voz/audio como un primer tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama actual de señal de voz/audio y la correlación entre la señal de frecuencia estrecha de la trama actual y la señal de frecuencia estrecha de la trama histórica, donde en una forma de realización, el primer tipo de señal es una señal fricativa, y el segundo tipo de señal es una señal no fricativa.

30 En una forma de realización, cuando el parámetro de inclinación de espectro *inclinación* es superior a 5 y un parámetro de correlación *cor* es inferior a un valor dado, la señal de frecuencia estrecha se clasifica como fricativa y el resto como no fricativas. El parámetro *cor* que muestra la correlación entre la señal de frecuencia estrecha de la trama actual y la señal de frecuencia estrecha de la trama histórica puede determinarse según una relación de magnitud de energía entre señales de una misma banda de frecuencia, o puede determinarse según una relación de energía entre varias bandas de la misma frecuencia, o puede calcularse según una fórmula que muestre una autocorrelación o una correlación cruzada entre señales de dominio de tiempo o que muestre una autocorrelación o una correlación cruzada entre señales de excitación de dominio de tiempo.

40 S2022: Cuando la trama actual de señal de voz/audio es un primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia. Es decir, cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es inferior o igual al primer valor predeterminado, un valor original del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es superior al primer valor predeterminado, el primer valor predeterminado se usa como el valor de límite de parámetro de inclinación de espectro.

50 Cuando la trama actual de señal de voz/audio es una señal fricativa, el parámetro de ganancia global de dominio de tiempo *ganancia'* se obtiene según la siguiente fórmula:

$$ganancia' = \begin{cases} inclinación, & inclinación \leq \partial 1 \\ \partial 1, & inclinación > \partial 1 \end{cases}$$

donde *inclinación* es el parámetro de inclinación de espectro, y $\partial 1$ es el primer valor predeterminado.

55 S2023: Cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia. Es decir, cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio pertenece al primer intervalo, un valor original del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de señal de voz/audio es mayor que un límite superior del primer intervalo, el límite superior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro; cuando el parámetro de inclinación de espectro de la trama actual de la señal de voz/audio es inferior a un límite inferior del

primer intervalo, el límite inferior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro.

5 Cuando la trama actual de señal de voz/audio no es una señal fricativa, el parámetro de ganancia global de dominio de tiempo *ganancia'* se obtiene según la siguiente fórmula:

$$ganancia' = \begin{cases} \text{inclinación, inclinación} \in [a, b] \\ a, \text{ inclinación} < a \\ b, \text{ inclinación} > b \end{cases},$$

10 donde *inclinación* es el parámetro de inclinación de espectro y $[a, b]$ es el primer intervalo.

En una forma de realización se obtiene un parámetro de inclinación de espectro *inclinación* de una señal de frecuencia estrecha y un parámetro *cor* que muestra una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica; la trama actual de señales se clasifica en dos tipos, fricativa y no fricativa, según los parámetros *inclinación* y *cor*; cuando el parámetro de inclinación de espectro *inclinación* es mayor que 5 y el parámetro de correlación *cor* es inferior a un valor dado, la señal de frecuencia estrecha se clasifica como fricativa, donde el resto son no fricativas; el parámetro *inclinación* está limitado dentro del intervalo de valores $0,5 \leq \text{inclinación} \leq 1,0$ y se usa como un parámetro de ganancia global de dominio de tiempo de una señal no fricativa, y el parámetro *inclinación* está limitado al intervalo de valores $\text{inclinación} \leq 8,0$ y se usa como un parámetro de ganancia global de dominio de tiempo de una señal fricativa. En lo que respecta a una señal fricativa, un parámetro de inclinación de espectro puede ser cualquier valor superior a 5, y en lo que respecta a una señal no fricativa, un parámetro de inclinación de espectro puede ser cualquier valor inferior o igual a 5, o puede ser superior a 5. Para garantizar que un parámetro de inclinación de espectro *inclinación* pueda usarse como un parámetro predicho de ganancia global, el parámetro *inclinación* se limita dentro de un intervalo de valores y después se usa como un parámetro de ganancia global de dominio de tiempo. Es decir, cuando *inclinación* > 8, se determina que *inclinación*=8 y 8 se usa como un parámetro de ganancia global de dominio de tiempo de una señal fricativa; cuando *inclinación*<0,5 se determina que *inclinación*=0,5; o cuando *inclinación*>1,0, se determina que *inclinación*=1,0, y 0,5 o 1,0 se usa como un parámetro de ganancia global de dominio de tiempo de una señal no fricativa.

30 S403: Corregir la señal inicial de alta frecuencia usando el parámetro de ganancia global de dominio de tiempo para obtener una señal corregida de dominio de tiempo de alta frecuencia.

En una forma de realización, la señal inicial de alta frecuencia se multiplica por el parámetro de ganancia global de dominio de tiempo para obtener la señal corregida de dominio de tiempo de alta frecuencia.

35 En otra forma de realización, la etapa S403 puede incluir:

llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica y la energía de la señal inicial de alta frecuencia de la trama actual; y corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia; es decir, la señal inicial de alta frecuencia se multiplica por el parámetro predicho de ganancia global con el fin de obtener una señal corregida de dominio de tiempo de alta frecuencia.

Opcionalmente, antes de la etapa S403, el procedimiento puede incluir además:

50 obtener un parámetro de envolvente de dominio de tiempo correspondiente a la señal inicial de alta frecuencia, y la corrección de la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global incluye: corregir la señal inicial de alta frecuencia usando el parámetro de envolvente de dominio de tiempo y el parámetro de ganancia global de dominio de tiempo.

