

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 807 529**

51 Int. Cl.:

G10L 25/78 (2013.01)

G10L 25/21 (2013.01)

G10L 21/0232 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.10.2016 PCT/CN2016/101444**

87 Fecha y número de publicación internacional: **20.04.2017 WO17063516**

96 Fecha de presentación y número de la solicitud europea: **08.10.2016 E 16854895 (6)**

97 Fecha y número de publicación de la concesión europea: **10.06.2020 EP 3364413**

54 Título: **Método para la determinación de señal de ruido y aparato del mismo**

30 Prioridad:

13.10.2015 CN 201510670697

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

23.02.2021

73 Titular/es:

**ADVANCED NEW TECHNOLOGIES CO., LTD.
(100.0%)**

**Cayman Corporate Centre, 27 Hospital Road
George Town, Grand Cayman KY1-9008, KY**

72 Inventor/es:

DU, ZHIJUN

74 Agente/Representante:

VIDAL GONZÁLEZ, Maria Ester

ES 2 807 529 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 para la determinación de señal de ruido y aparato del mismo

5 La presente solicitud reivindica prioridad a la solicitud de patente china No. 201510670697.8, presentada el 13 de octubre de 2015 y titulada " NOISE SIGNAL DETERMINING METHOD AND APPARATUS AND VOICE DENOISING METHOD AND APPARATUS".

Campo técnico

10 La presente solicitud se refiere al campo de las tecnologías de eliminación de ruido de voz, y en particular, a un método y aparato de determinación de señal de ruido y a un método y aparato de eliminación de ruido de voz.

Antecedentes de la Técnica

15 Una tecnología de eliminación de ruido de voz puede mejorar la calidad de la voz al eliminar los ruidos ambientales de una señal de voz. Un espectro de potencia de una señal de ruido en una señal de voz debe determinarse primero en el proceso de eliminación de ruido de voz, y luego la señal de voz puede eliminarse de acuerdo con el espectro de potencia determinado de la señal de ruido.

20 En la técnica anterior, un espectro de potencia de una señal de ruido en una señal de voz generalmente se puede determinar de la siguiente manera: analizando las primeras N señales de trama en un segmento de señal de voz suponiendo que las primeras N señales de trama son señales de ruido (es decir, que no incluyen ninguna señal de voz humana), para obtener los espectros de potencia de las señales de ruido en la señal de voz.

25 En un escenario de aplicación real, las primeras N señales de trama en una señal de voz que se supone que son señales de ruido en la técnica anterior usualmente son inconsistentes con las señales de ruido reales, y por lo tanto la precisión de los espectros de potencia de la señal de ruido obtenida se ve afectada.

30 El documento CN 101968957B describe un método de detección de voz en condiciones de ruido, que incluye las etapas de: convertir señales de entrada a un dominio de frecuencia y dividir las en subbandas; calcular un espectro de potencia de cada subbanda para formar un espectro de potencia de subbanda; calcular la desviación cuadrática media del espectro de potencia de subbanda de cada trama y comparar la desviación cuadrática media que sirve como característica de detección con un umbral de detección de voz adaptativo para determinar si la trama actual contiene señales de voz.

35 El documento EP 2546831 describe un dispositivo de supresión de ruido en el que una unidad de separación de banda lleva a cabo una división de banda de una pluralidad de espectros de potencia dentro de la que se convierte una señal de entrada por una unidad de conversión de tiempo a frecuencia para combinar los espectros de potencia en cada subbanda. Una unidad generadora de componentes representativos de banda define un espectro de potencia que tiene un máximo entre la pluralidad de espectros de potencia dentro de cada subbanda como un espectro de potencia representativo. Una unidad generadora de cantidad de supresión de ruido calcula una cantidad de supresión de ruido para cada subbanda utilizando el espectro de potencia representativo y un espectro de ruido, y una unidad de supresión de ruido suprime las amplitudes de los espectros de potencia de acuerdo con la cantidad de supresión de ruido.

