

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 610 102**

51 Int. Cl.:

G10L 25/78 (2013.01)

G10L 25/90 (2013.01)

G10L 19/005 (2013.01)

G10L 25/87 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.12.2013 PCT/CN2013/089983**

87 Fecha y número de publicación internacional: **03.07.2014 WO2014101713**

96 Fecha de presentación y número de la solicitud europea: **19.12.2013 E 13867161 (5)**

97 Fecha y número de publicación de la concesión europea: **05.10.2016 EP 2927906**

54 Título: **Método y aparato para detectar una señal de voz**

30 Prioridad:

27.12.2012 CN 201210580541

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.04.2017

73 Titular/es:

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

72 Inventor/es:

XU, LIJING

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 610 102 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

Método y aparato para detectar una señal de voz.

Campo técnico

5 La presente invención se refiere al campo del procesamiento de audio y, más específicamente, a un método y un aparato para detectar una señal de voz.

Antecedentes

10 Para facilitar el análisis, en tecnologías de audio, un comienzo abrupto (comienzo abrupto) y/o final abrupto (final abrupto) de una señal de voz en esta memoria descriptiva indica/n dos tipos de situaciones: Una situación es que el final abrupto y comienzo abrupto ocurren en un par en una misma sección de un segmento de voz y duran un período de tiempo relativamente breve, y, para abreviar, se hace referencia a ello como "interrupción abrupta" en este contexto. Por ejemplo, en un proceso de habla, una pérdida de una parte de información en la mitad de un segmento de señales de voz puede provocar una interrupción abrupta. La otra situación es que el comienzo abrupto ocurra individualmente o que el final abrupto ocurra individualmente, y, para abreviar, se hace referencia a ello como "comienzo abrupto" o "final abrupto" en este contexto. Por ejemplo, ocurre un comienzo abrupto de una señal de voz cuando comienza el habla u ocurre un final abrupto de una señal de voz cuando finaliza el habla. En la siguiente descripción, una excepción abrupta de una señal de voz puede incluir una de las siguientes situaciones: interrupción abrupta, comienzo abrupto o final abrupto de una señal de voz.

20 La excepción abrupta de una señal de voz principalmente está provocada por una pérdida de paquetes y una determinación errónea del VAD (detector de actividad de voz) en un proceso de procesamiento de señal y puede causar daños en la semántica (semántico) y sintaxis (sintáctico) de la señal de voz después de que se restaura la señal de voz. Debido a que la semántica y la sintaxis están relacionadas con el contenido del lenguaje (contenido del lenguaje), en comparación con un examinador de idioma no nativo, a un examinador de idioma nativo le afecta más un comienzo abrupto o final abrupto de una señal de voz. Cuando se utiliza un modelo de evaluación de calidad de voz para evaluar la calidad de una señal de voz, en general, no se analiza el contenido del lenguaje y, por lo tanto, no se puede reflejar un impacto de la excepción abrupta de una señal de voz en la calidad acústica. Para resolver este problema, además de un modelo de evaluación básico, es necesario que se pueda detectar una excepción abrupta de una señal de voz, para que la evaluación de calidad se realice sobre una excepción abrupta individual de una señal de voz que ocurre en todas las señales de voz.

En la técnica anterior, la precisión al detectar una excepción abrupta de una señal de voz es relativamente baja.

30 El documento WO 2002/047068 A2 describe una técnica de clasificación de habla para una clasificación robusta de modos de discurso variables con el fin de permitir un rendimiento máximo de las técnicas de codificación de tasa de bits variable multimodo. Un clasificador de habla clasifica de manera precisa un alto porcentaje de segmentos de habla para codificarlos a tasas de bits mínimas, cumpliendo con requisitos de tasas de bits más bajas. Una clasificación de habla de alta precisión produce una tasa de bits codificada promedio más baja, y una mayor calidad de habla descodificada. El clasificador de habla considera una cantidad máxima de parámetros para cada trama de habla, generando clasificaciones de modo de habla numerosas y precisas para cada trama. El clasificador de habla clasifica correctamente numerosos modos de habla en condiciones ambientales variables. El clasificador de habla ingresa parámetros de clasificación desde componentes externos, genera parámetros de clasificación interna a partir de los parámetros de entrada, establece un umbral de Función de Coeficiente de Autocorrelación Normalizada y selecciona un analizador de parámetros según el entorno de la señal, y luego analiza los parámetros para crear una clasificación de modo de habla.

45 El documento US 5.774.847 describe que en los métodos y aparatos para distinguir señales estacionarias de señales no estacionarias, se determina un conjunto de coeficientes de codificación de predicción lineal (LPC, por sus siglas en inglés) con propiedades espectrales de la señal para cada uno de los múltiples intervalos de tiempo sucesivos, incluyendo un intervalo de tiempo actual. Los coeficientes LPC se promedian entre múltiples intervalos de tiempo sucesivos que anteceden al intervalo de tiempo actual, y se determina una correlación cruzada de los coeficientes LPC para el intervalo de tiempo actual con los coeficientes LPC promediados. La señal se define como estacionaria en el intervalo de tiempo actual cuando la correlación cruzada excede un valor umbral, y se define como no estacionaria en el intervalo de tiempo actual cuando la correlación cruzada es menor que el valor umbral. Los métodos y aparatos se pueden aplicar especialmente a la detección de transiciones entre un estado de ausencia de habla, caracterizado por una señal estacionaria, y un estado de presencia de habla, caracterizado por una señal no estacionaria.

Compendio

55 En vista de lo anterior, las realizaciones de la presente invención ofrecen un método y un aparato para detectar una señal de voz, de modo que pueda resolverse un problema de precisión relativamente baja al detectar una excepción abrupta de una señal de voz.

Según un primer aspecto, se ofrece un método para detectar una señal de voz, que incluye: realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía de cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, donde la excepción abrupta potencial de una señal de voz incluye una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto, y final abrupto de una señal de voz, y en donde una interrupción abrupta corresponde a una ocurrencia de un par que comprende un final abrupto y un comienzo abrupto en la misma sección de un segmento de la señal de voz; realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que incluye el primer período de tiempo meta es un segundo período de tiempo meta; y procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, en donde el procesamiento de la característica de tono comprende realizar una transformada de Fourier rápida en cada uno de los segundos períodos de tiempo para adquirir un espectro de densidad de potencia, determinar un punto máximo local según el espectro de densidad de potencia, y analizar un segmento de un intervalo de dominio de frecuencia centrado en el punto máximo local para determinar si existe un componente tonal en una banda de frecuencia en la que está ubicado el punto máximo; y determinar, mediante el análisis de la característica de tono adquirida de al menos uno de los segundos períodos de tiempo incluyendo al menos uno de los primeros períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

En una primera forma de implementación posible, el método incluye: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural.

Con referencia a la primera forma de implementación posible del primer aspecto, en una segunda forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-1)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde a_1 y a_2 es un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

En referencia a la primera forma de implementación posible del primer aspecto, en una tercera forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-2)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidos como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

En referencia a la primera forma de implementación posible del primer aspecto, en una cuarta forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-3)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{\text{ésima}}$ y la trama $(i-3)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

En referencia a la primera forma de implementación posible del primer aspecto, en una quinta forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2)$ y $(trama_energía_corta(i-1) < a_1)$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

En referencia a la primera forma de implementación posible del primer aspecto, en una sexta forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2)$ y $(trama_energía_corta(i-2) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidos como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

En referencia a la primera forma de implementación posible del primer aspecto, en una séptima forma de implementación posible, el método incluye: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2)$ y $(trama_energía_corta(i-3) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas $(i-1)^{ésima}$ a la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama, y la 2^{da} trama están preestablecidos como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

En referencia al primer aspecto o a cualquiera de las formas de implementación posibles anteriores del primer aspecto, en una octava forma de implementación posible, el método incluye: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total (SPL; por sus siglas en inglés) $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{ésima}$, donde la trama $k^{ésima}$ es el segundo período de tiempo $k^{ésimo}$ en los múltiples segundos períodos de tiempo y k es un número natural.

En referencia a la octava forma de implementación posible del primer aspecto, en una novena forma de implementación posible, el método incluye: si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz; o, si una característica de tono del segundo período de tiempo meta cumple con $(a_4 \leq spl_tonal(k) < a_3)$ y $(spl_total(k) \geq a_5)$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz, donde a_3 , a_4 y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente

En referencia a la octava forma de implementación posible del primer aspecto, en una décima forma de implementación posible, el método incluye: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con: $(spl_tonal(k+1) \geq a_1)$, $(spl_tonal(k) < a_8)$, $(spl_tonal(k+1) - spl_no_tonal(k) > 0)$, y $(spl_no_tonal(k-1) < a_9)$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un comienzo abrupto real de una señal de voz; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con: $(spl_tonal(k+2) \geq a_{10})$, $(spl_total(k+1) < a_{11})$, $(spl_tonal(k+2) - spl_no_tonal(k+1) > 0)$, y $(spl_no_tonal(k) < a_{12})$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un comienzo abrupto real de una señal de voz, donde a_7 a a_{12} es un séptimo umbral preestablecido a un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k+1) - spl_total(k-1) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k) - spl_total(k-2) \geq a_6)$, $(spl_total(k) > spl_total(k-1))$, $(spl_total(k-1) > spl_total(k-2))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que el $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente, y a_6 es un sexto umbral preestablecido; o si la característica de tono del segundo período de tiempo no cumple con ninguna de las dos condiciones anteriores determinar que el $spl_tonal(k)$ aumenta ligeramente.

En referencia a la octava forma de implementación posible del primer aspecto, en una undécima forma de implementación posible, el método incluye: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con: $(spl_tonal(k-1) \geq a_1)$, $(spl_tonal(k) < a_8)$, $(spl_tonal(k-1) - spl_no_tonal(k) > 0)$, y $(spl_no_tonal(k+1) < a_9)$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz, donde $k \geq 1$; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con: $(spl_tonal(k-2) \geq a_{10})$, $(spl_tonal(k-1) < a_{11})$, $(spl_tonal(k-1) - spl_no_tonal(k-2) > 0)$, y $(spl_no_tonal(k) < a_{12})$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz, donde $k \geq 2$, y a_7 a a_{12} es un séptimo umbral preestablecido a un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-1) - spl_total(k) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-2) - spl_total(k) \geq a_6)$, $(spl_total(k-1) > spl_total(k))$, y $(spl_total(k-2) > spl_total(k-1))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente, o si ninguna de las dos condiciones anteriores se cumple determinar que $spl_total(k)$ disminuye ligeramente, donde a_6 es un sexto umbral preestablecido.

Según un segundo aspecto, se ofrece un aparato para detectar una señal de voz, que incluye una primera unidad de detección, una unidad de entramado, y una segunda unidad de detección, donde la primera unidad de detección está configurada para: realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía de cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, donde la excepción abrupta potencial de una señal de voz incluye una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto, y final abrupto de una señal de voz, y en donde una interrupción abrupta corresponde a una ocurrencia de un par que comprende un final abrupto y un comienzo abrupto en la misma sección de un segmento de la señal de voz; la unidad de entramado está configurada para realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que incluye el primer período de tiempo meta es un segundo período de tiempo meta; y la segunda unidad de detección está configurada para: procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, en donde el procesamiento de característica de tono comprende realizar una transformada de Fourier rápida en cada uno de los segundos períodos de tiempo para adquirir un espectro de densidad de potencia, determinar un punto máximo local según el espectro de densidad de potencia, y analizar un segmento de un intervalo de dominio de frecuencia centrado en el punto máximo local para determinar si existe un componente tonal en una banda de frecuencia en la que está ubicado el punto máximo local; en donde la segunda unidad de detección además está configurada para determinar, mediante el análisis de la característica de tono adquirida de al menos uno de los segundos períodos de tiempo que incluye al menos un primer período de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

En una primera forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para: realizar el entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-1)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$ determinar que la trama $i^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

En referencia al segundo aspecto, en una segunda forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; donde el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-2)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como los primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

En referencia al segundo aspecto, en una tercera forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; donde el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-3)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama, y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

En referencia al segundo aspecto, en una cuarta forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2)$ y $(trama_energía_corta(i-1) < a_1)$ determinar que la trama $i^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

En referencia al segundo aspecto, en una quinta forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2)$ y $(trama_energía_corta(i-2) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0, y la 1^{era} trama están preestablecidas como los primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

En referencia al segundo aspecto, en una sexta forma de implementación posible, la primera unidad de detección incluye un primer módulo de adquisición y un primer módulo de determinación, donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es el $i^{ésimo}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2)$ y $(trama_energía_corta(i-3) \geq a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

En referencia al segundo aspecto o a cualquiera de las formas de implementación posibles anteriores del segundo aspecto, en una séptima forma de implementación posible, la segunda unidad de detección incluye un segundo módulo de adquisición y un segundo módulo de determinación, donde el segundo módulo de adquisición está configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico, y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{ésima}$, donde la trama $k^{ésima}$ es el $k^{ésimo}$ segundo período de tiempo en los múltiples segundos períodos de tiempo, y k es un número natural; y el segundo módulo de determinación está configurado para: si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz; o, si una característica de tono del segundo período de tiempo meta cumple con $(a_4 \leq spl_tonal(k) < a_3)$ y $(spl_total(k) \geq a_5)$, determinar que la excepción abrupta potencial de una señal incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz, donde a_3 , a_4 , y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente.

En referencia al segundo aspecto o a cualquiera de las formas de implementación posibles anteriores del segundo aspecto, en una octava forma de implementación posible, la segunda unidad de detección incluye un segundo módulo de adquisición y un segundo módulo de determinación, donde el segundo módulo de adquisición está configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{ésima}$, donde la trama $k^{ésima}$ es el $k^{ésimo}$ segundo período de tiempo en los múltiples segundos períodos de tiempo y k es un número natural; y el segundo módulo de determinación está configurado para: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k+1) \geq a_7),$$

$$(spl_tonal(k) < a_8),$$

$$(spl_tonal(k+1) - spl_no_tonal(k) > 0), \text{ y}$$

$$(spl_no_tonal(k-1) < a_9),$$

- 5 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k+2) \geq a_{10}),$$

10 $(spl_tonal(k+1) < a_{11}),$

$$(spl_tonal(k+2) - (spl_no_tonal(k+1)) > 0), \text{ y}$$

$$(spl_no_tonal(k) < a_{12}),$$

- 15 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz, donde a_7 a a_{12} es un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y el segundo módulo de determinación está además configurado para determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k) - spl_total(k-1)) \geq a_6$ y $spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente; o si la
- 20 característica de tono del segundo período de tiempo cumple con $(spl_total(k) - spl_total(k-2)) \geq a_6$, $(spl_total(k) > spl_total(k-1))$, $(spl_total(k-1) > spl_total(k-2))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumenta ligeramente, y a_6 es un sexto umbral preestablecido; o si la característica de tono del segundo período de tiempo no cumple ninguna de las dos
- 25 condiciones anteriores determinar que $spl_tonal(k)$ aumenta ligeramente.