55 S404: Sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

En la forma de realización anterior, cuando una banda de frecuencia ancha conmuta a una banda de frecuencia estrecha, un parámetro de ganancia global de dominio de tiempo de una señal de alta frecuencia se obtiene según un parámetro de inclinación de espectro y una correlación de intertrama. Usando el parámetro de inclinación de espectro de frecuencia estrecha, una relación de energía entre una señal de frecuencia estrecha y una señal de alta frecuencia puede estimarse correctamente con el fin de estimar mejor la energía de la señal de alta frecuencia.

Usando la correlación intertrama, una correlación intertrama entre señales de alta frecuencia puede estimarse usando de manera apropiada la correlación entre tramas de frecuencia estrecha. De esta manera, cuando la ponderación se realiza para obtener una ganancia global de alta frecuencia, puede usarse la anterior información real, y no se introduce ningún ruido no deseable. La señal de alta frecuencia se corrige usando el parámetro de ganancia global de dominio de tiempo con el fin de implementar una transición gradual de la parte de alta frecuencia entre la banda de frecuencia ancha y al banda de frecuencia estrecha, por lo que se elimina de manera eficaz la molestia auditiva producida por la conmutación entre la banda de frecuencia ancha y la banda de frecuencia estrecha.

En asociación con las anteriores formas de realización de procedimiento, la presente invención proporciona además un aparato de procesamiento de señales de voz/audio. El aparato puede estar ubicado en un dispositivo terminal, un dispositivo de red o un dispositivo de prueba. El aparato de procesamiento de señales de voz/audio puede implementarse mediante un circuito de hardware o puede implementarse mediante software en combinación con hardware. Por ejemplo, con referencia a la FIG. 5, un procesador invoca al aparato de procesamiento de señales de voz/audio para implementar un procesamiento de señales de voz/audio. El aparato de procesamiento de señales de voz/audio puede ejecutar los procedimientos y procesos de las anteriores formas de realización de procedimiento.

Con referencia a la FIG. 6, una forma de realización de un aparato de procesamiento de señales de voz/audio incluye:

- una unidad de adquisición 601, configurada para: cuando una señal de voz/audio conmuta el ancho de banda, obtener una señal inicial de alta frecuencia correspondiente a una trama actual de la señal de voz/audio;
- una unidad de obtención de parámetros 602, configurada para obtener un parámetro de ganancia global de dominio de tiempo correspondiente a la señal inicial de alta frecuencia;
- una unidad de procesamiento de ponderación 603, configurada para llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica y la energía de la señal inicial de alta frecuencia de la trama actual;
- una unidad de corrección 604, configurada para corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia; y
- una unidad de sintetización 605, configurada para sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

En una forma de realización, la conmutación de ancho de banda se produce desde una señal de frecuencia ancha a una señal de frecuencia estrecha, y la unidad de obtención de parámetros 602 incluye:

- una unidad de obtención de parámetros de ganancia global, configurada para obtener el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia según un parámetro de inclinación de espectro de la trama actual de señal de voz/audio y una correlación entre una trama actual de señal de voz/audio y una señal de frecuencia estrecha de la trama histórica.

Con referencia a la FIG. 7, en otra forma de realización, la conmutación de ancho de banda se produce desde una señal de frecuencia ancha a una señal de frecuencia estrecha, y la unidad de obtención de parámetros 602 incluye:

- una unidad de obtención de envolvente de dominio de tiempo 701, configurada para usar una serie de valores prefijados como un parámetro de envolvente de dominio de tiempo de alta frecuencia de la trama actual de señal de voz/audio; y
- una unidad de obtención de parámetros de ganancia global 702, configurada para obtener el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia según un parámetro de inclinación de espectro de la trama actual de señal de voz/audio y una correlación entre una trama actual de señal de voz/audio y una señal de frecuencia estrecha de la trama histórica.

Por lo tanto, la unidad de corrección 604 está configurada para corregir la señal inicial de alta frecuencia usando el parámetro de envolvente de dominio de tiempo y el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia.

Con referencia a la FIG. 8, además, una forma de realización de la unidad de obtención de parámetros de ganancia global 702 incluye:

- una unidad de clasificación 801, configurada para clasificar la trama actual de señal de voz/audio como un primer tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama

actual de señal de voz/audio y la correlación entre la trama actual de señal de voz/audio y la señal de frecuencia estrecha de la trama histórica;

5 una primera unidad de limitación 802, configurada para: cuando la trama actual de señal de voz/audio es un primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia; y

10 una segunda unidad de limitación 803, configurada para: cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia.

15 Además, en una forma de realización, el primer tipo de señal es una señal fricativa y el segundo tipo de señal es una señal no fricativa; cuando el parámetro de inclinación de espectro *inclinación* es mayor que 5 y un parámetro de correlación *cor* es inferior a un valor dado, la señal de frecuencia estrecha se clasifica como una señal fricativa y el resto son señales no fricativas; el primer valor predeterminado es 8; y el primer intervalo predeterminado es [0,5; 1].

20 Con referencia a la FIG. 9, en una forma de realización, la unidad de adquisición 601 incluye:

una unidad de obtención de señales de excitación 901, configurada para predecir una señal de excitación de la señal de alta frecuencia según la trama actual de señal de voz/audio;

una unidad de obtención de coeficientes LPC 902, configurada para predecir un coeficiente LPC de la señal de alta frecuencia; y

25 una unidad de generación 903, configurada para sintetizar la señal de excitación de la señal de alta frecuencia y el coeficiente LPC de la señal de alta frecuencia para obtener la señal predicha de alta frecuencia.