Resumen de la invención

50 Los objetivos de las realizaciones de la presente solicitud son proporcionar un método de determinación de señal de ruido y un aparato para resolver el problema en la técnica anterior de que la precisión de los espectros de potencia de señal de ruido obtenidos se ve afectada ya que las primeras N señales de trama que se supone que son señales de ruido son inconsistentes con las reales señales de ruido.

55 Para resolver el problema técnico de más arriba, se proporcionan el método de determinación de la señal de ruido según se establece en la reivindicación 1 y el aparato para determinar las señales de ruido según se establece en la reivindicación 6. Realizaciones particulares se definen en las reivindicaciones dependientes. Todos los sucesos siguientes de las palabras "realización(es)", si se refieren a combinaciones de características diferentes de las definidas por las reivindicaciones independientes, se refieren a ejemplos que se presentaron originalmente, pero que no representan realizaciones de la invención reivindicada actualmente; estos ejemplos todavía se muestran solo con fines ilustrativos.

60 Como se puede ver en las soluciones técnicas de más arriba proporcionadas en las realizaciones de la presente solicitud, realizando la transformada de Fourier en un segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama, determinando una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a analizar en varias frecuencias, y finalmente determinando si la señal de trama es una señal de ruido en base a la varianza, el método de determinación de señal de ruido y el aparato proporcionados

en las realizaciones de la presente solicitud pueden obtener con precisión varias tramas de ruido incluidas en el segmento de señal de voz a analizar. Se puede eliminar ruido de la voz a procesar en función de una potencia promedio de las tramas de ruido determinadas en el proceso de eliminación de ruido de voz, y por lo tanto se mejora el efecto de la eliminación de ruido de voz.

5 Breve Descripción de los Dibujos

Para describir las soluciones técnicas en las realizaciones de la presente solicitud o la técnica anterior más claramente, a continuación, se presentan brevemente los dibujos adjuntos utilizados para describir las realizaciones o la técnica anterior. Aparentemente, los dibujos adjuntos que se describen a continuación son meramente algunas realizaciones mencionadas en la presente solicitud, y los expertos en la técnica aún pueden deducir, sin esfuerzos creativos, otros dibujos de estos dibujos adjuntos.

10 La Figura 1 es un diagrama de flujo de un método de determinación de señal de ruido de acuerdo con una realización de la presente solicitud;

15 La Figura 2 es un diagrama de flujo de etapas para determinar si una señal de trama es una señal de ruido de acuerdo con una realización de la presente solicitud;

La Figura 3 es un diagrama de flujo de etapas para determinar una varianza de los valores de potencia de una señal de trama en varios puntos de muestreo de acuerdo con una realización de la presente solicitud;

20 La Figura 4 es un gráfico de curva de las varianzas de los valores de potencia de acuerdo con una realización de la presente solicitud;

La Figura 5 es un diagrama de flujo de un método de eliminación de ruido de voz de acuerdo con una realización de la presente solicitud;

La Figura 6 es un diagrama de bloques de un aparato de determinación de señal de ruido de acuerdo con una realización de la presente solicitud;

25 La Figura 7 es un diagrama de bloques de un aparato de eliminación de ruido de voz de acuerdo con una realización de la presente solicitud; y

La Figura 8 es un diagrama estructural esquemático de un ejemplo de implementación de hardware de un aparato proporcionado en la presente solicitud.

30 Descripción Detallada

Para que los expertos en la técnica entiendan mejor las soluciones técnicas de la presente solicitud, las soluciones técnicas en las realizaciones de la presente solicitud se describirán clara y completamente a continuación con referencia a los dibujos adjuntos en las realizaciones de la presente solicitud.

35 La Figura 1 muestra un diagrama de flujo de un método de determinación de señal de ruido de acuerdo con una realización de la presente solicitud. Para determinar una señal de ruido en un segmento de señal de voz a analizar, el método de determinación de la señal de ruido de esta realización incluye las siguientes etapas:

40 S101: Se realiza la transformada de Fourier en cada señal de trama en el segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama en el segmento de señal de voz.