- En referencia al segundo aspecto o a cualquiera de las formas de implementación posibles anteriores del segundo aspecto, en una novena forma posible de implementación, la segunda unidad de detección incluye un segundo módulo de adquisición y un segundo módulo de determinación, donde el segundo módulo de adquisición está configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo
- 30 según un orden cronológico; y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{\text{ésima}}$, donde la trama $k^{\text{ésima}}$ es el $k^{\text{ésimo}}$ segundo período de tiempo en los múltiples segundos períodos de tiempo y k es un número natural; y el segundo módulo de determinación está configurado para: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y
- 35 $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k-1) \geq a_7),$$

$$(spl_tonal(k) < a_8),$$

$$(spl_tonal(k-1) - spl_no_tonal(k) > 0), \text{ y}$$

40 $(spl_no_tonal(k+1) < a_9),$

- determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz donde $(k \geq 1)$; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

- 45 $(spl_tonal(k-2) \geq a_{10}),$
- $$(spl_tonal(k-1) < a_{11}),$$
- $$(spl_tonal(k-1) - spl_no_tonal(k-2) > 0), \text{ y}$$
- $$(spl_no_tonal(k) < a_{12}),$$

- 50 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz, donde $k \geq 2$, y a_7 a a_{12} es un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y la determinación de si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente

rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-1)-spl_total(k) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-2)-spl_total(k) \geq a_6)$, $(spl_total(k-1) > spl_total(k))$, $(spl_total(k-2) > spl_total(k-1))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si ninguna de las dos condiciones anteriores se cumplen determinar que $spl_total(k)$ disminuye ligeramente, donde a_6 es un sexto umbral preestablecido.

- 10 Según la solución técnica mencionada anteriormente, se puede determinar una excepción abrupta real de una señal de voz al detectar primero una excepción abrupta potencial de una señal de voz, y analizar adicionalmente una característica de tono de la excepción abrupta potencial de una señal de voz, de tal manera que la precisión en la detección de una excepción abrupta de una señal de voz mejore de manera eficaz.

Breve descripción de las figuras

- 15 Para describir las soluciones técnicas en las realizaciones de la presente invención de manera más clara, a continuación se describen brevemente los dibujos que acompañan esta memoria, los cuales resultan necesarios para describir las realizaciones de la presente invención. Según parece, los dibujos que acompañan la siguiente descripción simplemente muestran algunas realizaciones de la presente invención, y una persona con experiencia ordinaria en la técnica puede incluso obtener otros dibujos a partir de los dibujos que la acompañan sin requerir esfuerzos creativos.

La Figura 1A y la Figura 1B son pantallazos esquemáticos de resultados de detección de detectar una excepción abrupta de una señal de voz en tecnologías relacionadas;

La Figura 2A y la Figura 2B son pantallazos esquemáticos de resultados de detección de detectar una excepción abrupta de una señal de voz en tecnologías relacionadas;

- 25 La Figura 3 es un diagrama de flujo esquemático de un método para detectar una excepción abrupta de una señal de voz según una realización de la presente invención;

La Figura 4 es un diagrama de flujo esquemático de un método para detectar una excepción abrupta de una señal de voz según otra realización de la presente invención;

- 30 La Figura 5A y la Figura 5B son diagramas esquemáticos de curvas de distribución de niveles de presión sonora según otra realización de la presente invención;

La Figura 6A y la Figura 6B son diagramas esquemáticos de curvas de distribución de niveles de presión sonora según otra realización de la presente invención;

Cada una de las figuras 7A y 7B es un diagrama de bloque esquemático de un aparato para detectar una señal de voz según una realización de la presente invención; y

- 35 La Figura 8 es un diagrama de bloque esquemático de un aparato para detectar una señal de voz según otra realización de la presente invención.

Descripción de las realizaciones

A continuación se describen claramente las soluciones técnicas en las realizaciones de la presente invención haciendo referencia a los dibujos que acompañan esta memoria en las realizaciones de la presente invención. Según parece, las realizaciones descritas son algunas pero no todas las realizaciones de la presente invención. Cualquier otra realización obtenida por una persona con experiencia ordinaria en la técnica a partir de las realizaciones de la presente invención que no implique esfuerzos creativos estará comprendida dentro del alcance de protección de la presente invención.

- 45 La Figura 1A y la Figura 1B son pantallazos esquemáticos de resultados de detección de detectar una excepción abrupta de una señal de voz en tecnologías relacionadas. La Figura 1A muestra un resultado de detección definido manualmente mediante comparación con una voz original y la Figura 1B es un resultado de detección en la técnica anterior. En la Figura 1A y la Figura 1B, un eje horizontal representa puntos de muestreo y un eje vertical representa la amplitud normalizada. Para una interrupción abrupta que ocurre en un mismo segmento de señales de voz y que dura un período de tiempo relativamente breve, y con el fin de simplificar su representación, en la Figura 1A y la Figura 1B sólo se han marcado las ubicaciones de finales abruptos, tal y como lo indican los segmentos de línea 11 en las figuras. En comparación con el resultado de detección definido manualmente, en la Figura 1B, no se detecta una interrupción más abrupta de una señal de voz, que dura un breve período de tiempo y se indica mediante las flechas 12 en la figura.

- 55 La Figura 2A y la Figura 2B son pantallazos esquemáticos de resultados de detección de detectar una excepción abrupta de una señal de voz en tecnologías relacionadas. La Figura 2A muestra un resultado de detección definido

- manualmente mediante comparación con una voz original y la Figura 2B muestra un resultado de detección en la técnica anterior. En la Figura 2A y la Figura 2B, un eje horizontal representa puntos de muestreo y un eje vertical representa la amplitud normalizada. Para una interrupción abrupta que ocurre en un mismo segmento de señales de voz y que dura un período de tiempo relativamente breve, y con el fin de simplificar su representación, en la Figura 2A y la Figura 2B sólo se han marcado las ubicaciones de finales abruptos y, además, también se han marcado comienzos abruptos y finales abruptos que ocurren individualmente, tal y como lo indican los segmentos de línea 21 en las figuras. En comparación con el resultado de detección definido manualmente, en la Figura 2B, no se detecta un comienzo abrupto o final abrupto de una señal de voz con energía relativamente baja, lo cual se indica mediante flechas 22 en la figura.
- 10 Para resolver un problema, de la tecnología relacionada, sobre la precisión relativamente baja al detectar una excepción abrupta de una señal de voz, las realizaciones de la presente invención ofrecen un método para detectar una señal de voz, donde una excepción abrupta de una señal de voz se puede detectar en base a un análisis de una característica de tono, de manera tal que la precisión al detectar la excepción abrupta de una señal de voz se mejore de manera eficaz.
- 15 La Figura 3 es un diagrama de flujo esquemático de un método 30 para detectar una excepción abrupta de una señal de voz según una realización de la presente invención. El método 30 incluye el siguiente contenido:
- E31. Realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía en cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, donde la excepción abrupta potencial de una señal de voz incluye una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto, y final abrupto de una señal de voz.
- 20 Tal y como se menciona antes, una excepción abrupta de una señal de voz puede incluir una de las siguientes situaciones: interrupción abrupta, comienzo abrupto y final abrupto de una señal de voz. Un primer período de tiempo que incluye una excepción abrupta potencial de una señal de voz puede determinarse comparando la energía de los múltiples primeros períodos de tiempo y comparando la energía de un primer período de tiempo específico y un umbral preestablecido y valores similares. En este contexto también se hace referencia al primer período de tiempo que incluye una excepción abrupta potencial de una señal de voz como un primer período de tiempo meta.
- 25 E32. Realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que incluye el primer período de tiempo meta es un segundo período de tiempo meta.
- E33. Procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, y determinar, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que incluye al menos uno de los segundos períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.
- 35 Una excepción abrupta de una señal de voz también se denomina excepción abrupta, para abreviar, una excepción abrupta potencial de una señal de voz también se denomina excepción abrupta potencial para abreviar, y un comienzo abrupto de una señal de voz o un final abrupto de una señal de voz también se denomina, para abreviar, comienzo abrupto o final abrupto, respectivamente. Una interrupción abrupta es un final abrupto y un comienzo abrupto que ocurren en pares en una misma sección de un segmento de voz y duran un período de tiempo relativamente breve. Un comienzo abrupto o un final abrupto implica que el comienzo abrupto ocurre individualmente o que el final abrupto ocurre individualmente, respectivamente.
- 40 Cuando la longitud de trama de segundo período de tiempo es una integral múltiple del primer período de tiempo, después de realizar el entramado de la muestra de voz continua en una unidad de longitud de trama de segundo período de tiempo, se obtienen uno o más segundos períodos de tiempo. Un segundo período de tiempo puede incluir múltiples primeros períodos de tiempo. Sin embargo, en todos los segundos períodos de tiempo, uno o algunos segundos períodos de tiempo pueden incluir, por separado, un primer período de tiempo meta. Este tipo de segundo período de tiempo es un objeto para la detección y análisis detallado en esta realización de la presente invención y también se hace referencia al mismo en esta memoria como un segundo período de tiempo meta. Como una tecnología existente, para eliminar un efecto de límite durante el procesamiento de señal de voz, se pueden superponer parcialmente dos segundos períodos de tiempo próximos. Por ejemplo, si un primer segundo período de tiempo va desde el punto de muestreo 0 al punto de muestreo 511°, un segundo segundo período de tiempo va desde el punto de muestreo 255° al punto de muestreo 767°. A continuación, el procesamiento de característica de tono que incluye una transformada de Fourier rápida y operación similar se realiza en cada uno de los segundos períodos de tiempo, y luego, se analiza si uno o más segundos períodos de tiempo cumplen una relación predeterminada, de tal manera que pueda determinarse si una excepción abrupta potencial de una señal de voz incluida en un segundo período de tiempo meta en uno o más de los segundos períodos de tiempo es una excepción
- 45
- 50
- 55
- 60

abrupta real de una señal de voz, donde se conoce que el segundo período de tiempo meta determinado incluye un primer período de tiempo meta.

Esta realización de la presente invención ofrece un método para detectar una señal de voz, donde se puede determinar una excepción abrupta real de una señal de voz al detectar primero una excepción abrupta potencial de una señal de voz, y al analizar adicionalmente una característica de tono de la excepción abrupta potencial de una señal de voz, de manera que la precisión en la detección de una excepción abrupta de una señal de voz se mejore de manera eficaz.

La Figura 4 es un diagrama de flujo esquemático de un método 40 para detectar una excepción abrupta de una señal de voz según otra realización de la presente invención. El método 40 incluye el siguiente contenido:

E41. Realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo.

El entramado se realiza en un segmento de una muestra de voz continua en una unidad de longitud de trama de primer período de tiempo para obtener múltiples primeros períodos de tiempo continuos. Se hace referencia a la trama $i^{\text{ésima}}$ en los múltiples primeros períodos de tiempo como el $i^{\text{ésimo}}$ primer período de tiempo y a continuación y para abreviar se hace referencia al mismo como trama $i^{\text{ésima}}$.

E42. Calcular la energía de cada uno de los primeros períodos de tiempo.

Suponiendo que $\text{trama_energía_corta}(i)$ representa la energía de la trama $i^{\text{ésima}}$, donde i es un número natural:

$$\text{trama_energía_corta}(i) = 10 * \lg \sum_{n=0}^{N_1-1} \text{tiempo_señal_breve}^2(n) \quad \text{Fórmula 1}$$

donde $\text{tiempo_señal_breve}(n)$ representa una señal de entrada en la trama $i^{\text{ésima}}$, n representa puntos de muestreo, N_1 representa la longitud de trama de primer período de tiempo, y en esta realización se establecen 32 puntos de muestreo. Al seleccionar un primer período de tiempo de una longitud de trama apropiada, se puede mejorar la precisión en la detección o se puede equilibrar la relación entre la precisión en la detección y la complejidad de un algoritmo.

E43. Determinar un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros períodos de tiempo. La etapa E43 puede incluir la etapa E43-1 o la etapa E43-2.

Se detecta energía de varias tramas previas a la trama $i^{\text{ésima}}$ y energía de la trama $i^{\text{ésima}}$, donde la trama $(i-1)^{\text{ésima}}$ es una trama previa a la trama $i^{\text{ésima}}$, la trama $(i-2)^{\text{ésima}}$ es una trama previa a la trama $(i-1)^{\text{ésima}}$, y la trama $(i-3)^{\text{ésima}}$ es una trama previa a la trama $(i-2)^{\text{ésima}}$, y así sucesivamente.

E43-1. Si la energía de la trama $i^{\text{ésima}}$ disminuye rápidamente, es decir, si una de las siguientes condiciones se cumple, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz.

a) $(\text{trama_energía_corta}(i-1) - \text{trama_energía_corta}(i) \geq a_2)$ y $(\text{trama_energía_corta}(i) < a_1)$.

Generalmente, se preestablece que la trama 0 no es un primer período de tiempo meta que incluye un final abrupto potencial. Cuando $i \geq 1$, se puede determinar, según la condición a), si la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un final abrupto potencial.

b) $(\text{trama_energía_corta}(i-2) - \text{trama_energía_corta}(i) \geq a_2)$ y $(\text{trama_energía_corta}(i) < a_1)$ y

ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial, donde $i \geq 2$, y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

Por ejemplo, cuando $i=2$, la trama 0 y la 1^{era} trama ya están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial, y luego se puede determinar si la 2^{da} trama es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, y así sucesivamente.

c) $(\text{trama_energía_corta}(i-3) - \text{trama_energía_corta}(i) \geq a_2)$ y $(\text{trama_energía_corta}(i) < a_1)$ y

ninguna de las tramas $(i-1)^{\text{ésima}}$ a la trama $(i-3)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial, donde $i \geq 3$, y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

Por ejemplo, cuando $i=3$, la trama 0, la 1^{era} trama y la 2^{da} ya están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial, y luego se puede determinar si la 3^{era} trama es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, y así sucesivamente.

5 En la aplicación real, una muestra de voz continua es relativamente extensa y, generalmente, se procesa en un orden cronológico, y algunos primeros períodos de tiempo previos se pueden preestablecer como primeros períodos de tiempo que no incluyen un final abrupto potencial según uno de los métodos anteriores. Debido a que en la aplicación real cada trama dura sólo decenas de milisegundos, la omisión de resultados de detección de diversas tramas iniciales no afecta la precisión de la detección de voz.