30 En una forma de realización, la conmutación de ancho de banda se produce desde una señal de frecuencia estrecha a una señal de frecuencia ancha, y el aparato de procesamiento de señales de voz/audio incluye además:

una unidad de ajuste de factor de ponderación, configurada para: cuando señales de banca estrecha de la trama de audio actual de señal de voz/audio y una trama anterior de señal de voz/audio tienen una correlación predeterminada, usar un valor obtenido atenuando, según un valor diferencial específico, un valor alfa de factor de ponderación de la relación de energía correspondiente a la trama anterior de señal de voz/audio como un factor de ponderación de la relación de energía correspondiente a la trama de audio actual, donde la atenuación se realiza trama a trama hasta que el valor alfa sea 0.

40 Con referencia a la FIG. 10, otra forma de realización de un aparato de procesamiento de señales de voz/audio incluye:

una unidad de predicción 1001, configurada para: cuando una señal de voz/audio conmuta desde una señal de frecuencia ancha a una señal de frecuencia estrecha, obtener una señal inicial de alta frecuencia correspondiente a una trama actual de señal de voz/audio;

45 una unidad de obtención de parámetros 1002, configurada para obtener un parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia según un parámetro de inclinación de espectro de la trama actual de señal de voz/audio y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica;

50 una unidad de corrección 1003, configurada para corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia; y

una unidad de sintetización 1004, configurada para sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

55 Con referencia a la FIG. 8, la unidad de obtención de parámetros 1002 incluye:

una unidad de clasificación 801, configurada para clasificar la trama actual de señal de voz/audio como un primer tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama actual de señal de voz/audio y la correlación entre la trama actual de señal de voz/audio y la señal de frecuencia estrecha de la trama histórica;

60 una primera unidad de limitación 802, configurada para: cuando la trama actual de señal de voz/audio es un primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia; y

65 una segunda unidad de limitación 803, configurada para: cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia.

5 una segunda unidad de limitación 803, configurada para: cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta frecuencia.

10 Además, en una forma de realización, el primer tipo de señal es una señal fricativa y el segundo tipo de señal es una señal no fricativa; cuando el parámetro de inclinación de espectro *inclinación* es mayor que 5 y un parámetro de correlación *cor* es inferior a un valor dado, la señal de frecuencia estrecha se clasifica como una señal fricativa, y el resto son señales no fricativas; el primer valor predeterminado es 8; y el primer intervalo predeterminado es [0,5; 1].

Opcionalmente, en una forma de realización, el aparato de procesamiento de señales de voz/audio incluye además:

15 una unidad de procesamiento de ponderación, configurada para llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica y la energía de la señal inicial de alta frecuencia de la trama actual; y

20 la unidad de corrección está configurada para corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener la señal corregida de dominio de tiempo de alta frecuencia.

25 En otra forma de realización, la unidad de obtención de parámetros está configurada además para obtener un parámetro de envolvente de dominio de tiempo correspondiente a la señal inicial de alta frecuencia; y la unidad de corrección está configurada para corregir la señal inicial de alta frecuencia usando el parámetro de envolvente de dominio de tiempo y el parámetro de ganancia global de dominio de tiempo.

30 Un experto en la técnica puede entender que todos o parte de los procesos de los procedimientos de las formas de realización pueden implementarse mediante un programa informático que da instrucciones a hardware pertinente. El programa puede almacenarse en un medio de almacenamiento legible por ordenador. Cuando el programa se ejecuta se llevan a cabo los procesos de los procedimientos de las formas de realización. El medio de almacenamiento puede incluir: un disco magnético, un disco óptico, una memoria de solo lectura (ROM) o una memoria de acceso aleatorio (RAM).

35 Lo que antecede es simplemente formas de realización a modo de ejemplo para ilustrar la presente invención, pero el alcance de la presente invención no se limita a esto. Modificaciones o variaciones resultarán evidentes a los expertos en la técnica anterior. El alcance de la presente invención está definido en las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un procedimiento de procesamiento de señales de voz/audio, que comprende:

5 cuando una señal de voz/audio conmuta desde una señal de frecuencia ancha a una señal de frecuencia estrecha, obtener una señal inicial de alta frecuencia correspondiente a una trama actual de señal de voz/audio;
 obtener un parámetro de ganancia global de dominio de tiempo de la señal inicial de alta frecuencia;
 llevar a cabo un procesamiento de ponderación en una relación de energía y el parámetro de ganancia global
 10 de dominio de tiempo, y usar un valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica y la energía de la señal inicial de alta frecuencia de la trama actual;
 corregir la señal inicial de alta frecuencia usando el parámetro predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia; y
 15 sintetizar una señal de dominio de tiempo de frecuencia estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la señal sintetizada.

2. El procedimiento según la reivindicación 1, en el que el obtener el parámetro de ganancia global de dominio de tiempo de la señal inicial de alta frecuencia comprende:

20 obtener un parámetro de ganancia global de dominio de tiempo de la señal inicial de alta frecuencia según un parámetro de inclinación de espectro de la trama actual de señal de voz/audio y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica.

3. El procedimiento según la reivindicación 2, en el que el obtener el parámetro de ganancia global de dominio de tiempo de la señal inicial de alta frecuencia según un parámetro de inclinación de espectro de una trama actual de señal de voz/audio y una correlación entre una señal de frecuencia estrecha de la trama actual y una señal de frecuencia estrecha de la trama histórica comprende:

30 clasificar la trama actual de señal de voz/audio como un primer tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama actual de señal de voz/audio y la correlación entre la señal de frecuencia estrecha de la trama actual y la señal de frecuencia estrecha de la trama histórica, en el que el primer tipo de señal es una señal fricativa y el segundo tipo de señal es una señal no fricativa;
 cuando la trama actual de señal de voz/audio es un primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro;
 cuando la trama actual de señal de voz/audio es un segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro; y
 40 usar el valor de límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal inicial de alta frecuencia.