El segmento de señal de voz a analizar puede capturarse a partir de una voz a procesar en base a una determinada regla. El segmento de señal de voz a analizar puede ser un "segmento de trama de ruido sospechoso" que posiblemente incluye muchas tramas de ruido en base a una determinación preliminar. Preferentemente, antes de la etapa S101, el método incluye, además:

45 determinar un segmento de señal de voz con una variación de amplitud menor que un umbral preestablecido en la voz a procesar como el segmento de señal de voz a analizar en base a una variación de amplitud de una señal de dominio de tiempo de la voz a procesar; o capturar las primeras N señales de voz de trama en la voz a procesar como el segmento de señal de voz a analizar.

50 En la realización de la presente solicitud, en un dominio de tiempo de una señal de voz, una señal de ruido es generalmente un segmento de señal de voz que tiene una pequeña variación de amplitud o que tiene amplitudes consistentes, mientras que un segmento de señal de voz que incluye una voz de voz humana generalmente fluctúa mucho en variación de amplitud. En base a dicha regla, se puede establecer de antemano un umbral preestablecido utilizado para reconocer un "segmento de trama de ruido sospechoso" incluido en una voz a procesar (es decir, una voz a cuyo ruido será eliminado). Por lo tanto, un segmento de señal de voz que tiene una variación de amplitud menor que el umbral preestablecido en la voz a procesar puede determinarse como el segmento de señal de voz a analizar.

60 En la realización de la presente solicitud, el encuadre puede realizarse primero en una señal de voz. Una señal de trama se refiere a una señal de voz de trama única, y un segmento de señal de voz puede incluir varias señales de trama. Una señal de trama puede incluir varios puntos de muestreo, por ejemplo, 1.024 puntos de muestreo. Dos señales de trama adyacentes pueden superponerse entre sí (por ejemplo, una relación de superposición puede ser del 50%). En esta realización, se puede realizar una transformada de Fourier a corto plazo (STFT) en una señal de voz en un dominio de tiempo para adquirir un espectro de potencia (dominio de frecuencia) de la señal de voz. El espectro de potencia puede incluir múltiples valores de potencia correspondientes a diferentes frecuencias, por

ejemplo, 1.024 valores de potencia.

En la realización de la presente solicitud, generalmente se puede suponer de forma predeterminada que una señal de voz en un período de tiempo (por ejemplo, 1,5 s) antes de que una persona hable es una señal de ruido (un ruido ambiental) en un segmento de señal de voz que incluye una voz humana. Por lo tanto, se puede determinar en la realización de la presente solicitud que la señal de voz a analizar son las primeras N señales de trama en un segmento de señal de voz. Por ejemplo, la señal de voz a analizar es una señal de voz en los primeros 1,5 s: $\{f_1', f_2', \dots, f_n'\}$, en el que f_1', f_2', \dots, f_n' representan señales de trama incluidas en la señal de voz, respectivamente. La realización de la presente solicitud tiene como objetivo determinar las señales de ruido a partir de las señales de trama en la señal de voz analizada.

Se pueden calcular múltiples valores de potencia correspondientes a cada señal de trama en función del espectro de potencia de la señal de voz a analizar: $\{f_1', f_2', \dots, f_n'\}$ obtenida después del STFT. Suponga que el espectro de potencia de una señal de trama a una frecuencia es $a+bi$, en el que la parte real a puede representar la amplitud y la parte imaginaria b puede representar la fase. Entonces, un valor de potencia de la señal de trama a la frecuencia puede ser: a^2+b^2 . Los valores de potencia de cada señal de trama a diferentes frecuencias se pueden obtener en base al proceso de más arriba. Por ejemplo, si cada una de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ incluye 1.024 puntos de muestreo, se pueden obtener 1.024 valores de potencia de cada señal de trama a diferentes frecuencias en función del espectro de potencia. Por ejemplo, los valores de potencia correspondientes a la señal de trama f_1' son $\{p^1_1, p^1_2, \dots, p^1_{1024}\}$, los valores de potencia correspondientes a la señal de trama f_2' son $\{p^2_1, p^2_2, \dots, p^2_{1024}\}, \dots$, y los valores de potencia correspondientes a la señal de trama f_n' son $\{p^n_1, p^n_2, \dots, p^n_{1024}\}$.