10 E43-2. Comparar la energía de diversas tramas previas a la trama i ^{ésima} y la energía de la trama i ^{ésima}. Si la energía de la trama i ^{ésima} aumenta rápidamente, es decir, si una de las siguientes condiciones se cumple, determinar que la trama i ^{ésima} es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz.

d) $(trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2)$ y $(trama_energía_corta(i-1) < a_1)$, donde $i \geq 1$.

15 Generalmente, se preestablece que la trama 0 no es un primer período de tiempo meta que incluye un comienzo abrupto potencial. Cuando $i \geq 1$, se puede determinar, según la condición d), si la 1^{era} trama es el primer período de tiempo meta que incluye un comienzo abrupto potencial.

e) $(trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2)$ y $(trama_energía_corta(i-2) < a_1)$ y

ni la trama $(i-1)$ ^{ésima} ni la trama $(i-2)$ ^{ésima} es un primer período de tiempo meta que incluye un comienzo abrupto potencial, donde $i \geq 2$, y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

20 Por ejemplo, cuando $i=2$, ya se ha preestablecido si la trama 0 y la 1^{era} trama han sido preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial, y luego se puede determinar si la 2^{da} trama es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, y así sucesivamente.

f) $(trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2)$ y $(trama_energía_corta(i-3) < a_1)$ y

25 ninguna de las tramas comprendidas entre la trama $(i-1)$ ^{ésima} y la trama $(i-3)$ ^{ésima} es un primer período de tiempo meta que incluye un comienzo abrupto potencial, donde $i \geq 3$, y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

30 Por ejemplo, cuando $i=3$, la trama 0, la 1^{era} trama y la 2^{da} trama ya están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial, y luego se puede determinar si la 3^{era} trama es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, y así sucesivamente.

En la aplicación real, una muestra de voz continua es relativamente extensa y, generalmente, se procesa en un orden cronológico, y algunos primeros períodos de tiempo previos se pueden preestablecer como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial según uno de los métodos anteriores.

35 Debido a que en la aplicación real cada trama dura sólo decenas de milisegundos, la omisión de resultados de detección de diversas tramas iniciales no afecta la precisión de la detección de voz.

40 En esta realización de la presente invención $a_1 = 38$ y $a_2 = 40$. A_1 y a_2 , a_3 hasta a_{12} en las siguientes realizaciones, y valores similares, son todos umbrales preestablecidos en las condiciones y, generalmente, necesitan ser determinados teniendo en consideración varios aspectos. Por ejemplo, los umbrales se obtienen adiestrando una gran cantidad de muestras según un tipo de una secuencia de prueba. Además, los umbrales son significativos para el volumen de sonido de la secuencia de prueba.

En las condiciones b, c, e, y f, si las diversas tramas previas a la trama i ^{ésima} son una excepción abrupta potencial, es una condición conocida.

Los procesos anteriores desde E41 a E43 son una detección aproximada, y luego, desde la E44 a la E46, se realiza la detección detallada.

45 E44. Realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde cada longitud de trama de segundo período de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y realizar un procesamiento de detección de tono en cada una de los segundos períodos de tiempo según un orden cronológico.

50 En la aplicación real, una muestra de voz continua procesada es relativamente extensa y, generalmente, se pueden detectar múltiples excepciones abruptas potenciales. A partir de lo anterior se conoce que un segundo período de tiempo incluye múltiples primeros períodos de tiempo, y el segundo período de tiempo es más extenso que el primer período de tiempo. Por lo tanto, el segundo período de tiempo también se usa para indicar un período de tiempo extenso, y el primer período de tiempo también se usa para indicar un período de tiempo breve.

El entramado se realiza sobre la muestra de voz continua en una unidad de longitud de trama de segundo período de tiempo para obtener uno o más segundos períodos de tiempo, donde algunos segundos períodos de tiempo incluyen los primeros períodos de tiempo meta determinados mediante una detección aproximada, los primeros períodos de tiempo meta incluyen una excepción abrupta potencial de una señal de voz, y también se hace referencia a estos segundos períodos de tiempo como segundos períodos de tiempo meta. Se hace referencia a la trama $k^{\text{ésima}}$ en los múltiples segundos períodos de tiempo como el segundo período de tiempo $k^{\text{ésimo}}$ y, a continuación, se hace referencia al mismo como trama $k^{\text{ésima}}$ para abreviar. La trama $(k-2)^{\text{ésima}}$, la trama $(k-1)^{\text{ésima}}$, la trama $k^{\text{ésima}}$, la trama $(k+1)^{\text{ésima}}$, y la trama $(k+2)^{\text{ésima}}$ son múltiples segundos períodos de tiempo dispuestos en orden.

Una etapa de procesamiento de detección de tono incluye: realizar una transformada de Fourier rápida (FFT, por sus siglas en inglés) en cada uno de los segundos períodos de tiempo para adquirir un espectro de densidad de potencia; determinar un punto máximo local según el espectro de densidad de potencia; y analizar un segmento de un intervalo de dominio de frecuencia centrado en el punto máximo local, para determinar si existe un componente tonal en una banda de frecuencia en la que se ubica el punto máximo local. En esta etapa, se utiliza un algoritmo de detección de tono del modelo psicoacústico 1 del MPEG (Moving Pictures Experts Group). Para descripciones detalladas, se hace referencia a la etapa 1 y etapa 4 del documento 11173-3 y Anexo D.1 (modelo psicoacústico 1) (modelo psicoacústico 1) de la ISO/IEC (Organización Internacional de Normalización/Comisión Electrotécnica Internacional).

Lo que resulta especial en esta realización de la presente invención es que no sólo se analiza un nivel de presión sonora total, es decir, una característica, de una trama actual, sino que también se analiza de forma separada un componente tonal y un componente no tonal de la trama actual. A continuación, el componente tonal y el componente no tonal se utilizan para calcular otras dos características de tono: un nivel de presión sonora de componente tonal y un nivel de presión sonora de componente no tonal, respectivamente. Se puede conocer una situación de distribución de un componente tonal y un componente no tonal de cada uno de los segundos períodos de tiempo en un dominio de frecuencia mediante la detección del componente tonal, y luego se puede calcular un nivel de presión sonora de componente tonal y un nivel de presión sonora de componente no tonal.

Las etapas posteriores en esta realización de la presente invención se utilizan para además determinar si una excepción abrupta potencial de una señal de voz es una excepción abrupta real de una señal de voz. Por ejemplo, a pesar de que la trama $(k-1)^{\text{ésima}}$ puede no incluir un primer período de tiempo que incluya una excepción abrupta potencial de una señal de voz, la trama $(k-1)^{\text{ésima}}$ es un segundo período de tiempo próximo a la trama $k^{\text{ésima}}$, y, por lo tanto, es necesario calcular un nivel de presión sonora total, un nivel de presión sonora de componente tonal y un nivel de presión sonora de componente no tonal de la trama $(k-1)^{\text{ésima}}$, para que se aplique a una o más de una de las condiciones determinantes citadas a continuación, determinando así si la excepción abrupta potencial de una señal de voz incluida en un primer período de tiempo meta incluido en la trama $k^{\text{ésima}}$ es una excepción abrupta real de una señal de voz.

E45. Después del procesamiento de detección de tono, adquirir un nivel de presión sonora total, un nivel de presión sonora de componente tonal, y un nivel de presión sonora de componente no tonal de cada uno de los segundos períodos de tiempo.

E45-1. Adquirir un nivel de presión sonora total de la trama $k^{\text{ésima}}$ según la siguiente Fórmula 2.

Suponiendo que $spl_total(k)$ representa el nivel de presión sonora total de la trama $k^{\text{ésima}}$:

$$spl_total(k) = 10 * \lg \left(\sum_{f=0}^{N_2/2-1} 10^{\frac{pot_espec(f)}{10}} \right) dB \quad \text{Fórmula 2}$$

donde $pot_espec(f)$ representa un espectro de densidad de potencia del segundo período de tiempo $k^{\text{ésimo}}$, $f=0,1,2,\dots,(N_2/2-1)$, y N_2 indica la longitud del segundo período de tiempo, y en esta realización se establecen 512 puntos de muestreo. El nivel de presión sonora se corresponde con la intensidad sonora, donde mayor intensidad sonora naturalmente se corresponde con mayor energía. Por lo tanto, el nivel de presión sonora puede reflejar una situación de energía. En esta realización de la presente invención, la característica, es decir, el nivel de presión sonora total, se utiliza para reflejar la energía total del segundo período de tiempo.

E45-2. Adquirir un nivel de presión sonora de componente tonal según la siguiente Fórmula 3.

Suponiendo que $spl_total(k)$ representa un nivel de presión sonora de componente tonal de la trama $k^{\text{ésima}}$:

$$spl_tonal(k) = 10 * \lg \left(\sum_{n=0}^{N_k-1} \left(10^{\frac{pot_espec(f_tonal(n)-1)}{10}} + 10^{\frac{pot_espec(f_tonal(n))}{10}} + 10^{\frac{pot_espec(f_tonal(n)+1)}{10}} \right) \right) dB \quad \text{Fórmula 3}$$

donde N_k representa una cantidad de componentes tonales detectados en la trama actual, y las ubicaciones de los componentes tonales se marcan como $\{f_tonal(0), f_tonal(1), f_tonal(2), \dots, f_tonal(N_k)\}$.

La característica, es decir, el nivel de presión sonora de componente tonal, se utiliza para describir una situación de energía de un componente tonal en el segundo período de tiempo. Si $spl_tonal(k)$ es relativamente elevado, indica que la trama $k^{ésima}$ está ubicada en un área con componentes tonales relativamente ricos.

E45-3. Adquirir un nivel de presión sonora de componente no tonal según la siguiente Fórmula 4.

Suponiendo que $spl_no_tonal(k)$ representa un nivel de presión sonora de componente no tonal de la trama $k^{ésima}$:

$$spl_non_tonal(k) = 10 * \lg \left(\sum_{f \in \Phi_{tonal}} 10^{\frac{pot_espec(f)}{10}} \right) dB \quad \text{Fórmula 4}$$

donde Φ_{tonal} representa las ubicaciones de un componente tonal y un componente próximo del componente tonal en un dominio de frecuencia:

$$\Phi_{tonal} = \{f_tonal(0)-1, f_tonal(0), f_tonal(0)+1, f_tonal(1)-1, f_tonal(1), f_tonal(1)+1, f_tonal(2)-1, f_tonal(2), f_tonal(2)+1, \dots, f_tonal(N_k)-1, f_tonal(N_k), f_tonal(N_k)+1\}$$

Fórmula 5

La característica, es decir, el nivel de presión sonora de componente no tonal, se utiliza para describir una situación de energía de un componente no tonal en el segundo período de tiempo. Si $spl_no_tonal(k)$ es relativamente elevado, indica que la trama $k^{ésima}$ está ubicada en un área con componentes no tonales relativamente ricos.

En esta realización de la presente invención, el análisis de situación de energía se realiza particularmente en un componente tonal y un componente no tonal de cada uno de los segundos períodos de tiempo, lo que difiere de la técnica anterior. El análisis facilita determinar si la excepción abrupta potencial de una señal de voz incluida en el segundo período de tiempo es una excepción abrupta real de una señal de voz en la siguiente etapa.

E46. Determinar, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que incluye al menos un segundo período de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

Un método de determinación incluye E46-1 o E46-2. En E46-1, se puede determinar una interrupción abrupta real de una señal de voz, y en E46-2, se puede determinar un comienzo abrupto real o un final abrupto real de una señal de voz. Las etapas E46-1 y E46-2 se describen por separado a continuación:

E46-1. Si el nivel de presión sonora de componente tonal de la trama $k^{ésima}$ cumple con cualquiera de las siguientes condiciones, condición g y condición h, determinar que la excepción abrupta potencial incluida en el primer período de tiempo meta incluido en la trama $k^{ésima}$ es una interrupción abrupta real.

g) $spl_tonal(k)$ es lo suficientemente elevado, tal y como se expresa en la siguiente fórmula:

$$spl_tonal(k) \geq a_3 \quad \text{Fórmula 6}$$

h) $spl_tonal(k)$ es relativamente elevado, tal y como se expresa en la siguiente fórmula:

$$(a_4 \leq spl_tonal(k) < a_5) \vee (spl_total(k) \geq a_5) \quad \text{Fórmula 7}$$

En esta realización de la presente invención, $a_3 = 55$, $a_4 = 30$ y $a_5 = 58$.

Según la condición g o la condición h, se puede determinar de manera secuencial si una excepción abrupta potencial incluida en el primer período de tiempo meta incluido en cada segundo período de tiempo meta es una interrupción abrupta real.

Si $spl_tonal(k)$ y $spl_total(k)$ cumplen las condiciones anteriores, indica que la trama $k^{ésima}$ está ubicada en un área con componentes tonales relativamente ricos. En una situación normal, es imposible hallar cambios de energía repentinos y breves en detecciones aproximadas realizadas en un área relativamente con componentes tonales relativamente ricos. Si se puede detectar una interrupción de una señal de voz en una detección aproximada, indica que la interrupción detectada es una interrupción abrupta real.

La Figura 5A y la Figura 5B son diagramas esquemáticos de curvas de distribución de niveles de presión sonora según una realización de la presente invención. En referencia a la Figura 5A, 51 es una señal de entrada, un eje horizontal representa puntos de muestreo y un eje vertical representa la amplitud normalizada. La figura incluye una interrupción abrupta que ocurre en múltiples ubicaciones y que tiene una duración relativamente breve. En la Figura 5B, se ofrecen por separado curvas de un nivel de presión sonora total 52, un nivel de presión sonora 53 de componente tonal, y un nivel de presión sonora 54 de componente no tonal, donde un eje horizontal representa puntos de muestreo, y un eje vertical representa un valor de un nivel de presión sonora. Debido a que todas las características de niveles de presión sonora en ubicaciones de interrupción 55 en la Figura 5A cumplen con la condición anterior, esto indica que la interrupción en estas ubicaciones está ubicada en un área con componentes tonales relativamente ricos y que es una interrupción abrupta real.

E46-2. Para otro resultado detectado en la detección aproximada, incluyendo un comienzo abrupto o un final abrupto que ocurren individualmente, se puede determinar, según un cambio de un nivel de presión sonora de componente tonal de la trama $k^{\text{ésima}}$, si la excepción abrupta potencial de una señal de voz es una excepción abrupta real.

Para una señal de voz normal, se puede detectar un cambio repentino de energía relativamente evidente al comienzo de la detección aproximada. Sin embargo, un proceso variable en el que un componente tonal de la señal de voz normal aumenta de repente es inevitablemente una transición natural. Si $spl_tonal(k)$ aumenta excesivamente rápido, indica que el proceso variable en el que el componente tonal de la señal de voz normal aumenta de repente no es natural, y su comienzo correspondiente es un comienzo abrupto. Un principio para detectar un final abrupto es similar a éste.