4. El procedimiento según la reivindicación 3, en el que limitar el parámetro de inclinación de espectro a menos de o igual a un primer valor predeterminado para obtener un valor de límite de parámetro de inclinación de espectro comprende:

45 cuando un valor del parámetro de inclinación de espectro es menor que o igual al primer valor predeterminado, el valor del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro;
 cuando un valor del parámetro de inclinación de espectro es mayor que el primer valor predeterminado, el primer valor predeterminado se usa como el valor de límite de parámetro de inclinación de espectro.
 50

5. El procedimiento según la reivindicación 3 o 4, en el que el primer valor predeterminado es 8.

6. El procedimiento según la reivindicación 3, en el que limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para obtener un valor de límite de parámetro de inclinación de espectro comprende:

55 cuando un valor del parámetro de inclinación de espectro pertenece al primer intervalo, el valor del parámetro de inclinación de espectro se mantiene como el valor de límite de parámetro de inclinación de espectro;
 cuando un valor del parámetro de inclinación de espectro es mayor que un límite superior del primer intervalo, el límite superior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro;
 60 cuando un valor del parámetro de inclinación de espectro es menor que un límite inferior del primer intervalo, el límite inferior del primer intervalo se usa como el valor de límite de parámetro de inclinación de espectro.

7. El procedimiento según la reivindicación 3 o 6, en el que primer intervalo es [0,5; 1].

8. El procedimiento según una cualquiera de las reivindicaciones 1-7, en el que el obtener una señal inicial de alta frecuencia correspondiente a una trama actual de señal de voz/audio comprende:

predecir una señal de excitación de alta frecuencia según la trama actual de señal de voz/audio;
 predecir un coeficiente LPC de la señal de alta frecuencia; y
 sintetizar la señal de excitación de alta frecuencia y el coeficiente LPC de la señal de alta frecuencia para
 5 obtener la señal inicial de alta frecuencia.

9. Un aparato de procesamiento de señales de voz/audio, que comprende:

10 una unidad de adquisición, configurada para, cuando una señal de voz/audio conmuta desde una señal de
 frecuencia ancha a una señal de frecuencia estrecha, obtener una señal inicial de alta frecuencia
 correspondiente a una trama actual de señal de voz/audio;
 una unidad de obtención de parámetros, configurada para obtener un parámetro de ganancia global de
 dominio de tiempo correspondiente a la señal inicial de alta frecuencia;
 15 una unidad de procesamiento de ponderación, configurada para llevar a cabo un procesamiento de
 ponderación en una relación de energía y el parámetro de ganancia global de dominio de tiempo, y usar un
 valor ponderado obtenido como un parámetro predicho de ganancia global, donde la relación de energía es
 una relación entre la energía de una señal de dominio de tiempo de alta frecuencia de una trama histórica y la
 energía de la señal inicial de alta frecuencia de la trama actual;
 20 una unidad de corrección, configurada para corregir la señal inicial de alta frecuencia usando el parámetro
 predicho de ganancia global para obtener una señal corregida de dominio de tiempo de alta frecuencia; y
 una unidad de sintetización, configurada para sintetizar una señal de dominio de tiempo de frecuencia
 estrecha de la trama actual y la señal corregida de dominio de tiempo de alta frecuencia y proporcionar la
 señal sintetizada.

25 10. El aparato según la reivindicación 9, en el que la unidad de obtención de parámetros comprende:

una unidad de obtención de parámetros de ganancia global, configurada para obtener el parámetro de
 ganancia global de dominio de tiempo de la señal de alta frecuencia según un parámetro de inclinación de
 espectro de la trama actual de señal de voz/audio y una correlación entre una señal de frecuencia estrecha
 30 de la trama actual y una señal de frecuencia estrecha de la trama histórica.

11. El aparato según la reivindicación 10, en el que la unidad de obtención de parámetros de ganancia global
 comprende:

35 una unidad de clasificación, configurada para clasificar la trama actual de señal de voz/audio como un primer
 tipo de señal o un segundo tipo de señal según el parámetro de inclinación de espectro de la trama actual de
 señal de voz/audio y la correlación entre la señal de frecuencia estrecha de la trama actual y la señal de
 frecuencia estrecha de la trama histórica, en el que el primer tipo de señal es una señal fricativa y el segundo
 tipo de señal es una señal no fricativa;
 40 una primera unidad de limitación, configurada para, cuando la trama actual de señal de voz/audio es un
 primer tipo de señal, limitar el parámetro de inclinación de espectro a un valor inferior o igual a un primer valor
 predeterminado para obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de
 límite de parámetro de inclinación de espectro como el parámetro de ganancia global de dominio de tiempo
 45 de la señal de alta frecuencia; y
 una segunda unidad de limitación, configurada para, cuando la trama actual de señal de voz/audio es un
 segundo tipo de señal, limitar el parámetro de inclinación de espectro a un valor de un primer intervalo para
 obtener un valor de límite de parámetro de inclinación de espectro, y usar el valor de límite de parámetro de
 inclinación de espectro como el parámetro de ganancia global de dominio de tiempo de la señal de alta
 50 frecuencia.

12. El aparato según la reivindicación 11, donde el aparato está configurado además para:

cuando un valor del parámetro de inclinación de espectro es menor que o igual al primer valor
 predeterminado, mantener el valor del parámetro de inclinación de espectro como el valor de límite de
 parámetro de inclinación de espectro;

55 cuando un valor del parámetro de inclinación de espectro es mayor que el primer valor predeterminado, usar
 el primer valor predeterminado como el valor de límite de parámetro de inclinación de espectro.