S102: Una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a varias frecuencias se determina en función del espectro de potencia de la señal de trama.

En base a los valores de potencia de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ en varias frecuencias, varianzas $\{Var(f_1'), Var(f_2'), \dots, Var(f_n')\}$ de los valores de potencia de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ se pueden calcular de acuerdo con una fórmula de cálculo de la varianza. Por ejemplo, si cada señal de trama incluye 1024 puntos de muestreo, $Var(f_1')$ es una varianza de $\{p^1_1, p^1_2, \dots, p^1_{1024}\}$, $Var(f_2')$ es una varianza de $\{p^2_1, p^2_2, \dots, p^2_{1024}\}, \dots$, y $Var(f_n')$ es una varianza de $\{p^n_1, p^n_2, \dots, p^n_{1024}\}$.

S103: Se determina si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza.

En la realización de la presente solicitud, la energía (es decir, un valor de potencia) de una señal de trama que incluye un segmento de voz generalmente varía mucho con las bandas, mientras que la energía de una señal de trama sin un segmento de voz (es decir, una señal de ruido) varía con las bandas ligeramente y se distribuye uniformemente. Por lo tanto, se puede determinar si cada señal de trama es una señal de ruido en base a una varianza de los valores de potencia de la señal de trama.

La Figura 2 muestra un diagrama de flujo de etapas para determinar si una señal de trama es una señal de ruido de acuerdo con una realización de la presente solicitud. En la realización de la presente solicitud, la etapa de más arriba S103 puede incluir las siguientes etapas:

S1031: Se determina si la varianza de los valores de potencia de la señal de trama es mayor que un primer umbral T_1 .

S1032: Si no, la señal de trama se determina como una señal de ruido.

Si una varianza de los valores de potencia de una señal de trama excede el primer umbral T_1 , se indica que una amplitud de variación de la energía (es decir, valores de potencia) de la señal de trama con bandas excede el primer umbral T_1 . Por lo tanto, se puede determinar que la señal de trama no es una señal de ruido. En contraste, si una varianza de los valores de potencia de una señal de trama no excede el primer umbral T_1 , se indica que una amplitud de variación de la energía (es decir, valores de potencia) de la señal de trama con bandas no excede el primer umbral T_1 . Por lo tanto, se puede determinar que la señal de trama es una señal de ruido.

En base al proceso de más arriba, las señales de trama de ruido $\{f_1', f_2', \dots, f_m'\}$ y las señales de trama sin ruido $\{f_{m+1}', f_{m+2}', \dots, f_n'\}$ se puede determinar secuencialmente en las señales de voz a analizar $\{f_1', f_2', \dots, f_n'\}$. Por lo tanto, las señales de ruido incluidas en un segmento de señal de voz pueden determinarse, y la eliminación de ruido de voz puede realizarse de acuerdo con estas señales de ruido $\{f_1', f_2', \dots, f_n'\}$.

Con referencia a la Figura 3, en la realización de la presente solicitud, la etapa de más arriba S102 puede incluir específicamente las siguientes etapas:

S1021: Los valores de potencia de cada una de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ en varias frecuencias se clasifican al menos en un primer conjunto de valores de potencia correspondiente a un primer intervalo de frecuencia y un segundo conjunto de valores de potencia correspondiente a un segundo intervalo de frecuencia de acuerdo con los intervalos de frecuencia a los que pertenecen las frecuencias correspondientes al espectro de potencia de la señal de trama, siendo el primer intervalo de frecuencia inferior al segundo intervalo de frecuencia.

En una realización específica, se puede adquirir una varianza de cada señal de trama en el dominio de frecuencia a través de estadísticas. Las señales sin ruido generalmente se concentran en las bandas de frecuencia media baja, mientras que las señales de ruido generalmente se distribuyen uniformemente en todas las bandas de frecuencia. Por lo tanto, se puede adquirir una varianza de los valores de potencia de cada señal de trama a varias frecuencias a través de estadísticas en al menos dos bandas de frecuencia diferentes (es decir, los intervalos de frecuencia de más arriba).