La Figura 6A y la Figura 6B son diagramas esquemáticos de curvas de distribución de niveles de presión sonora según otra realización de la presente invención. En referencia a la Figura 6A, 61 es una señal de entrada, un eje horizontal representa puntos de muestreo y un eje vertical representa la amplitud normalizada. En la Figura 6B, se ofrece por separado un nivel de presión sonora total 62, un nivel de presión sonora 63 de componente tonal y un nivel de presión sonora 64 de componente no tonal. En la Figura 6B, una flecha 65 representa un cambio de tendencia de $spl_tonal(k)$ en una ubicación de comienzo natural y una flecha 66 representa un cambio de tendencia de $spl_tonal(k)$ en una ubicación de comienzo abrupto. Tal y como se muestra en la figura, el $spl_tonal(k)$ en la ubicación de comienzo abrupto aumenta rápidamente, y ocurre una transición natural en el cambio de tendencia de $spl_tonal(k)$ en la ubicación del comienzo natural.

Las etapas para detectar un comienzo abrupto incluyen E46-2-1 y E46-2-2. Si E46-2-1 es verdadera, se determina además si E46-2-2 es verdadera. Si E46-2-2 es verdadera, el comienzo abrupto potencial de una señal de voz es un comienzo abrupto real; y si E46-2-2 es falsa, el comienzo abrupto no es un comienzo abrupto real. Si E46-2-1 es falsa, no es necesario determinar si E46-2-2 es verdadera, y el comienzo abrupto potencial de una señal de voz ciertamente no es un comienzo abrupto real.

E46-2-1. Determinar si se cumple cualquiera de las condiciones siguientes j o m.

j) $(spl_total(k)-spl_total(k-1) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente.

m) $(spl_total(k)-spl_total(k-2) \geq a_6)$,

$(spl_total(k) > spl_total(k-1))$,

$(spl_total(k-1) > spl_total(k-2))$, y

$(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente.

Si se cumple cualquiera de las condiciones j o m, se determina que $spl_total(k)$ de la trama $k^{\text{ésima}}$ aumenta excesivamente rápido. Entonces, se realiza la etapa E46-2-2. Si no se cumple ninguna de las condiciones j o m, no es necesario determinar si E46-2-2 es verdadera, y el comienzo abrupto potencial de una señal de voz ciertamente no es un comienzo abrupto real.

Que el nivel de presión sonora total aumente ligeramente es diferente a que el nivel de presión sonora total aumente excesivamente rápido. El aumento lento se refiere a que no se cumple ninguna de las condiciones anteriores j y m para determinar que el aumento es excesivamente rápido. Se ha de notar específicamente en la presente memoria que, en el procesamiento real, se establecen inicialmente diversas tramas iniciales para que aumenten ligeramente, y que la determinación sólo comienza en una trama posterior a las diversas tramas anteriores. Debido a que cada trama dura sólo decenas de milisegundos en la aplicación real, se omiten los resultados de detección de las diversas tramas iniciales.

E46-2-2. Si se detecta, según la condición j o m, que uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, determinar si se cumple alguna de las siguientes condiciones: n y p.

n) $(spl_total(k+1) \geq a_7)$,

$(spl_tonal(k) < a_8)$,
 $(spl_tonal(k+1) - sp_no_tonal(k) > 0)$,
 y $(spl_no_tonal(k-1) < a_9)$.
 p) $(spl_tonal(k+2) \geq a_{10})$,
 5 $(spl_tonal(k+1) < a_{11})$,
 $(spl_tonal(k+2) - sp_no_tonal(k+1) > 0)$, y
 $(spl_no_tonal(k) < a_{12})$.

Si se cumple cualquiera de las condiciones, n o p, la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en la trama $k^{ésima}$ es un comienzo abrupto real de una señal de voz. Si no se cumple ninguna de las condiciones n o p, la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en la trama $k^{ésima}$ no es un comienzo abrupto real.

Además, las etapas de detección de final abrupto incluyen E46-2-3 y E46-2-4. Si E46-2-3 es verdadera, se determina además si E46-2-4 es verdadera. Si E46-2-4 es verdadera, el final abrupto potencial de una señal de voz es un final abrupto real; y, si E46-2-4 es falsa, el final abrupto potencial de una señal de voz no es un final abrupto real. Si E46-2-3 es falsa, no es necesario determinar si E46-2-4 es verdadera, y el final abrupto potencial de una señal de voz ciertamente no es un final abrupto real. E46-2-3.

Determinar si se cumple cualquiera de las condiciones q o r.

q) $(spl_total(k-1) - spl_total(k) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente.

r) $(spl_total(k-2) - spl_total(k) \geq a_6)$,
 $(spl_total(k-1) > spl_total(k))$,
 $(spl_total(k-2) > spl_total(k-1))$, y

25 $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente.

Si $spl_tonal(k)$ disminuye excesivamente rápido, indica que $spl_total(k)$ de la trama $k^{ésima}$ disminuye excesivamente rápido. Entonces, se realiza la etapa E46-2-4. Si no se cumple ninguna de las condiciones q o r, no es necesario determinar además si E46-2-4 es verdadera, y el final abrupto potencial de una señal de voz ciertamente no es un final abrupto real.

30 Que el nivel de presión sonora total disminuya ligeramente es diferente a que el nivel de presión sonora total disminuya excesivamente rápido. La disminución lenta se refiere a que no se cumple ninguna de las condiciones anteriores q y r para determinar que la disminución es excesivamente rápida. Se ha de notar específicamente en la presente memoria que, en el procesamiento real, se establecen inicialmente diversas tramas iniciales para que disminuyan ligeramente y que, la determinación sólo comienza en una trama posterior a las diversas tramas anteriores. Debido a que cada trama dura sólo decenas de milisegundos en la aplicación real, se omiten los resultados de detección de las diversas tramas iniciales.

E46-2-4. Si se detecta, según la condición q o r, que uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, determinar si se cumple cualquiera de las siguientes condiciones s o t.

s) $(spl_tonal(k-1) \geq a_7)$,
 40 $(spl_tonal(k) < a_8)$,
 $(spl_tonal(k-1) - sp_no_tonal(k) > 0)$, y
 $(spl_no_tonal(k+1) < a_9)$, donde $i \geq 1$.

t) $(spl_tonal(k-2) \geq a_{10})$,
 $(spl_tonal(k-1) < a_{11})$,

45 $(spl_tonal(k-1) - sp_no_tonal(k-2) > 0)$, y
 $(spl_no_tonal(k) < a_{12})$, donde $i \geq 2$.

En esta realización, $a_6=25$, $a_7=47$, $a_{10}=50$, y $a_8=a_9=a_{11}=a_{12}=10$.

Si se cumple cualquiera de las condiciones s o t , la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz. Si no se cumple ninguna de las condiciones s o t , la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluido en la trama $k^{\text{ésima}}$ no es un final abrupto real.

Esta realización de la presente invención ofrece un método para detectar una señal de voz, donde se puede determinar una excepción abrupta real de una señal de voz al detectar primero una excepción abrupta potencial de una señal de voz, y analizar adicionalmente una característica de tono de la excepción abrupta potencial de una señal de voz, de manera que la precisión en la detección de una excepción abrupta de una señal de voz mejore de manera eficaz.

La Figura 7A es un diagrama de bloque esquemático de un aparato 70 para detectar una señal de voz según una realización de la presente invención. El aparato 70 incluye: una primera unidad de detección 71, una unidad de entramado 72, y una segunda unidad de detección 73.

La primera unidad de detección 71 está configurada para: realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía en cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, donde la excepción abrupta potencial de una señal de voz incluye una de las siguientes situaciones: una interrupción abrupta potencial, un comienzo abrupto, y un final abrupto de una señal de voz.

La unidad de detección 72 está configurada para: realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que incluye el primer período de tiempo meta es un segundo período de tiempo meta.

La segunda unidad de detección 73 está configurada para: procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, y determinar, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que incluye al menos uno de los segundos períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo incluida en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

Esta realización de la presente invención ofrece un aparato para detectar una señal de voz, donde se puede determinar una excepción abrupta real de una señal de voz, detectando primero una excepción abrupta potencial de una señal de voz, y analizando adicionalmente una característica de tono de la excepción abrupta potencial de una señal de voz, de manera que la precisión en la detección de una excepción abrupta de una señal de voz mejore de manera eficaz.

En otra realización, la Figura 7B es un diagrama de bloque esquemático de un aparato 70 para detectar una señal de voz según otra realización de la presente invención. A diferencia del aparato 70 de la Figura 7A, la primera unidad de detección 71 puede además incluir, específicamente: un primer módulo de adquisición 710 y un primer módulo de determinación 715; y la segunda unidad de detección 73 puede además incluir, específicamente: un segundo módulo de adquisición 730 y un segundo módulo de determinación 735.

El primer módulo de adquisición 710 está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $\text{trama_energía_corta}(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural.

De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(\text{trama_energía_corta}(i-1) - \text{trama_energía_corta}(i) \geq a_2)$ y $(\text{trama_energía_corta}(i) < a_1)$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(\text{trama_energía_corta}(i-2) - \text{trama_energía_corta}(i) \geq a_2)$ y $(\text{trama_energía_corta}(i) < a_1)$, donde a_1 y a_2 es un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-3)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

5
10 De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2)$ y $(trama_energía_corta(i-1) < a_1)$, determinar que la trama $i^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

15 De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2)$ y $(trama_energía_corta(i-2) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

20 De manera opcional, como una realización diferente, el primer módulo de determinación 715 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2)$ y $(trama_energía_corta(i-3) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas $(i-1)^{ésima}$ a la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

25 El segundo módulo de adquisición 730 está configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{ésima}$, donde la trama $k^{ésima}$ es el segundo período de tiempo $k^{ésimo}$ en los múltiples segundos períodos de tiempo y k es un número natural.

30 De manera opcional, como una realización diferente, el segundo módulo de determinación 735 está configurado para: si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz; o, si una característica de tono del segundo período de tiempo meta cumple con $(a_4 \leq spl_tonal(k) < a_3)$ y $(spl_total(k) \geq a_5)$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz, donde a_3 , a_4 y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente.

35 De manera opcional, como una realización diferente, el segundo módulo de determinación 735 está configurado para determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_total(k+1) \geq a_7),$$

$$(spl_total(k) < a_8),$$

$$(spl_total(k+1) - spl_no_tonal(k) > 0), \text{ y}$$

$$(spl_no_tonal(k-1) < a_9),$$

40 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un comienzo abrupto real de una señal de voz; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

45

$$(spl_total(k+2) \geq a_{10}),$$

50

$$(spl_total(k+1) < a_{11}),$$

55

$$(spl_total(k+2) - spl_no_tonal(k+1) > 0), \text{ y}$$

$$(spl_no_tonal(k) < a_{12}),$$

determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un comienzo abrupto real de una señal de voz, donde a_7 hasta a_{12} son un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k)-spl_total(k-1)) \geq a_6$ y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumenta ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k)-spl_total(k-2)) \geq a_6$, $(spl_total(k) > spl_total(k-1))$, $(spl_total(k-1) > spl_total(k-2))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente, y a_6 es un sexto umbral preestablecido; o si la característica de tono del segundo período de tiempo no cumple ninguna de las dos condiciones anteriores, determinar que $spl_tonal(k)$ aumenta ligeramente.

De manera opcional, como una realización diferente, el segundo módulo de determinación 735 está configurado para determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k-1) \geq a_7),$$

$$(spl_tonal(k) < a_8),$$

$$(spl_tonal(k-1) - spl_no_tonal(k) > 0), \text{ y}$$

$$(spl_no_tonal(k+1) < a_9),$$

determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz donde $k \geq 1$; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k-2) \geq a_{10}),$$

$$(spl_tonal(k-1) < a_{11}),$$

$$(spl_tonal(k-1) - spl_no_tonal(k-2) > 0), \text{ y}$$

$$(spl_no_tonal(k) < a_{12}),$$

determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz, donde $k \geq 2$, y a_7 hasta a_{12} son un séptimo umbral preestablecido a un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-1) - spl_total(k)) \geq a_6$ y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-2) - spl_total(k)) \geq a_6$, $(spl_total(k-1) > spl_total(k))$, $(spl_total(k-2) > spl_total(k-1))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si ninguna de las dos condiciones anteriores se cumplen, determinar que $spl_total(k)$ disminuye ligeramente, donde a_6 es un sexto umbral preestablecido.

El aparato 70 implementa los métodos 30 y 40. En aras de la brevedad, no se vuelven a brindar detalles específicos en la presente memoria.

La Figura 8 es un diagrama de bloque esquemático de un aparato 80 para detectar una señal de voz según otra realización de la presente invención. El aparato 80 incluye componentes, tal y como un procesador 81 y una memoria 82, donde los componentes se comunican entre sí mediante un bus.

El procesador 81 está configurado para ejecutar un programa de esta realización de la presente invención que se almacena en la memoria 82 y realizar una comunicación bi-direccional con otro aparato mediante el bus.

La memoria 82 puede incluir una memoria RAM y una ROM, o cualquier medio de almacenamiento fijo, o un medio de almacenamiento móvil, y se configura para almacenar un programa que pueda ejecutar esta realización de la presente invención, o datos que van a ser procesados en esta realización de la presente invención, o un resultado de detección para su posterior aplicación.

La memoria 82 y el procesador 81 pueden estar integrados en un módulo físico al que se aplica esta realización de la presente invención, y el programa que implementa esta realización de la presente invención se almacena y opera en el módulo físico.

En esta realización de la presente invención, el procesador 81 realiza, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detecta la energía de cada uno de los primeros períodos de tiempo, y determina un primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, donde la excepción abrupta potencial de una señal de voz incluye una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto, y final abrupto de una señal de voz; realiza, en una unidad de longitud trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que incluye el primer período de tiempo meta es un segundo período de tiempo meta; y procesa cada uno de los segundos períodos de tiempo para adquirir una característica de tono, y determina, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que incluye al menos uno de los segundos períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz incluida en el primer período de tiempo meta incluida en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

Después de determinar si la excepción abrupta potencial de una señal de voz es una excepción abrupta real de una señal de voz, el procesador puede enviar el resultado a la memoria para su almacenamiento, de tal manera que se realice otro procesamiento.