13. El aparato según la reivindicación 11 o 12, en el que el primer valor predeterminado es 8.

60 14. El aparato según la reivindicación 11, donde el aparato está configurado además para:

cuando un valor del parámetro de inclinación de espectro pertenece al primer intervalo, mantener el valor del
 parámetro de inclinación de espectro como el valor de límite de parámetro de inclinación de espectro;

cuando un valor del parámetro de inclinación de espectro es mayor que un límite superior del primer intervalo,
 usar el límite superior del primer intervalo como el valor de límite de parámetro de inclinación de espectro;

65 cuando un valor del parámetro de inclinación de espectro es menor que un límite inferior del primer intervalo,
 usar el límite inferior del primer intervalo como el valor de límite de parámetro de inclinación de espectro.

15. El aparato según la reivindicación 11 o 14, en el que el primer intervalo es [0,5; 1].
- 5 16. El aparato según una cualquiera de las reivindicaciones 9-15, en el que la unidad de adquisición comprende:
una unidad de obtención de señales de excitación, configurada para predecir una señal de excitación de la
señal de alta frecuencia según la trama actual de señal de voz/audio;
una unidad de obtención de coeficientes LPC, configurada para predecir un coeficiente LPC de la señal de
alta frecuencia; y
10 una señal de generación, configurada para sintetizar la señal de excitación de la señal de alta frecuencia y el
coeficiente LPC de la señal de alta frecuencia para obtener la señal inicial de alta frecuencia.