Por ejemplo, el primer intervalo de frecuencia puede ser 0~2000 Hz (banda de baja frecuencia), y el segundo intervalo de frecuencia puede ser 2000~4000 Hz (banda de alta frecuencia). Si cada señal de trama incluye 1024 puntos de muestreo, 1024 valores de potencia correspondientes a cada señal de trama se clasifican en un primer conjunto de valores de potencia A correspondiente a 0~2000 Hz y un segundo conjunto de valores de potencia B correspondiente a 2000~4000 Hz de acuerdo con los intervalos de frecuencia correspondientes a los valores de potencia. Usando la señal de trama f_1' como ejemplo, 1.024 valores de potencia correspondientes son $\{p^1_1, p^1_2, \dots, p^1_{1024}\}$. De acuerdo con los intervalos de frecuencia, se puede deducir que los valores de potencia incluidos en el primer conjunto de valores de potencia A son, por ejemplo, $\{p^1_1, p^1_2, \dots, p^1_{126}\}$, valores de potencia incluidos en el primer el conjunto de potencias A son, por ejemplo, $\{p^1_{127}, p^1_{128}, \dots, p^1_{1024}\}$, y el resto puede deducirse por analogía.

Debe observarse que las varianzas de los valores de potencia de la señal pueden adquirirse mediante estadísticas en más de dos bandas de frecuencia en otras realizaciones de la presente solicitud.

S1022: Se determina una primera varianza de los valores de potencia incluidos en el primer conjunto de valores de potencia.

Como se describió más arriba, utilizando la señal de trama f_1' como ejemplo, los valores de potencia incluidos en el primer conjunto de valores de potencia A son, por ejemplo, $\{p^1_{127}, p^1_{128}, \dots, p^1_{1024}\}$. Por lo tanto, se puede calcular una primera variación $Var_{alta}(f_1')$ de los valores de potencia $p^1_{127} \sim p^1_{1024}$ de acuerdo con una fórmula de varianza.

S1021: Se determina una segunda varianza de los valores de potencia incluidos en el segundo conjunto de valores de potencia.

Como se describió más arriba, utilizando la señal de trama f_1' como ejemplo, los valores de potencia incluidos en el segundo conjunto de valores de potencia B son, por ejemplo, $\{p^1_1, p^1_2, \dots, p^1_{126}\}$. Por lo tanto, se puede calcular una segunda variación $Var_{baja}(f_1')$ de los valores de potencia $p^1_1 \sim p^1_{126}$ de acuerdo con una fórmula de varianza.

La Figura 4 muestra un gráfico de curva de varianzas de acuerdo con una realización de la presente solicitud. En el gráfico, el eje horizontal indica un número de trama de una señal de trama, y el eje vertical indica la magnitud de una varianza. Una primera curva de varianza muestra la tendencia de una primera varianza de cada señal de trama, y la primera curva de varianza muestra la tendencia de una segunda varianza de cada señal de trama. Como se puede ver en el gráfico, la varianza fluctúa ligeramente en la banda de alta frecuencia 2000~4000 Hz, y la varianza fluctúa mucho en la banda de baja frecuencia 0~2000 Hz. Esto puede demostrar que las señales sin ruido se concentran principalmente en la banda de baja frecuencia.

Como se describió más arriba, en una realización preferida de la presente solicitud, la etapa S1031 puede incluir específicamente:

determinar si la primera varianza de los valores de potencia de la señal de trama es mayor que un primer umbral T_1 ; y en caso afirmativo, determinar la señal de trama como una señal de ruido. Usando la señal de trama f_1' como ejemplo, se determina si la primera varianza $Var_{alta}(f_1')$ es mayor que el primer umbral T_1 .

En la realización de la presente solicitud, la etapa de más arriba S103 puede incluir además específicamente:

determinar si una diferencia entre la primera varianza y la segunda varianza es mayor que un segundo umbral T_2 ; y en caso negativo, determinar la señal de trama como señal de ruido.