El procesador 81 puede realizar específicamente el entramado de una muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía $trama_energía_corta(i)$ de cada uno de los primeros períodos de tiempo, donde la trama $i^{ésima}$ es la trama $i^{ésima}$ en los múltiples primeros períodos de tiempo, e i es un número natural; y a continuación, mediante el análisis de relación entre la energía adquirida de los primeros períodos de tiempo y haciendo referencia a las condiciones a a f, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye una excepción abrupta potencial de una señal de voz.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-2)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i-3)-trama_energía_corta(i) \geq a_2)$ y $(trama_energía_corta(i) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas desde la $(i-1)^{ésima}$ a la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un final abrupto potencial, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un final abrupto potencial de una señal de voz.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2)$ y $(trama_energía_corta(i-1) < a_1)$, determinar que la trama $i^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2)$ y $(trama_energía_corta(i-2) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $(trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2)$ y $(trama_energía_corta(i-3) < a_1)$, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, determinar

que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un comienzo abrupto potencial de una señal de voz, donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

5 A continuación, el procesador 81 está configurado para: realizar un procesamiento de detección de tono en uno o más segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total ($spl_total(k)$), un nivel de presión sonora de componente tonal ($spl_tonal(k)$), y un nivel de presión sonora de componente no tonal ($spl_no_tonal(k)$) de la trama $k^{\text{ésima}}$, donde la trama $k^{\text{ésima}}$ es el segundo período de tiempo $k^{\text{ésimo}}$ en los múltiples segundos períodos de tiempo y k es un número natural. Finalmente, el procesador 81
10 determina, mediante el análisis de si la característica de tono del segundo período de tiempo meta cumple las condiciones g a t, si la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es una interrupción abrupta real de una señal de voz.

De manera opcional, como una realización diferente, el procesador 81 está configurado para: si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es una interrupción abrupta real de una señal de voz; o, si una
15 característica de tono del segundo período de tiempo meta cumple con ($a_4 \leq spl_tonal(k) < a_3$) y ($spl_total(k) \geq a_5$), determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es una interrupción abrupta real de una señal de voz, donde a_3 , a_4 y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente. De manera opcional, como una realización
20 diferente, el procesador 81 está configurado para: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$(spl_tonal(k+1) \geq a_7),$$

$$(spl_tonal(k) < a_8),$$

$$(spl_tonal(k+1) - spl_no_tonal(k) > 0), \text{ y}$$

$$25 (spl_no_tonal(k-1) < a_9),$$

determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$30 (spl_tonal(k+2) \geq a_{10}),$$

$$(spl_tonal(k+1) < a_{11}),$$

$$(spl_tonal(k+2) - spl_no_tonal(k+1) > 0), \text{ y}$$

$$(spl_no_tonal(k) < a_{12}),$$

35 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz, donde a_7 a a_{12} son un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con ($spl_total(k) - spl_total(k-1) \geq a_6$) y ($spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de
40 la 1^{era} trama aumenta ligeramente; o si la característica de tono del segundo período de tiempo cumple con ($spl_total(k) - spl_total(k-2) \geq a_6$), ($spl_total(k) > spl_total(k-1)$), ($spl_total(k-1) > spl_total(k-2)$), y ($spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente), determinar que $spl_tonal(k)$ aumenta excesivamente rápido, donde $k \geq 2$, se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente, y a_6 es un sexto umbral preestablecido; o si la característica de tono del segundo período de
45 tiempo no cumple ninguna de las dos condiciones anteriores, determinar que $spl_tonal(k)$ aumenta ligeramente.

De manera opcional, como una realización diferente, el procesador 81 está configurado para determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

$$50 (spl_tonal(k-1) \geq a_7),$$

$$(spl_tonal(k) < a_8),$$

$$(spl_tonal(k-1) - spl_no_tonal(k) > 0), \text{ y}$$

$$(spl_no_tonal(k+1) < a_9),$$

determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz, donde $k \geq 1$; o determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y la característica de tono del segundo período de tiempo cumple con:

- 5 $(spl_tonal(k-2) \geq a_{10})$,
 $(spl_tonal(k-1) < a_{11})$,
 $(spl_tonal(k-1) - spl_no_tonal(k-2) > 0)$, y
 $(spl_no_tonal(k) < a_{12})$,

- 10 determinar que la excepción abrupta potencial de una señal de voz incluida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz, donde $k \geq 2$, y a_7 a a_{12} son un séptimo umbral preestablecido a un duodécimo umbral preestablecido; y determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido incluye: si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-1) - spl_total(k) \geq a_6)$ y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si la característica de tono del segundo período de tiempo cumple con $(spl_total(k-2) - spl_total(k) \geq a_6)$, $(spl_total(k-1) > spl_total(k))$, $(spl_total(k-2) > spl_total(k-1))$, y $(spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente), determinar que $spl_total(k)$ disminuye excesivamente rápido, donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuye ligeramente; o si ninguna de las dos condiciones anteriores se cumple, determinar que $spl_total(k)$ disminuye ligeramente, donde a_6 es un sexto umbral preestablecido.
- 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100
 105
 110
 115
 120
 125
 130
 135
 140
 145
 150
 155
 160
 165
 170
 175
 180
 185
 190
 195
 200
 205
 210
 215
 220
 225
 230
 235
 240
 245
 250
 255
 260
 265
 270
 275
 280
 285
 290
 295
 300
 305
 310
 315
 320
 325
 330
 335
 340
 345
 350
 355
 360
 365
 370
 375
 380
 385
 390
 395
 400
 405
 410
 415
 420
 425
 430
 435
 440
 445
 450
 455
 460
 465
 470
 475
 480
 485
 490
 495
 500
 505
 510
 515
 520
 525
 530
 535
 540
 545
 550
 555
 560
 565
 570
 575
 580
 585
 590
 595
 600
 605
 610
 615
 620
 625
 630
 635
 640
 645
 650
 655
 660
 665
 670
 675
 680
 685
 690
 695
 700
 705
 710
 715
 720
 725
 730
 735
 740
 745
 750
 755
 760
 765
 770
 775
 780
 785
 790
 795
 800
 805
 810
 815
 820
 825
 830
 835
 840
 845
 850
 855
 860
 865
 870
 875
 880
 885
 890
 895
 900
 905
 910
 915
 920
 925
 930
 935
 940
 945
 950
 955
 960
 965
 970
 975
 980
 985
 990
 995
 1000
 1005
 1010
 1015
 1020
 1025
 1030
 1035
 1040
 1045
 1050
 1055
 1060
 1065
 1070
 1075
 1080
 1085
 1090
 1095
 1100
 1105
 1110
 1115
 1120
 1125
 1130
 1135
 1140
 1145
 1150
 1155
 1160
 1165
 1170
 1175
 1180
 1185
 1190
 1195
 1200
 1205
 1210
 1215
 1220
 1225
 1230
 1235
 1240
 1245
 1250
 1255
 1260
 1265
 1270
 1275
 1280
 1285
 1290
 1295
 1300
 1305
 1310
 1315
 1320
 1325
 1330
 1335
 1340
 1345
 1350
 1355
 1360
 1365
 1370
 1375
 1380
 1385
 1390
 1395
 1400
 1405
 1410
 1415
 1420
 1425
 1430
 1435
 1440
 1445
 1450
 1455
 1460
 1465
 1470
 1475
 1480
 1485
 1490
 1495
 1500
 1505
 1510
 1515
 1520
 1525
 1530
 1535
 1540
 1545
 1550
 1555
 1560
 1565
 1570
 1575
 1580
 1585
 1590
 1595
 1600
 1605
 1610
 1615
 1620
 1625
 1630
 1635
 1640
 1645
 1650
 1655
 1660
 1665
 1670
 1675
 1680
 1685
 1690
 1695
 1700
 1705
 1710
 1715
 1720
 1725
 1730
 1735
 1740
 1745
 1750
 1755
 1760
 1765
 1770
 1775
 1780
 1785
 1790
 1795
 1800
 1805
 1810
 1815
 1820
 1825
 1830
 1835
 1840
 1845
 1850
 1855
 1860
 1865
 1870
 1875
 1880
 1885
 1890
 1895
 1900
 1905
 1910
 1915
 1920
 1925
 1930
 1935
 1940
 1945
 1950
 1955
 1960
 1965
 1970
 1975
 1980
 1985
 1990
 1995
 2000
 2005
 2010
 2015
 2020
 2025
 2030
 2035
 2040
 2045
 2050
 2055
 2060
 2065
 2070
 2075
 2080
 2085
 2090
 2095
 2100
 2105
 2110
 2115
 2120
 2125
 2130
 2135
 2140
 2145
 2150
 2155
 2160
 2165
 2170
 2175
 2180
 2185
 2190
 2195
 2200
 2205
 2210
 2215
 2220
 2225
 2230
 2235
 2240
 2245
 2250
 2255
 2260
 2265
 2270
 2275
 2280
 2285
 2290
 2295
 2300
 2305
 2310
 2315
 2320
 2325
 2330
 2335
 2340
 2345
 2350
 2355
 2360
 2365
 2370
 2375
 2380
 2385
 2390
 2395
 2400
 2405
 2410
 2415
 2420
 2425
 2430
 2435
 2440
 2445
 2450
 2455
 2460
 2465
 2470
 2475
 2480
 2485
 2490
 2495
 2500
 2505
 2510
 2515
 2520
 2525
 2530
 2535
 2540
 2545
 2550
 2555
 2560
 2565
 2570
 2575
 2580
 2585
 2590
 2595
 2600
 2605
 2610
 2615
 2620
 2625
 2630
 2635
 2640
 2645
 2650
 2655
 2660
 2665
 2670
 2675
 2680
 2685
 2690
 2695
 2700
 2705
 2710
 2715
 2720
 2725
 2730
 2735
 2740
 2745
 2750
 2755
 2760
 2765
 2770
 2775
 2780
 2785
 2790
 2795
 2800
 2805
 2810
 2815
 2820
 2825
 2830
 2835
 2840
 2845
 2850
 2855
 2860
 2865
 2870
 2875
 2880
 2885
 2890
 2895
 2900
 2905
 2910
 2915
 2920
 2925
 2930
 2935
 2940
 2945
 2950
 2955
 2960
 2965
 2970
 2975
 2980
 2985
 2990
 2995
 3000
 3005
 3010
 3015
 3020
 3025
 3030
 3035
 3040
 3045
 3050
 3055
 3060
 3065
 3070
 3075
 3080
 3085
 3090
 3095
 3100
 3105
 3110
 3115
 3120
 3125
 3130
 3135
 3140
 3145
 3150
 3155
 3160
 3165
 3170
 3175
 3180
 3185
 3190
 3195
 3200
 3205
 3210
 3215
 3220
 3225
 3230
 3235
 3240
 3245
 3250
 3255
 3260
 3265
 3270
 3275
 3280
 3285
 3290
 3295
 3300
 3305
 3310
 3315
 3320
 3325
 3330
 3335
 3340
 3345
 3350
 3355
 3360
 3365
 3370
 3375
 3380
 3385
 3390
 3395
 3400
 3405
 3410
 3415
 3420
 3425
 3430
 3435
 3440
 3445
 3450
 3455
 3460
 3465
 3470
 3475
 3480
 3485
 3490
 3495
 3500
 3505
 3510
 3515
 3520
 3525
 3530
 3535
 3540
 3545
 3550
 3555
 3560
 3565
 3570
 3575
 3580
 3585
 3590
 3595
 3600
 3605
 3610
 3615
 3620
 3625
 3630
 3635
 3640
 3645
 3650
 3655
 3660
 3665
 3670
 3675
 3680
 3685
 3690
 3695
 3700
 3705
 3710
 3715
 3720
 3725
 3730
 3735
 3740
 3745
 3750
 3755
 3760
 3765
 3770
 3775
 3780
 3785
 3790
 3795
 3800
 3805
 3810
 3815
 3820
 3825
 3830
 3835
 3840
 3845
 3850
 3855
 3860
 3865
 3870
 3875
 3880
 3885
 3890
 3895
 3900
 3905
 3910
 3915
 3920
 3925
 3930
 3935
 3940
 3945
 3950
 3955
 3960
 3965
 3970
 3975
 3980
 3985
 3990
 3995
 4000
 4005
 4010
 4015
 4020
 4025
 4030
 4035
 4040
 4045
 4050
 4055
 4060
 4065
 4070
 4075
 4080
 4085
 4090
 4095
 4100
 4105
 4110
 4115
 4120
 4125
 4130
 4135
 4140
 4145
 4150
 4155
 4160
 4165
 4170
 4175
 4180
 4185
 4190
 4195
 4200
 4205
 4210
 4215
 4220
 4225
 4230
 4235
 4240
 4245
 4250
 4255
 4260
 4265
 4270
 4275
 4280
 4285
 4290
 4295
 4300
 4305
 4310
 4315
 4320
 4325
 4330
 4335
 4340
 4345
 4350
 4355
 4360
 4365
 4370
 4375
 4380
 4385
 4390
 4395
 4400
 4405
 4410
 4415
 4420
 4425
 4430
 4435
 4440
 4445
 4450
 4455
 4460
 4465
 4470
 4475
 4480
 4485
 4490 <

5 El producto de software está almacenado en un medio de almacenamiento, e incluye diversas instrucciones para indicar a un dispositivo informático (que puede ser un ordenador personal, un servidor o un dispositivo de red) que realice alguna o todas las etapas de los métodos descritos en las realizaciones de la presente invención. El medio de almacenamiento anterior incluye: cualquier medio que pueda almacenar un código de programa, tal y como un unidad flash USB, un disco duro removible, una memoria de solo lectura (ROM, memoria de sólo lectura), una memoria de acceso aleatorio (RAM, memoria de acceso aleatorio), un disco magnético, o un disco óptico.

Las descripciones anteriores son meramente formas de implementación específicas de la presente invención, pero no están concebidas para limitar el alcance de protección de la presente invención, el cual se define mediante las reivindicaciones adjuntas.

10

REIVINDICACIONES

1. Un método para detectar una señal de voz, que comprende:

realizar, en una unidad de longitud de trama de primer segmento de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía en cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, en donde la excepción abrupta potencial de una señal de voz comprende una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto y final abrupto de una señal de voz, y en donde una interrupción abrupta corresponde a una ocurrencia de un par que comprende un final abrupto y un comienzo abrupto en la misma sección de un segmento de la señal de voz;

realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, en donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que comprende el primer período de tiempo meta es un segundo período de tiempo meta; y

procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, en donde el procesamiento de característica de tono comprende realizar una transformada de Fourier rápida en cada uno de los segundos períodos de tiempo para adquirir un espectro de densidad de potencia, determinar un punto máximo local según el espectro de densidad de potencia, y analizar un segmento de un intervalo de dominio de frecuencia centrado en el punto máximo local para determinar si existe un componente tonal en una banda de frecuencia en la que se ubica el punto máximo local; y

determinar, mediante el análisis de la característica de tono adquirida de al menos uno de los segundos períodos de tiempo que comprende al menos uno de los primeros períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz comprendida en el primer período de tiempo meta comprendido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

2. El método según la reivindicación 1, en donde realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectando energía de cada uno de los primeros períodos de tiempo comprende:

realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico; y

adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde el período de tiempo $i^{\text{ésimo}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural.

3. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros períodos de tiempo comprende:

si la relación entre la energía de los primeros períodos de tiempo cumple con *trama_energía_corta(i-1)-trama_energía_corta(i)* $\geq a_2$ y *trama_energía_corta(i)* $< a_1$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

4. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros períodos de tiempo comprende:

si la relación entre la energía de los primeros períodos de tiempo cumple con *trama_energía_corta(i-2)-trama_energía_corta(i)* $\geq a_2$ y *trama_energía_corta(i)* $< a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, en donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no comprenden un final abrupto potencial de una señal de voz.

5. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros períodos de tiempo comprende:

si la relación entre la energía de los primeros periodos de tiempo cumple con $trama_energía_corta(i-3)-trama_energía_corta(i) \geq a_2$ y $trama_energía_corta(i) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que comprende un final abrupto potencial, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, en donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros periodos de tiempo que no comprenden un final abrupto potencial de una señal de voz.

- 5
6. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros periodos de tiempo comprende:

si la relación entre la energía de los primeros periodos de tiempo cumple con $trama_energía_corta(i)-trama_energía_corta(i-1) \geq a_2$ y $trama_energía_corta(i-1) < a_1$, determinar que la trama $i^{ésima}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

- 15
7. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros periodos de tiempo comprende:

si la relación entre la energía de los primeros periodos de tiempo cumple con $trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2$ y $trama_energía_corta(i-2) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{ésima}$ ni la trama $(i-2)^{ésima}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer a período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, en donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros periodos de tiempo que no comprenden un comienzo abrupto potencial de una señal de voz.

- 20
8. El método según la reivindicación 2, en donde la determinación de un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los primeros periodos de tiempo además comprende:

si la relación entre la energía de los primeros periodos de tiempo cumple con $trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2$ y $trama_energía_corta(i-3) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{ésima}$ y la trama $(i-3)^{ésima}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{ésima}$ es el primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, en donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros periodos de tiempo que no comprenden un comienzo abrupto potencial de una señal de voz.

- 30
9. El método según cualquiera de la reivindicación 1, en donde el procesamiento de cada uno de los segundos periodos de tiempo para adquirir una característica de tono comprende:

realizar un procesamiento de detección de tono en los múltiples segundos periodos de tiempo según un orden cronológico; y

adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{ésima}$ como características de tono de la trama $k^{ésima}$, en donde la trama $k^{ésima}$ es el $k^{ésimo}$ segundo período de tiempo en los múltiples segundos periodos de tiempo y k es un número natural.

- 35
10. El método según la reivindicación 9, en donde determinar, mediante el análisis de una característica de tono de al menos uno de los segundos periodos de tiempo que comprende al menos uno de los primeros periodos de tiempo meta, si la excepción abrupta potencial de una señal de voz comprendida en el primer período de tiempo comprendido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz comprende:

si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz; o

si una característica de tono del segundo período de tiempo meta cumple con $a_4 \leq spl_tonal(k) < a_3$ y $spl_total(k) \geq a_5$, determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{ésima}$ es una interrupción abrupta real de una señal de voz, en donde

a_3 , a_4 , y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente.

11. El método según la reivindicación 9, en donde determinar, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que comprende al menos uno de los primeros períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz comprendida en el primer período de tiempo meta comprendido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz comprende:

5 determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y

la característica de tono del segundo período de tiempo cumple con:

$$spl_tonal(k+1) \geq a_7,$$

$$10 \quad spl_tonal(k) < a_8,$$

$$spl_tonal(k+1) - spl_no_tonal(k) > 0, \text{ y}$$

$$spl_no_tonal(k-1) < a_9,$$

determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz; o

- 15 determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y

la característica de tono del segundo período de tiempo cumple con:

$$spl_tonal(k+2) \geq a_{10},$$

$$spl_tonal(k+1) < a_{11},$$

$$20 \quad spl_tonal(k+2) - spl_no_tonal(k+1) > 0, \text{ y}$$

$$spl_no_tonal(k) > a_{12},$$

determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz, en donde

a_7 a a_{12} son un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y

- 25 la determinación de si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido comprende:

30 si la característica de tono del segundo período de tiempo cumple con $spl_total(k) - spl_total(k-1) \geq a_6$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente, determinar que $spl_tonal(k)$ aumenta excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente; o

35 si la característica de tono del segundo período de tiempo cumple con $spl_total(k) - spl_total(k-2) \geq a_6$, $spl_total(k) > spl_total(k-1)$, $spl_total(k-1) - spl_total(k-2)$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente, determinar que $spl_tonal(k)$ aumenta excesivamente rápido, en donde $k \geq 2$, se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente, y a_6 es un sexto umbral preestablecido; o

si la característica de tono del segundo período de tiempo no cumple ninguna de las dos condiciones que determinan que $spl_tonal(k)$ aumenta ligeramente.

12. El método según la reivindicación 9, en donde determinar, mediante el análisis de una característica de tono de al menos uno de los segundos períodos de tiempo que comprende al menos uno de los primeros períodos de tiempo meta, si la excepción abrupta potencial de una señal de voz comprendida en el primer período de tiempo meta comprendido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz comprende:

40 determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye rápido, y

45 la característica de tono del segundo período de tiempo cumple con:

$$spl_tonal(k-1) \geq a_7,$$

$$spl_tonal(k) < a_8,$$

$spl_tonal(k-1)-sp_no_tonal(k)>0$, y

$spl_no_tonal(k+1)<a_9$,

determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz, en donde $k \geq 1$; o

5 determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y

la característica de tono del segundo período de tiempo cumple con:

$spl_tonal(k-2) \geq a_{10}$,

$spl_tonal(k-1) < a_{11}$,

10 $spl_total(k-1)-sp_no_tonal(k-2)>0$, y

$spl_no_tonal(k) < a_{12}$,

determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{ésima}$ es un final abrupto real de una señal de voz, en donde $k \geq 2$; y

a_7 a a_{12} son un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y

15 la determinación de si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido comprende:

20 si la característica de tono del segundo período de tiempo cumple con $spl_total(k-1)-spl_total(k) \geq a_6$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente, determinar que $spl_tonal(k)$ disminuye excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente; o

25 si la característica de tono del segundo período de tiempo cumple con $spl_total(k-2)-spl_no_tonal(k) \geq a_6$, $spl_total(k-1) > spl_total(k)$, $spl_total(k-2)-spl_total(k-1)$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente, determinar que $spl_tonal(k)$ disminuye excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente; o

si no se cumple ninguna de las dos condiciones anteriores que determinan que $spl_total(k)$ disminuye ligeramente, en donde

a_6 es un sexto umbral preestablecido.

13. Un método para detectar una señal de voz, que comprende:

30 una primera unidad de detección, configurada para: realizar, en una unidad de longitud de trama de primer período de tiempo, un entramado de una muestra de voz continua para obtener múltiples primeros períodos de tiempo, detectar energía en cada uno de los primeros períodos de tiempo, y determinar un primer período de tiempo meta que comprende una excepción abrupta potencial de una señal de voz mediante el análisis de una relación entre la energía de los múltiples primeros períodos de tiempo, en donde la excepción abrupta potencial de una señal de voz comprende una de las siguientes situaciones: interrupción abrupta potencial, comienzo abrupto, y final abrupto de una señal de voz, y en donde una interrupción abrupta corresponde a la ocurrencia de un par que comprende un final abrupto y un comienzo abrupto en la misma sección de un segmento de la señal de voz;

40 una unidad de entramado, configurada para realizar, en una unidad de longitud de trama de segundo período de tiempo, un entramado de la muestra de voz continua para obtener múltiples segundos períodos de tiempo, en donde una longitud de trama de cada uno de los segundos períodos de tiempo es una integral múltiple de la longitud de trama de primer período de tiempo, y un segundo período de tiempo que comprende el primer período de tiempo meta es un segundo período de tiempo meta; y

45 una segunda unidad de detección, configurada para: procesar cada uno de los segundos períodos de tiempo para adquirir una característica de tono, en donde la característica de tono comprende realizar una transformada de Fourier rápida en cada uno de los segundos períodos de tiempo para adquirir un espectro de densidad de potencia, determinar un punto máximo local según el espectro de densidad de potencia, y analizar un segmento de un intervalo de dominio de frecuencia centrado en el punto máximo local para determinar si existe un componente tonal en una banda de frecuencia en la que está ubicado el punto máximo local, en donde la segunda unidad de detección está además configurada para determinar, mediante el análisis de la característica de tono adquirida de al menos uno de los segundos períodos de tiempo que comprenden al menos uno de los primeros períodos de tiempo meta, si la excepción abrupta

potencial de una señal de voz comprendida en el primer período de tiempo meta comprendido en el segundo período de tiempo meta es una excepción abrupta real de una señal de voz.

14. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

5 un primer módulo de adquisición, configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo e i es un número natural; y

10 un primer módulo de determinación, configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $\text{trama_energía_corta}(i-1)-\text{trama_energía_corta}(i) \geq a_2$ y $\text{trama_energía_corta}(i) < a_1$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

15. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

15 un primer módulo de adquisición, en donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y

20 un primer módulo de determinación, en donde el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $\text{trama_energía_corta}(i-2)-\text{trama_energía_corta}(i) \geq a_2$ y $\text{trama_energía_corta}(i) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que incluye un final abrupto potencial de una señal de voz, en donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no comprenden un final abrupto potencial de una señal de voz.

16. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

30 un primer módulo de adquisición, en donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y

35 un primer módulo de determinación, en donde el primer módulo de determinación está configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $\text{trama_energía_corta}(i-3)-\text{trama_energía_corta}(i) \geq a_2$ y $\text{trama_energía_corta}(i) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{\text{ésima}}$ y la trama $(i-3)^{\text{ésima}}$ es un primer período de tiempo meta que comprende un final abrupto potencial, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que comprende un final abrupto potencial de una señal de voz, en donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no comprenden un final abrupto potencial de una señal de voz.

45 17. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

un primer módulo de adquisición, en donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y

50 un primer módulo de determinación, configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $\text{trama_energía_corta}(i)-\text{trama_energía_corta}(i-1) \geq a_2$ y $\text{trama_energía_corta}(i-1) < a_1$, determinar que la trama $i^{\text{ésima}}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, e $i \geq 1$.

18. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

un primer módulo de adquisición, en donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y

5

un primer módulo de determinación, configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $trama_energía_corta(i)-trama_energía_corta(i-2) \geq a_2$ y $trama_energía_corta(i-2) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ni la trama $(i-1)^{\text{ésima}}$ ni la trama $(i-2)^{\text{ésima}}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, en donde $i \geq 2$ y la trama 0 y la 1^{era} trama están preestablecidas como primeros períodos de tiempo que no comprenden un comienzo abrupto potencial de una señal de voz.

10

15

19. El aparato según la reivindicación 13, en donde la primera unidad de detección comprende:

un primer módulo de adquisición, en donde el primer módulo de adquisición está configurado para: realizar un entramado de la muestra de voz continua en una unidad de longitud de trama de primer período de tiempo, para dividir la muestra de voz continua en los múltiples primeros períodos de tiempo según un orden cronológico, y adquirir energía *trama_energía_corta(i)* de cada uno de los primeros períodos de tiempo, en donde la trama $i^{\text{ésima}}$ es el $i^{\text{ésimo}}$ primer período de tiempo en los múltiples primeros períodos de tiempo, e i es un número natural; y

20

un primer módulo de determinación, configurado para: si la relación entre la energía de los primeros períodos de tiempo cumple con $trama_energía_corta(i)-trama_energía_corta(i-3) \geq a_2$ y $trama_energía_corta(i) < a_1$, en donde a_1 y a_2 son un primer umbral preestablecido y un segundo umbral preestablecido, respectivamente, y ninguna de las tramas comprendidas entre la trama $(i-1)^{\text{ésima}}$ y la trama $(i-3)^{\text{ésima}}$ es un primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, determinar que la trama $i^{\text{ésima}}$ es el primer período de tiempo meta que comprende un comienzo abrupto potencial de una señal de voz, en donde $i \geq 3$ y la trama 0, la 1^{era} trama y la 2^{da} trama están preestablecidas como primeros períodos de tiempo que no incluyen un comienzo abrupto potencial de una señal de voz.

25

30

20. El aparato según cualquiera de las reivindicaciones 13 a 19, en donde la segunda unidad de detección comprende:

un segundo módulo de adquisición, configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico, y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{\text{ésima}}$, donde la trama $k^{\text{ésima}}$ es el segundo período de tiempo $k^{\text{ésimo}}$ en los múltiples segundos períodos de tiempo, y k es un número natural; y

35

un segundo módulo de determinación, configurado para: si una característica de tono del segundo período de tiempo meta cumple con $spl_tonal(k) \geq a_3$, determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es una interrupción abrupta real de una señal de voz; o

40

si una característica de tono del segundo período de tiempo cumple con $a_4 \leq spl_tonal(k) < a_3$ y $spl_total(k) \geq a_5$, determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es una interrupción abrupta real de una señal de voz, en donde

45

a_3 , a_4 , y a_5 son un tercer umbral preestablecido, un cuarto umbral preestablecido, y un quinto umbral preestablecido, respectivamente.

21. El aparato según cualquiera de las reivindicaciones 13 a 19, en donde la segunda unidad de detección comprende:

un segundo módulo de adquisición, configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{\text{ésima}}$, en donde la trama $k^{\text{ésima}}$ es el segundo período de tiempo $k^{\text{ésimo}}$ en los múltiples segundos períodos de tiempo y k es un número natural; y

50

un segundo módulo de determinación, configurado para: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y

55

la característica de tono del segundo período de tiempo cumple con:

$$spl_tonal(k+1) \geq a_7,$$

$$spl_tonal(k) < a_8,$$

$$spl_tonal(k+1) - spl_no_tonal(k) > 0, \text{ y}$$

$$spl_no_tonal(k-1) < a_9,$$

- 5 determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz; o

determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido, y

la característica de tono del segundo período de tiempo cumple con:

10 $spl_tonal(k+2) \geq a_{10},$

$$spl_tonal(k+1) < a_{11},$$

$$spl_tonal(k+2) - spl_no_tonal(k+1) > 0, \text{ y}$$

$$spl_no_tonal(k) < a_{12},$$

- 15 determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un comienzo abrupto real de una señal de voz, en donde

a_7 a a_{12} son un séptimo umbral preestablecido hasta un duodécimo umbral preestablecido; y

el segundo módulo de determinación está además configurado para determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ aumenta excesivamente rápido comprende:

- 20 si la característica de tono del segundo período de tiempo cumple con $spl_total(k) - spl_total(k-1) \geq a_6$ y con que $spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente, determinar que el $spl_tonal(k)$ aumenta excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente; o

- 25 si la característica de tono del segundo período de tiempo cumple con $spl_total(k) - spl_total(k-2) \geq a_6$, $spl_total(k) > spl_total(k-1)$, $spl_total(k-1) > spl_total(k-2)$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ aumentan ligeramente, determinar que $spl_tonal(k)$ aumenta excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama aumentan ligeramente, y a_6 es un sexto umbral preestablecido; o

si la característica de tono del segundo período de tiempo no cumple ninguna de las dos condiciones determinar que $spl_tonal(k)$ aumenta ligeramente.