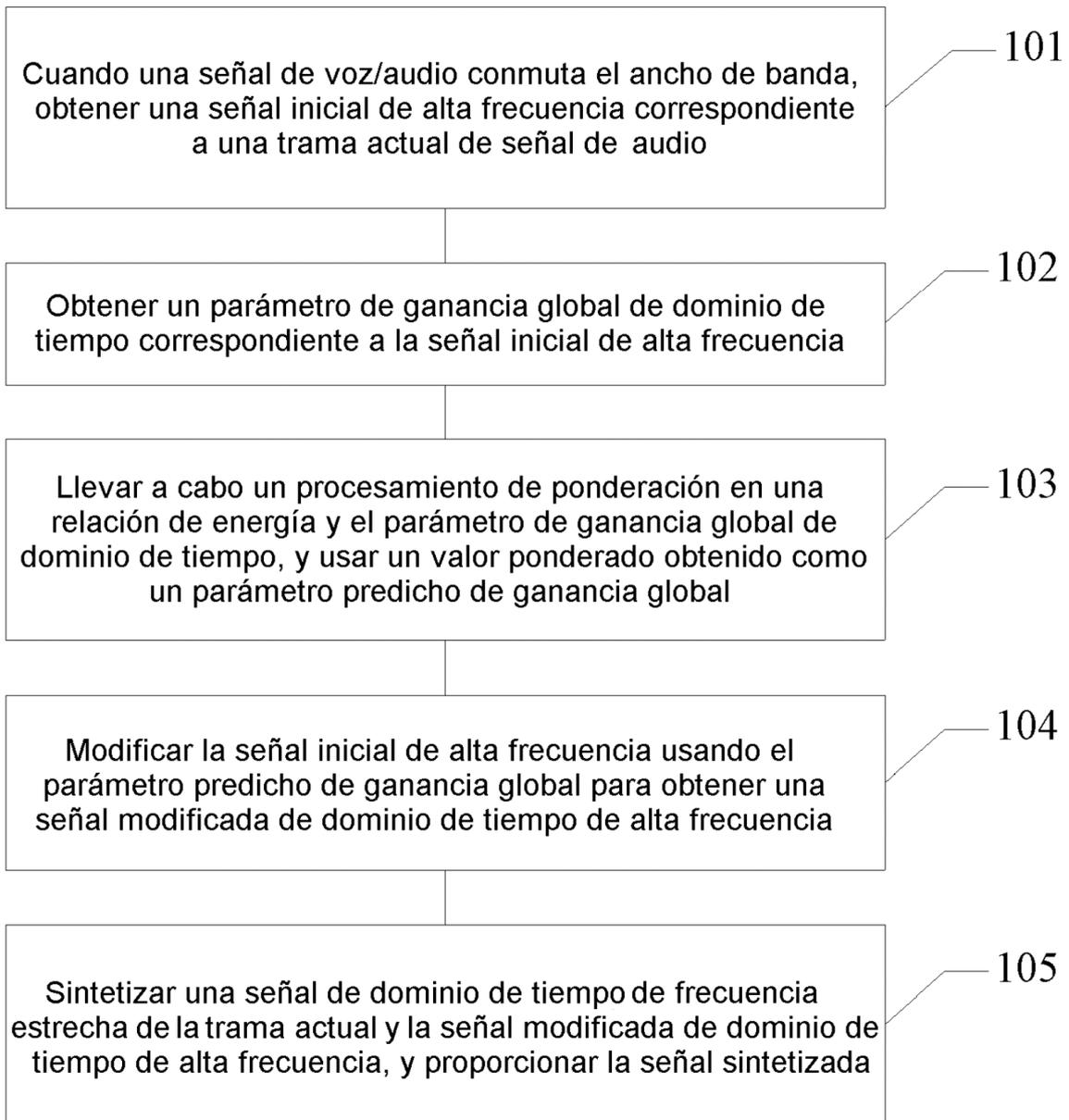


FIG. 1

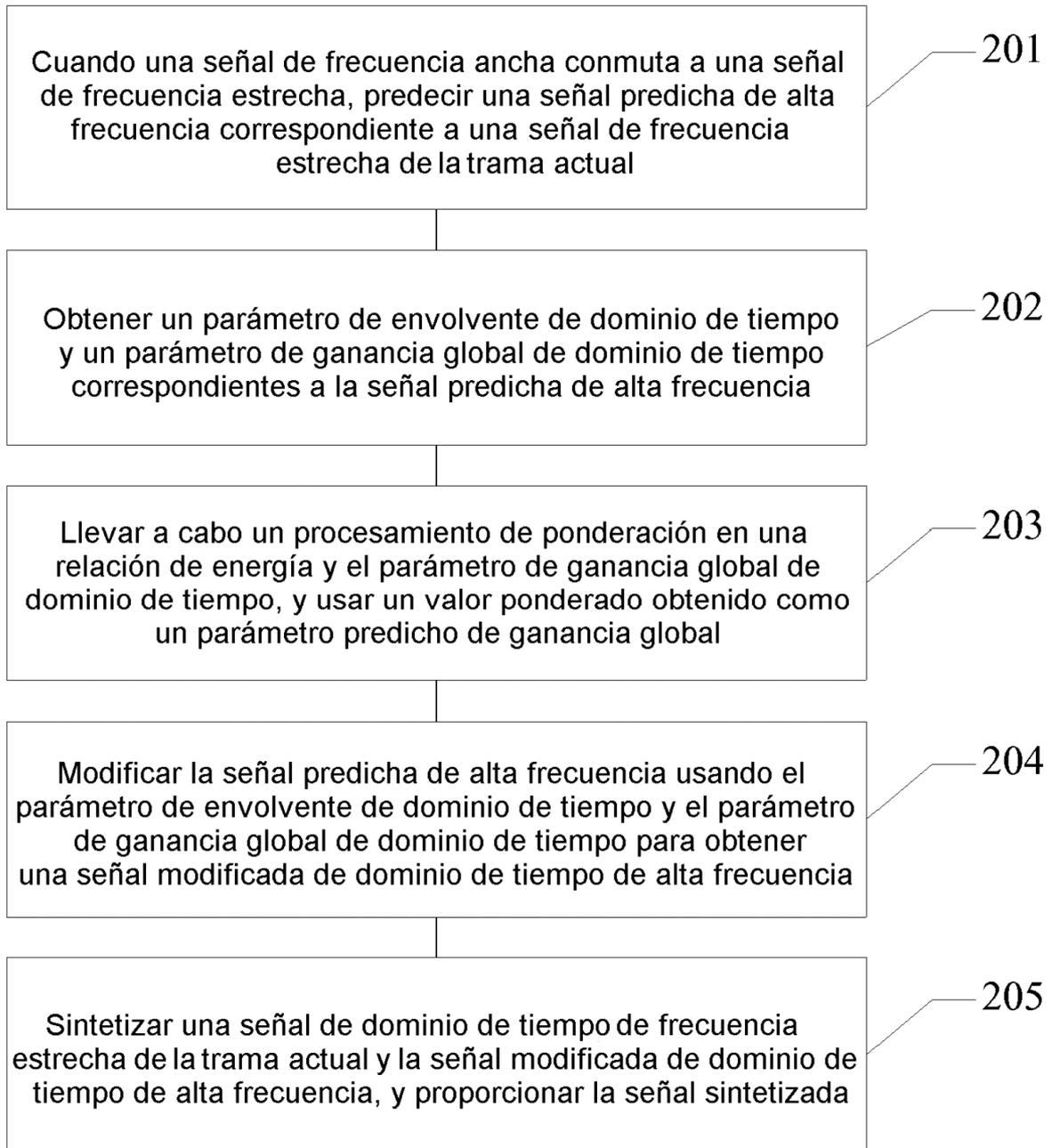


FIG. 2

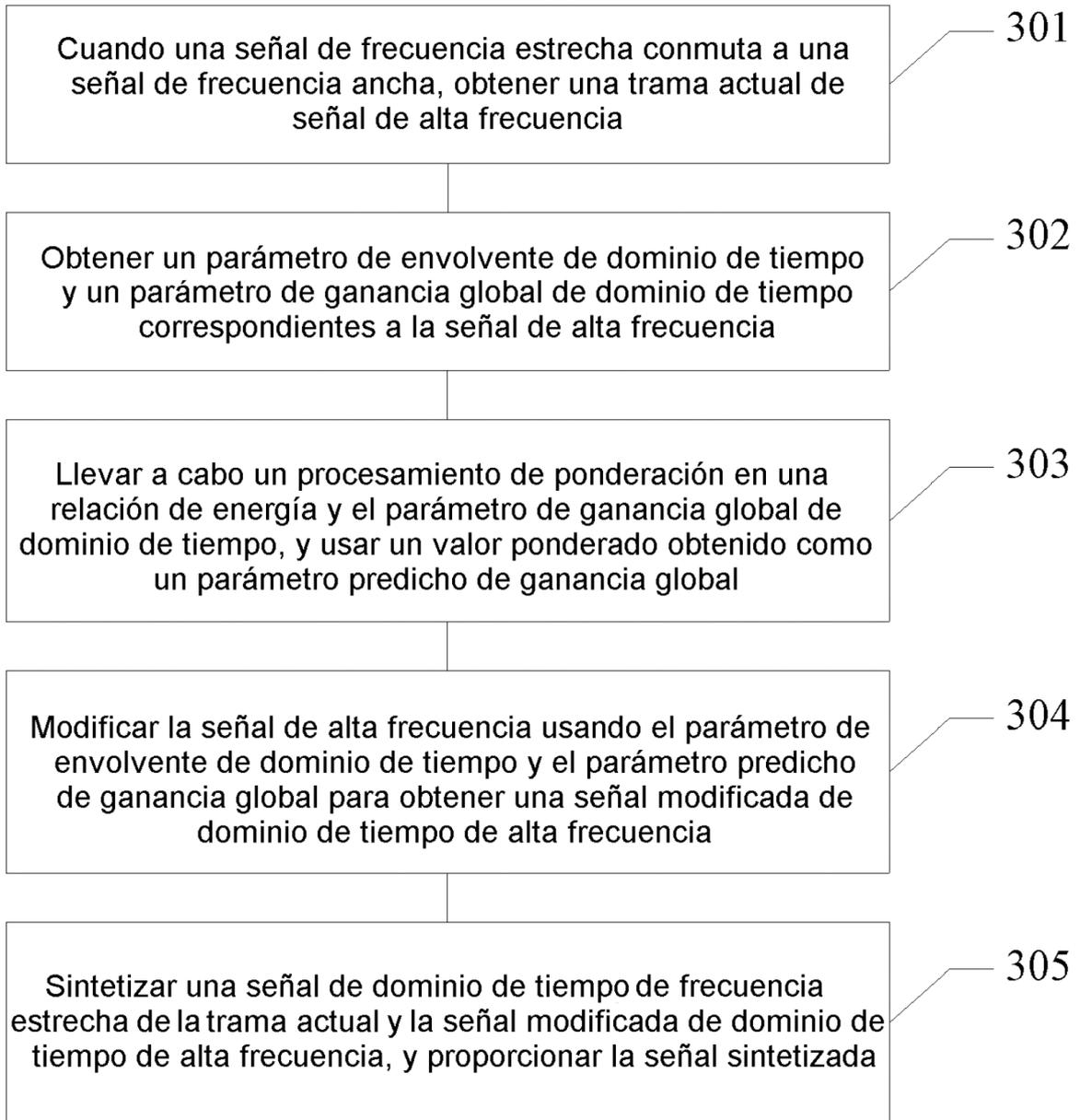


FIG. 3