Usando la señal de trama f_1' como ejemplo, una diferencia entre la primera varianza y la segunda varianza es $|Var_{alta}(f_1') - Var_{baja}(f_1')|$. Si $|Var_{alta}(f_1') - Var_{baja}(f_1')| < T_2$, la señal de trama f_1' se determina como una señal de ruido. Las señales de ruido se pueden determinar secuencialmente a partir de las señales de trama de voz a analizar $\{f_1', f_2', \dots, f_n'\}$ de acuerdo con esta etapa.

En la realización de la presente solicitud, entre la etapa S102 y la etapa S103, el método puede incluir, además: clasificar las señales de trama en el segmento de señal de voz a analizar de acuerdo con las magnitudes de las varianzas.

Luego, determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza incluye:

determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza de los valores de potencia de cada señal de trama clasificada en varias frecuencias.

Como se describió más arriba, las varianzas $\{Var(f_1'), Var(f_2'), \dots, Var(f_n')\}$ de los valores de potencia de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ se puede determinar en esta realización. Las señales de trama se pueden clasificar en orden ascendente de las varianzas de los valores de potencia. Una señal con una varianza menor es más probable que sea una señal de ruido. Por lo tanto, las señales de trama de ruido en la señal de voz a analizar pueden clasificarse al frente. En la realización de la presente solicitud, si las varianzas se adquieren respectivamente a través de estadísticas en la banda de baja frecuencia (por ejemplo, 0~2000 Hz) y la banda de alta frecuencia (por ejemplo, 2000~4000 Hz), los valores de potencia de cada una de las señales de trama $\{f_1', f_2', \dots, f_n'\}$ en varias frecuencias se pueden clasificar en un primer conjunto de valores de potencia A correspondiente a un primer intervalo de frecuencia (por ejemplo, 0~2000 Hz) y un segundo conjunto de valores de potencia B correspondiente a un segundo intervalo de frecuencia (por ejemplo, 2000~4000 Hz) de acuerdo con los intervalos de frecuencia a los que pertenecen las frecuencias correspondientes al espectro de potencia de la señal de trama. Luego, las primeras varianzas $\{Var_{baja}(f_1'), Var_{baja}(f_2'), \dots, Var_{baja}(f_n')\}$ de los valores de potencia incluidos en los primeros conjuntos de valores de potencia correspondientes a las señales de trama $\{f_1', f_2', \dots, f_n'\}$ se pueden determinar respectivamente, y segundas varianzas $\{Var_{alta}(f_1'), Var_{alta}(f_2'), \dots, Var_{alta}(f_n')\}$ de los valores de potencia incluidos en los segundos conjuntos de valores de potencia correspondientes a las señales de trama $\{f_1', f_2', \dots, f_n'\}$ se pueden determinar respectivamente. En la etapa S104 de más arriba, en base a las estadísticas de varianza a frecuencias altas y bajas, las señales de ruido incluidas en las señales de voz a analizar (que pueden ser señales de voz clasificadas de acuerdo con las magnitudes de las varianzas) se pueden determinar de la siguiente manera:

$$\begin{aligned} Var_{baja}(f_i') > T_1 & \quad (1); \\ |Var_{alta}(f_i') - Var_{baja}(f_i')| > T_2 & \quad (2); \\ Var_{alta}(f_{i+1}') - Var_{alta}(f_{i-1}') > T_3 & \quad (3); \\ Var_{alta}(f_{i+1}') - Var_{alta}(f_{i-1}') > T_3 & \quad (4); \end{aligned}$$

$i \in (1, n)$. Se puede determinar con base en la fórmula (1) si una primera varianza de los valores de potencia de cada señal de trama f_i' es mayor que un primer umbral T_1 . Si no, la señal de trama f_i' se determina como una señal de trama de ruido. Un conjunto de señales de trama de ruido determinadas se determina como una señal de ruido.

Se puede determinar con base en la fórmula (2) si una segunda varianza de los valores de potencia de cada señal de trama f_i' es mayor que un segundo umbral T_2 . Si no, la señal de trama f_i' se determina como una señal de trama de ruido. Un conjunto de señales de trama de ruido determinadas se determina como una señal de ruido.