- 30 22. El aparato según cualquiera de las reivindicaciones 13 a 19, en donde la segunda unidad de detección comprende: un segundo módulo de adquisición, configurado para: realizar un procesamiento de detección de tono en los múltiples segundos períodos de tiempo según un orden cronológico; y adquirir un nivel de presión sonora total $spl_total(k)$, un nivel de presión sonora de componente tonal $spl_tonal(k)$, y un nivel de presión sonora de componente no tonal $spl_no_tonal(k)$ de la trama $k^{\text{ésima}}$, en donde la trama $k^{\text{ésima}}$ es el segundo período de tiempo $k^{\text{ésimo}}$ en los múltiples segundos períodos de tiempo y k un número natural; y

- 35 un segundo módulo de determinación, configurado para: determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y

- 40 la característica de tono del segundo período de tiempo cumple con:

$$spl_tonal(k-1) \geq a_7,$$

$$spl_tonal(k) < a_8,$$

$$spl_tonal(k-1) - spl_no_tonal(k) > 0, \text{ y}$$

45 $spl_no_tonal(k+1) < a_9,$

determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz, donde $k \geq 1$; o

ES 2 610 102 T3

determinar si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y si uno de $spl_total(k)$, $spl_total(k-1)$, y $spl_total(k+1)$ disminuye excesivamente rápido, y

la característica de tono del segundo período de tiempo cumple con:

5 $spl_tonal(k-2) \geq a_{10}$,

$$spl_tonal(k-1) < a_{11},$$

$$spl_tonal(k-1) - spl_no_tonal(k-2) > 0, \text{ y}$$

$$spl_no_tonal(k) < a_{12},$$

10 determinar que la excepción abrupta potencial de una señal de voz comprendida en la trama $k^{\text{ésima}}$ es un final abrupto real de una señal de voz, en donde $k \geq 2$, y

a_7 a a_{12} son un séptimo umbral preestablecido a un duodécimo umbral preestablecido; y

15 la determinación de si uno de $spl_total(k)$, $spl_total(k-1)$, y el $spl_total(k+1)$ aumenta excesivamente rápido comprende:

si la característica de tono del segundo período de tiempo cumple con $spl_total(k-1) - spl_total(k) \geq a_6$ y con que $spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente, determinar que $spl_tonal(k)$ disminuye excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente; o

20 si la característica de tono del segundo período de tiempo cumple con $spl_total(k-2) - spl_total(k) \geq a_6$, $spl_total(k-1) - spl_total(k)$, $spl_total(k-1) - spl_total(k)$, y con que $spl_total(k-1)$ y $spl_total(k-2)$ disminuyen ligeramente, determinar que $spl_tonal(k)$ disminuye excesivamente rápido, en donde $k \geq 2$, y se preestablece que un nivel de presión sonora total de la trama 0 y un nivel de presión sonora total de la 1^{era} trama disminuyen ligeramente; o

25 si no se cumple ninguna de las dos condiciones anteriores que determinan que $spl_total(k)$ disminuye ligeramente, en donde

a_6 es un sexto umbral preestablecido.