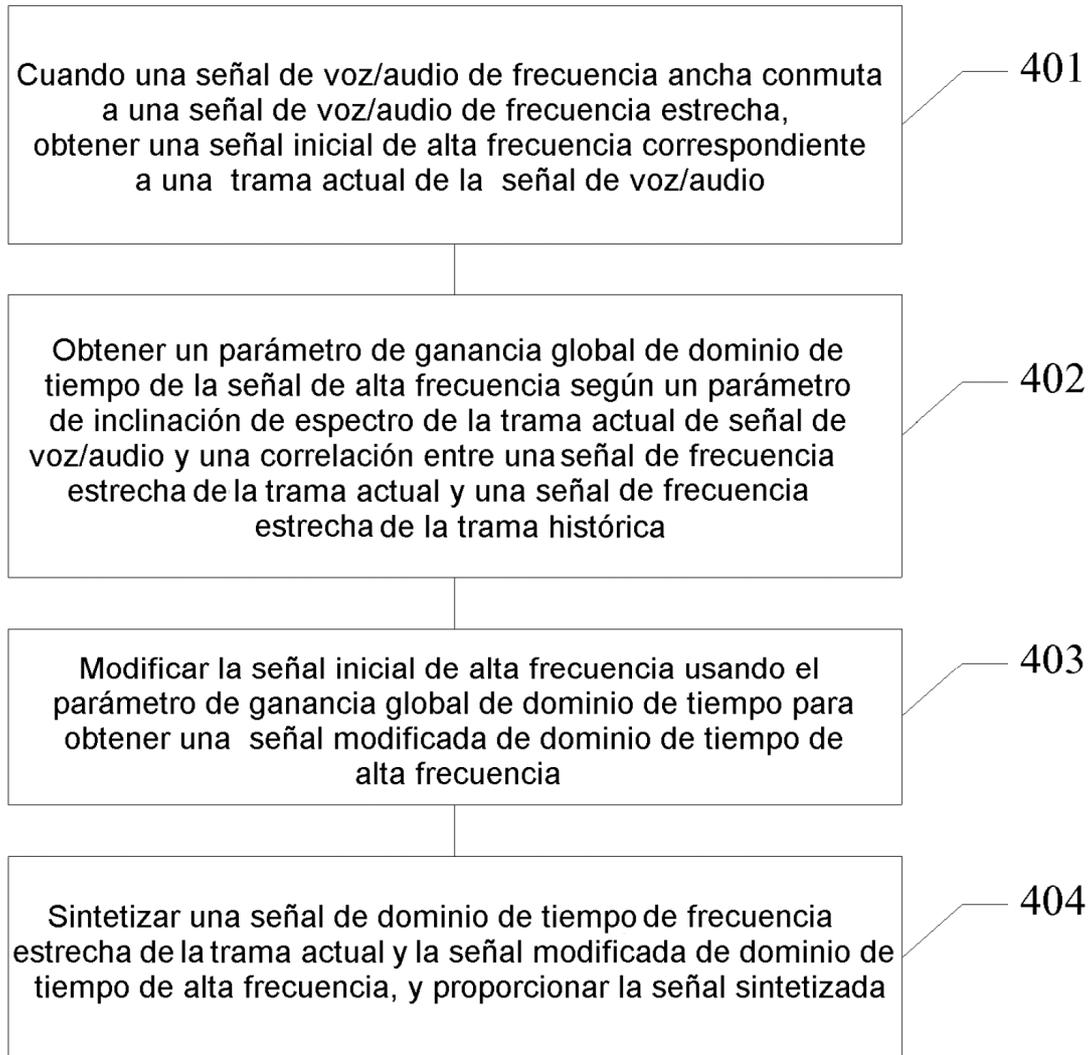


FIG. 4

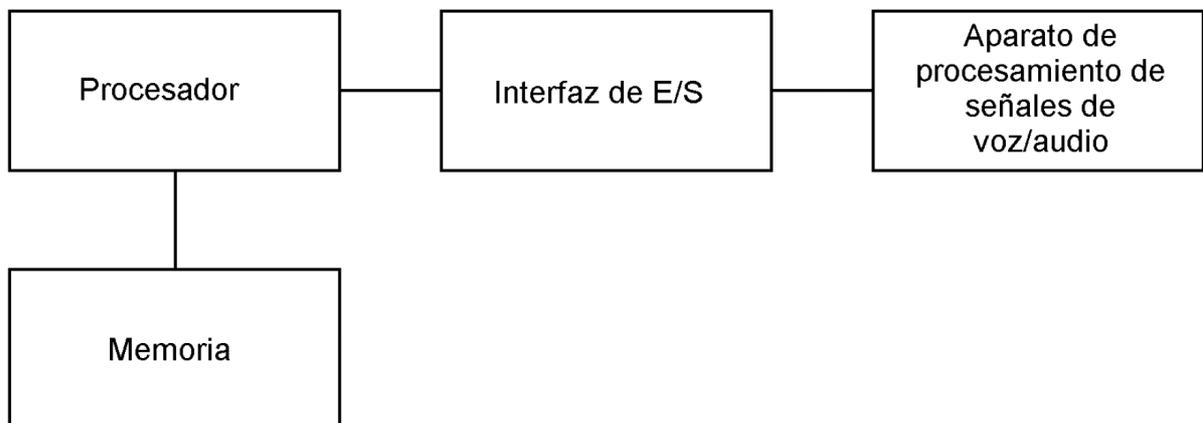


FIG. 5

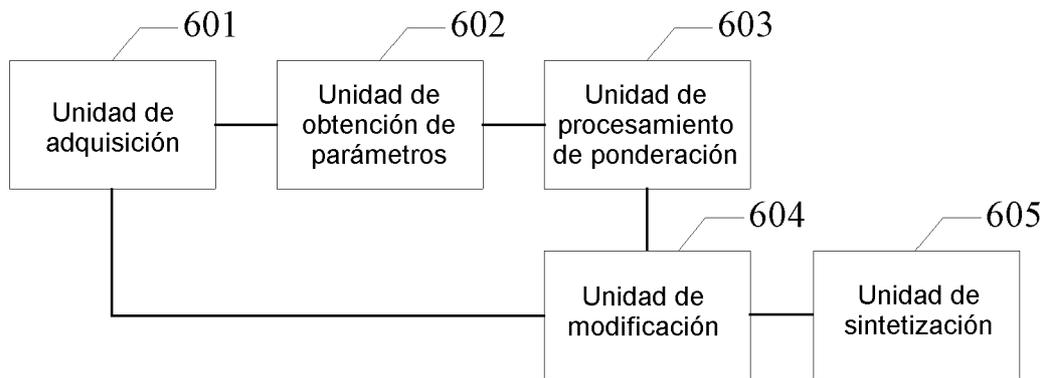


FIG. 6

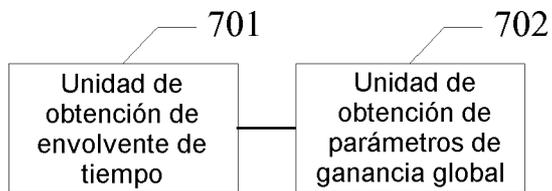


FIG. 7

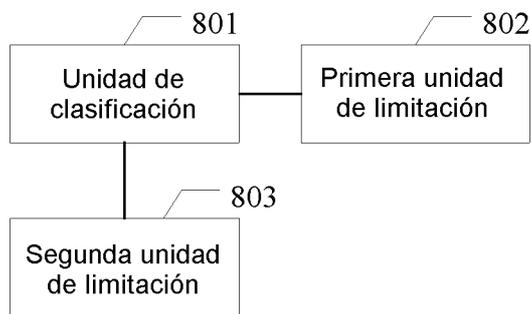


FIG. 8

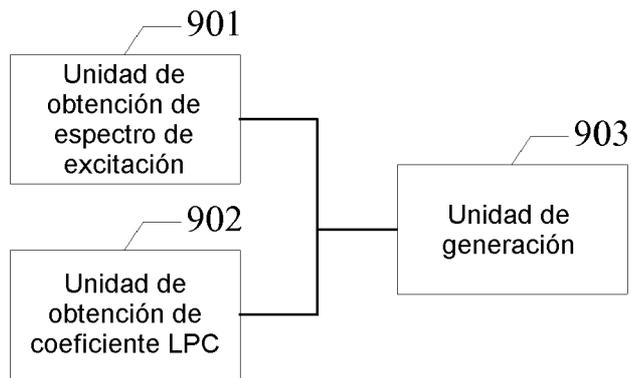


FIG. 9

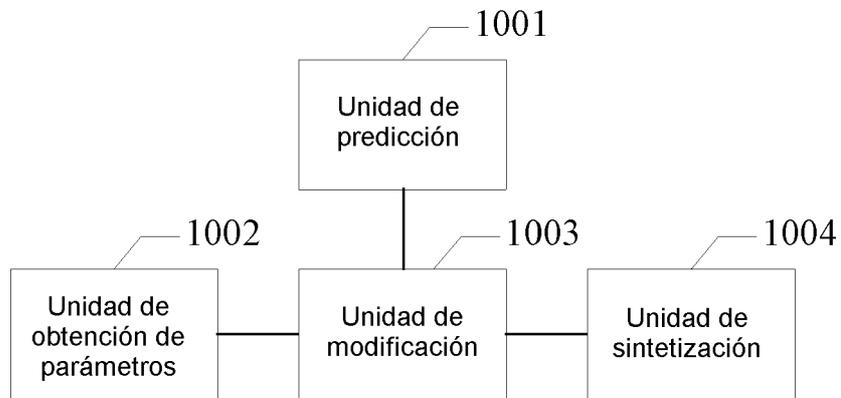


FIG. 10