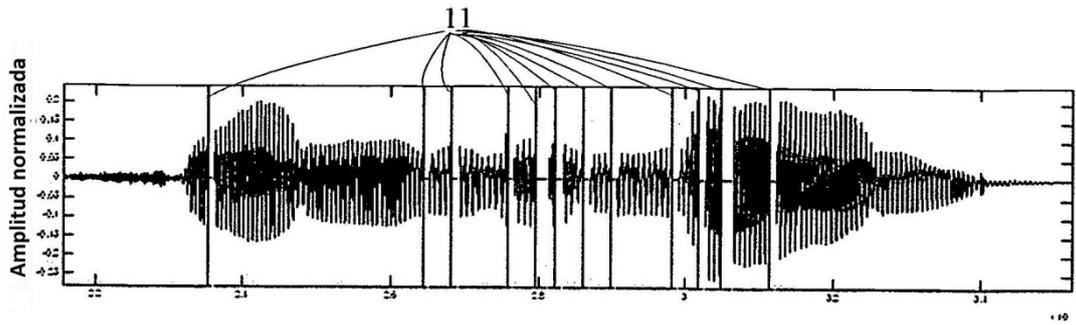


FIG. 1A

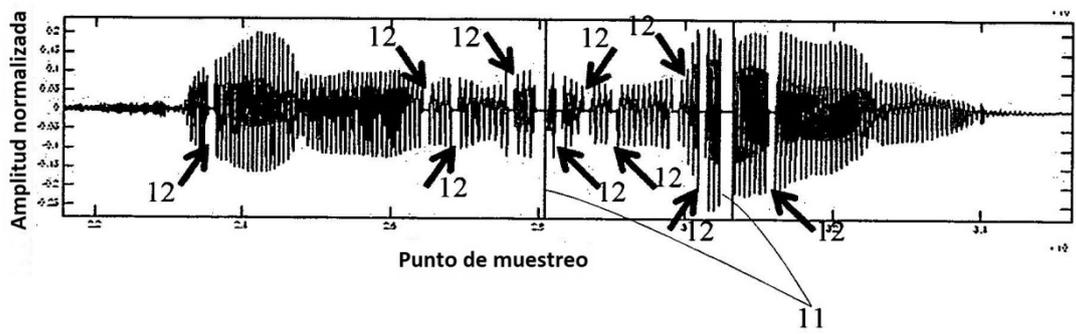


FIG. 1B

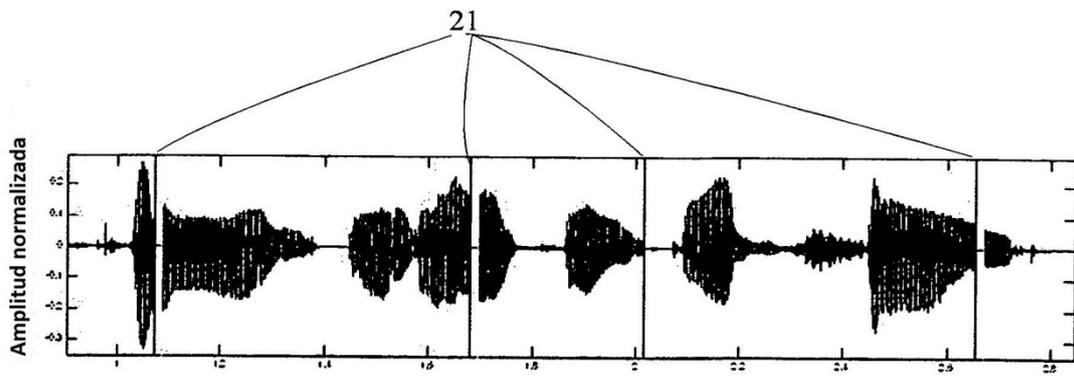


FIG. 2A

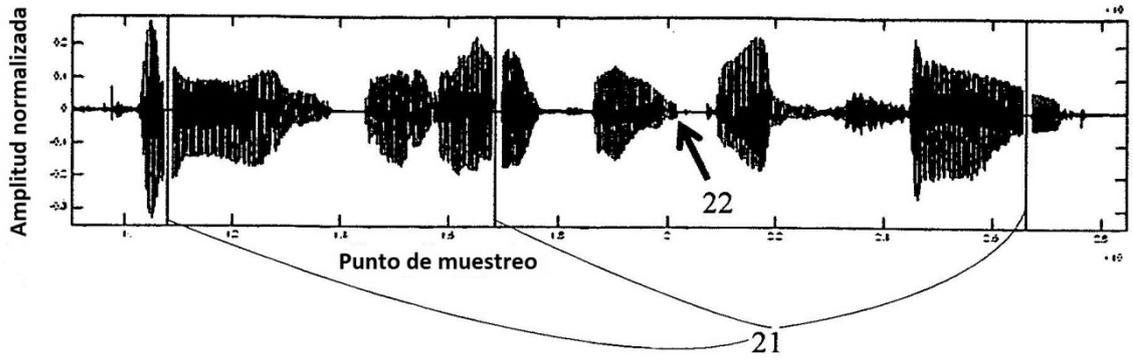


FIG. 2B

30

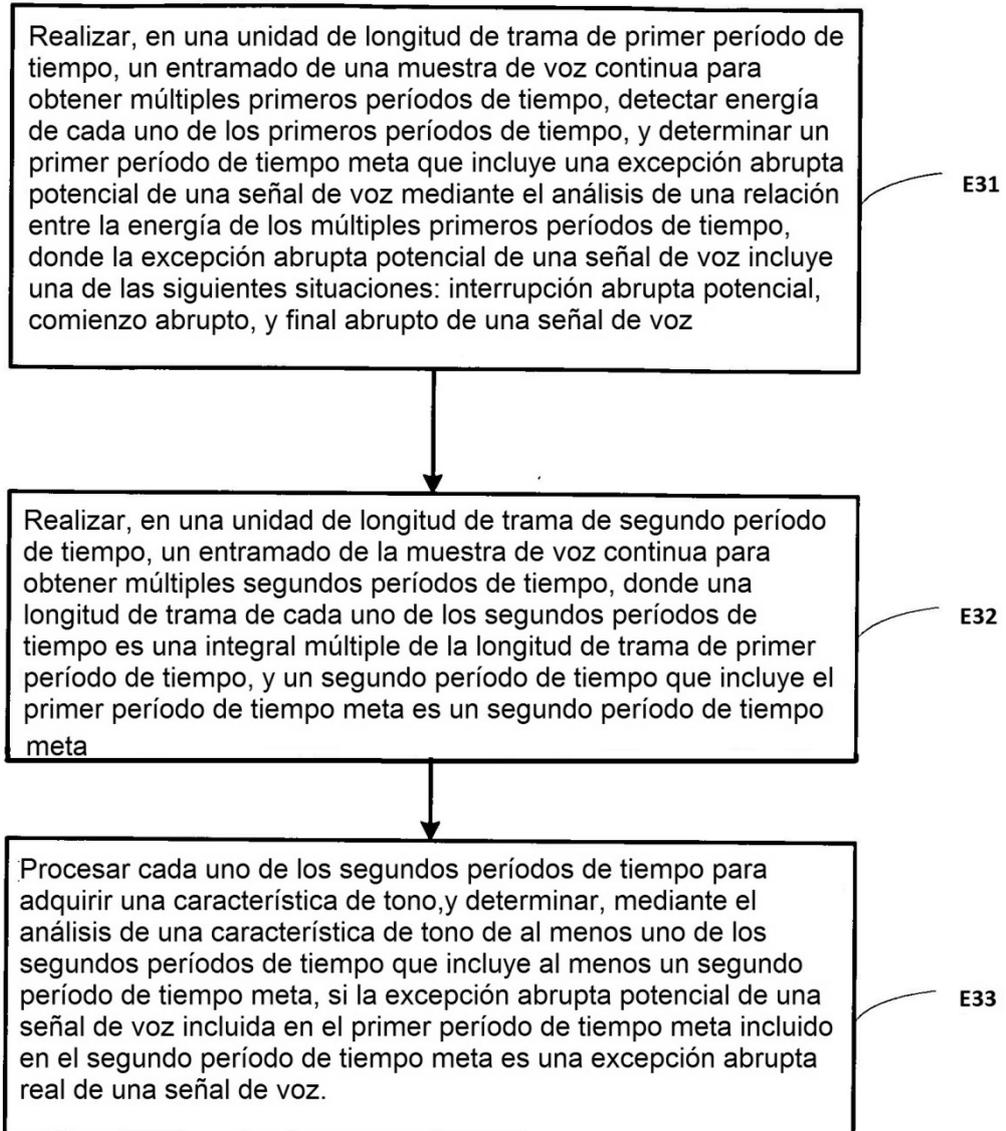


FIG. 3

40

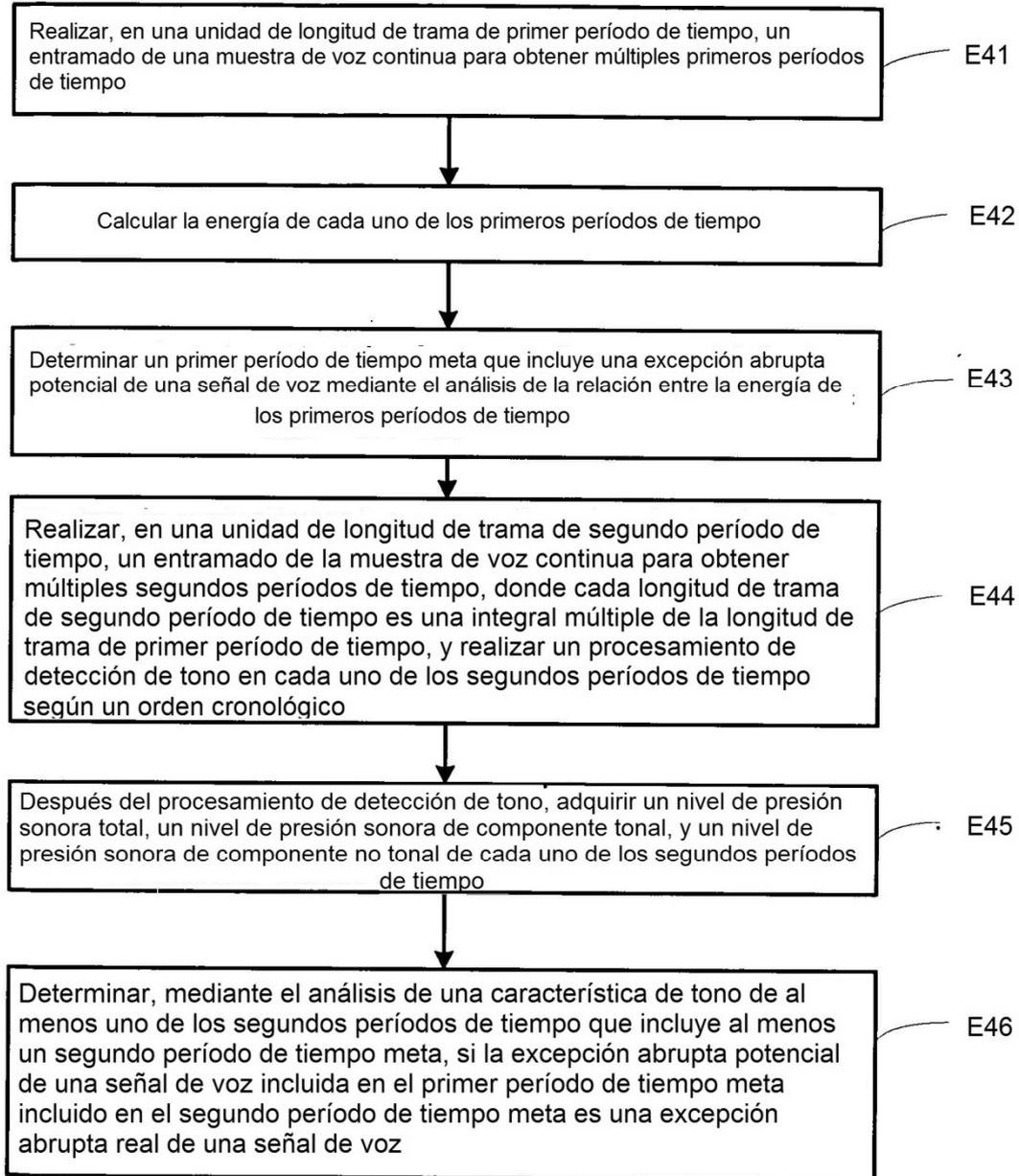


FIG. 4

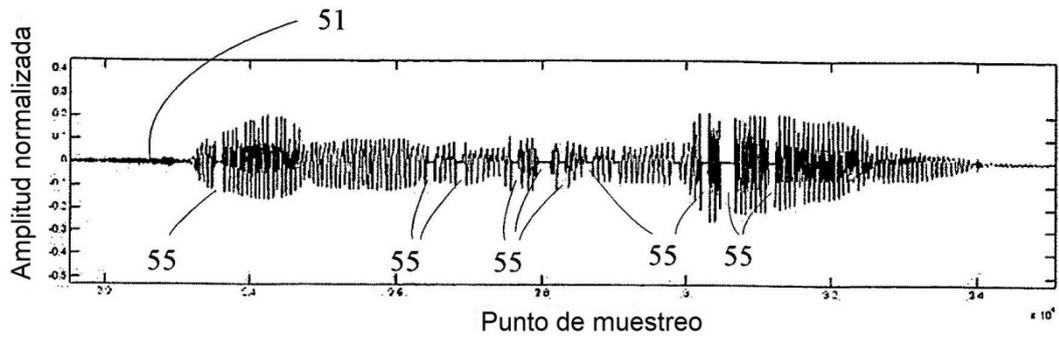


FIG. 5A

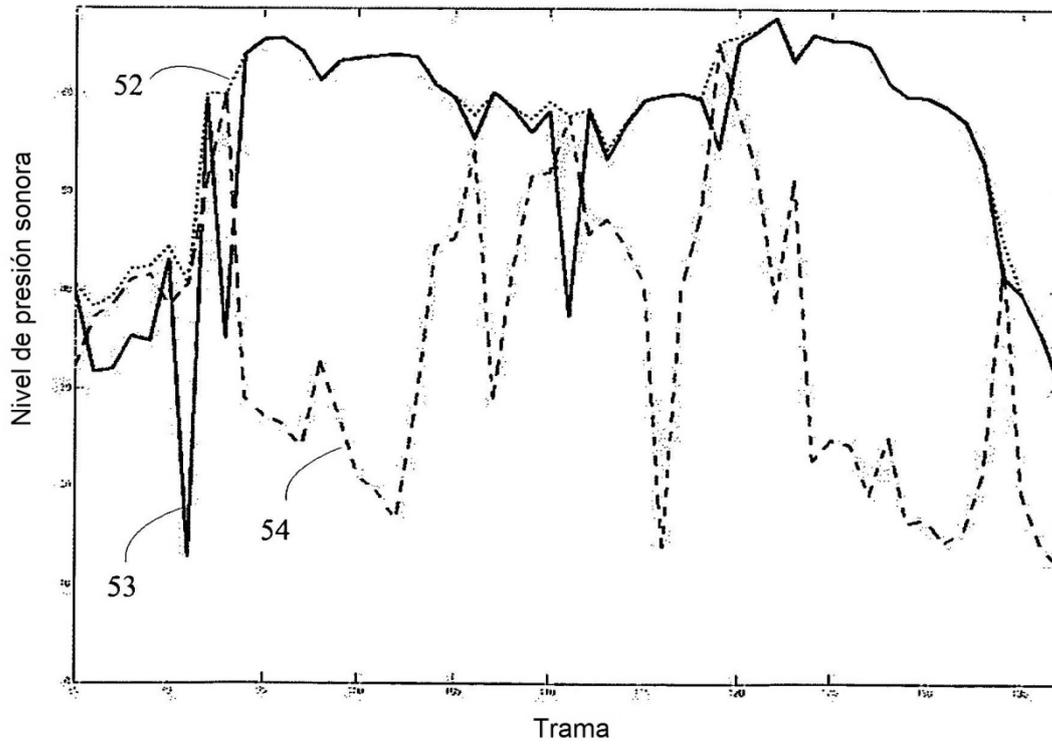


FIG. 5B

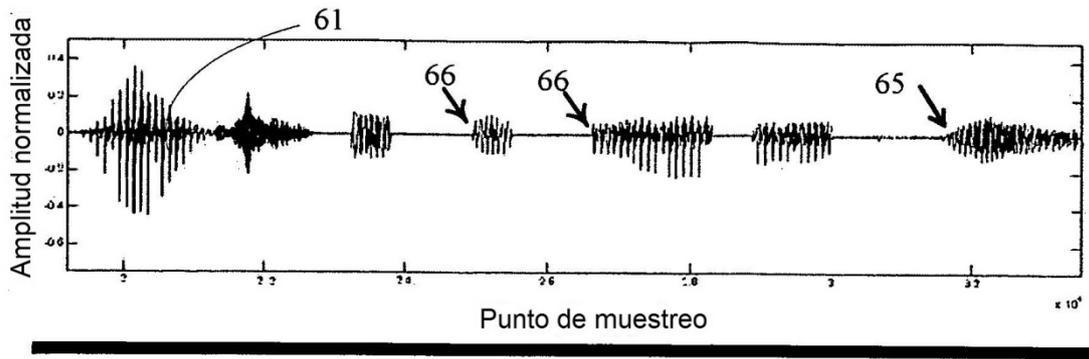


FIG. 6A

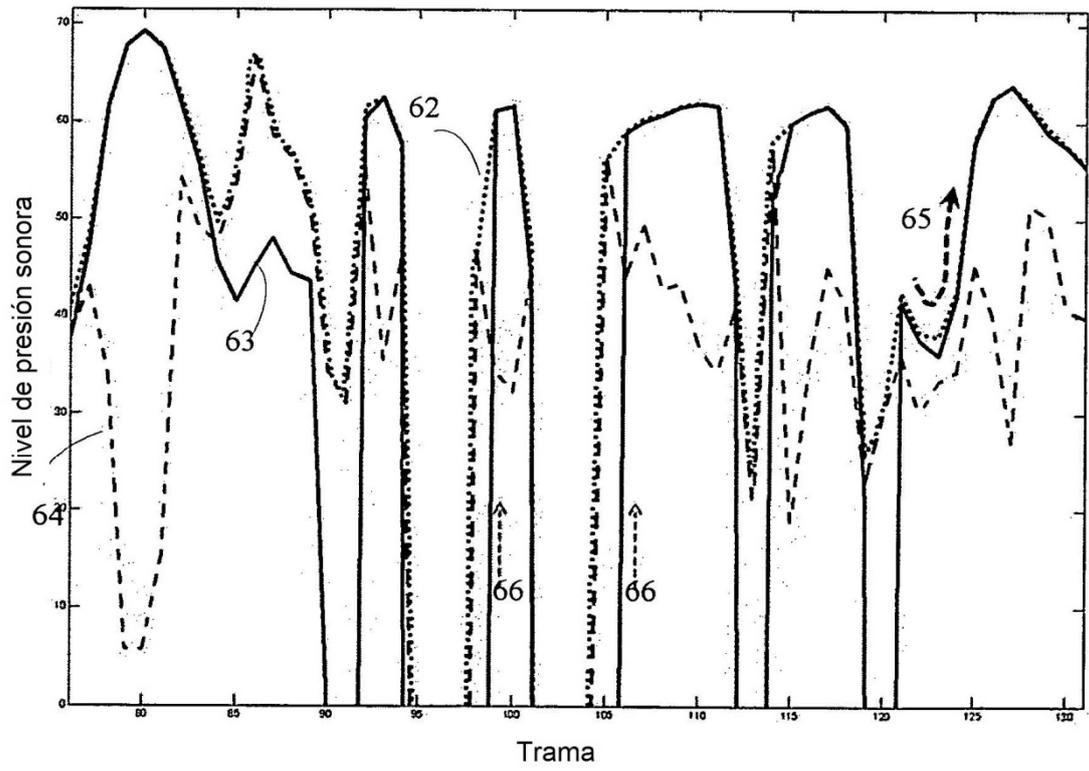


FIG. 6B

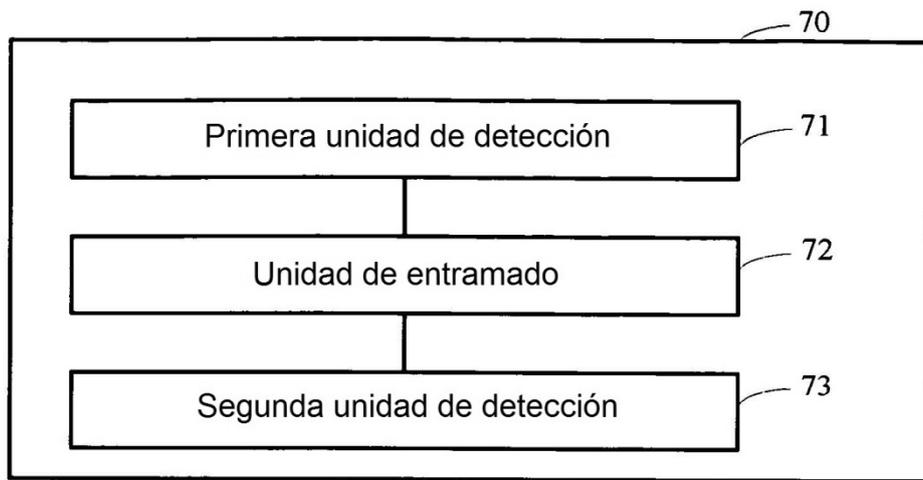


FIG. 7A

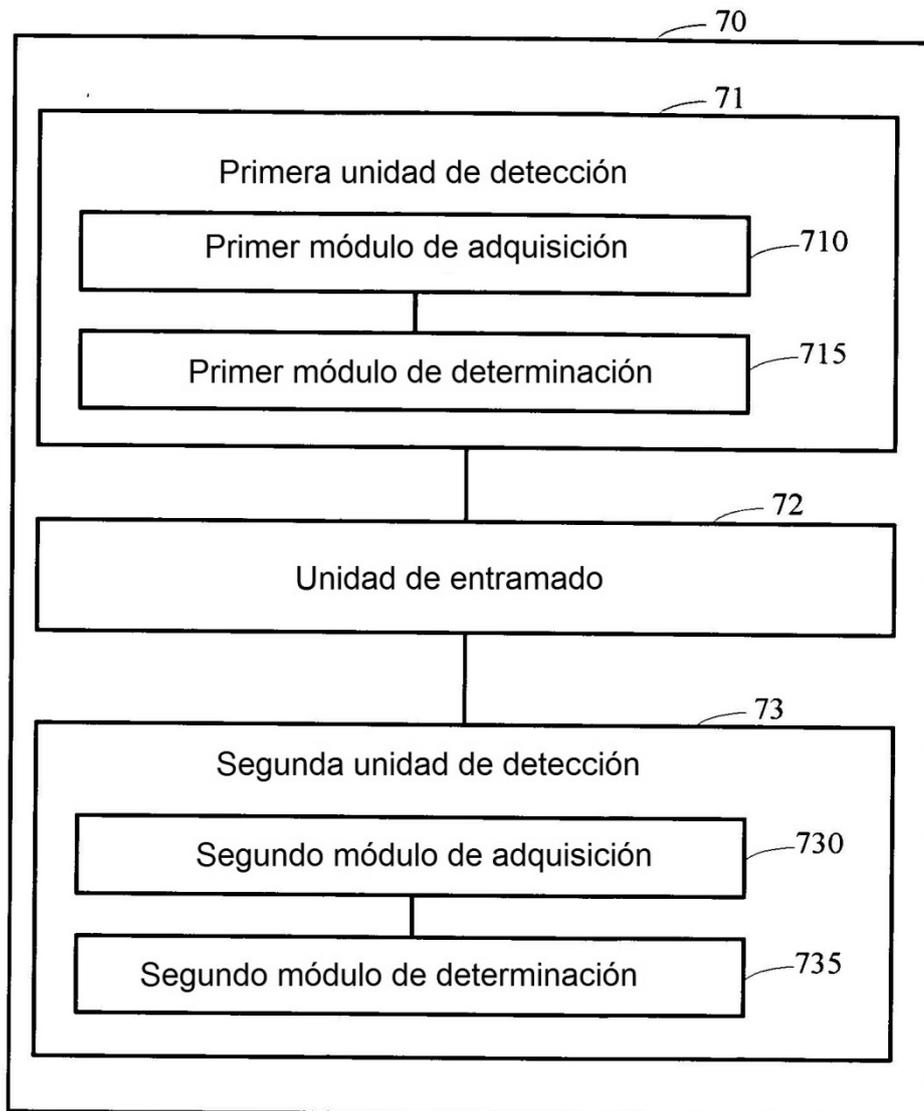


FIG. 7B

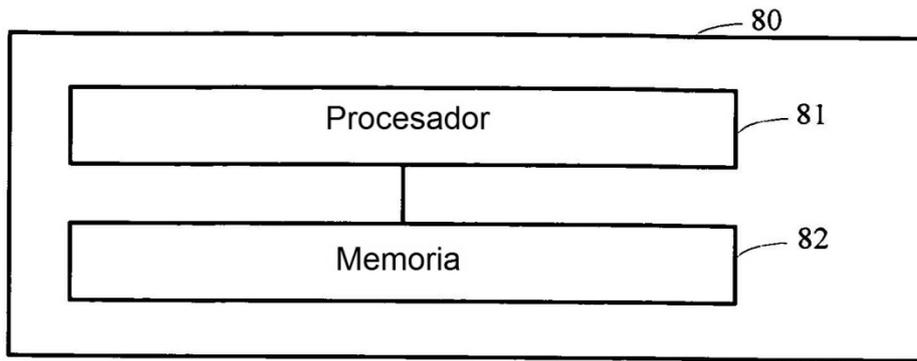


FIG. 8