Se puede determinar con base en la fórmula (3) si una diferencia $Var_{alta}(f_{i+1}') - Var_{alta}(f_{i-1}')$ entre una segunda varianza $Var_{alta}(f_{i-1}')$ de los valores de potencia de una la señal de trama f_{i-1} antes de una señal de trama f_i' y una segunda varianza $Var_{alta}(f_{i+1}')$ de los valores de potencia de una señal de trama f_{i+1} junto a la señal de trama f_i' es mayor de un tercer umbral T_3 . Si no, la señal de trama f_i' se determina como una señal de trama de ruido. Un conjunto de señales de trama de ruido determinadas se determina como una señal de ruido.

Se puede determinar con base en la fórmula (4) si una diferencia $Var_{baja}(f_{i+1}') - Var_{baja}(f_{i-1}')$ entre una primera varianza $Var_{baja}(f_{i-1}')$ de los valores de potencia de un señal de trama f_{i-1} antes de una señal de trama f_i' y una primera varianza $Var_{baja}(f_{i+1}')$ de los valores de potencia de una señal de trama f_{i+1} al lado de la señal de trama f_i' es mayor de un cuarto umbral T_4 . Si no, la señal de trama f_i' se determina como una señal de trama de ruido. Un conjunto de señales de trama de ruido determinadas se determina como una señal de ruido.

En la realización de la presente solicitud, las tramas de ruido incluidas en la señal de voz a analizar pueden reconocerse usando las fórmulas (1) a (4) de más arriba. Es decir, cualquier señal de trama f_i' que cumpla con cualquiera de las fórmulas de más arriba (1) a (4) puede determinarse como una señal sin ruido (una trama final de ruido). En otras palabras, cualquier señal de trama f_i' cumplir ninguna de las fórmulas de más arriba (1) a (4) se puede determinar como una señal de ruido. Se puede determinar una trama final de ruido f_m' en base al proceso de más arriba, y luego las tramas de ruido incluyen: $\{f_1', f_2', \dots, f_{m-1}'\}$.

Cabe señalar que, en otras realizaciones de la presente solicitud, la trama final de ruido puede determinarse en base a algunas de las fórmulas (1) a (4), tales como las fórmulas (1) y (2), o las fórmulas (2) y (3). Además, las fórmulas para determinar la trama final de ruido en la realización de la presente solicitud no se limitan a las fórmulas enumeradas más arriba. Los umbrales T_1 , T_2 , T_3 y T_4 se obtienen de las estadísticas de una gran cantidad de muestras de prueba.

La Figura 5 es un diagrama de flujo de un método de eliminación de ruido de voz de acuerdo con una realización de la presente solicitud, que incluye las siguientes etapas:

- S201: Se determina un segmento de señal de voz a analizar incluido en una voz a procesar.
- S202: Se realiza la transformada de Fourier en cada señal de trama en el segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama en el segmento de señal de voz.
- S203: Se determina una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a varias frecuencias en función del espectro de potencia de la señal de trama.
- S204: Se determina si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza, y se obtienen varias tramas de ruido incluidas en el segmento de señal de voz.

S205: Se determina una potencia promedio correspondiente a las varias tramas de ruido incluidas en el segmento de señal de voz, y la voz a procesar se elimina en función de la potencia promedio de las tramas de ruido.

5 En la realización de la presente solicitud, después de que las tramas de ruido $\{f_1', f_2', \dots, f_{m-1}'\}$ incluidas en un segmento de voz a analizar se adquieren de acuerdo con el método de más arriba, los números de trama de señales originales (antes de la clasificación) correspondientes a las tramas de ruido respectivamente, se puede determinar, y se puede obtener una potencia promedio de estas señales de trama a través de estadísticas para obtener un valor de estimación del espectro de potencia P_{ruido} de la señal de ruido. La voz puede ser eliminada después de obtener el
10 valor de estimación del espectro de potencia P_{ruido} de la señal de ruido. El método de eliminación de ruido es bien conocido por los expertos en la técnica y no se describirá específicamente aquí.

Definitivamente, en otras formas de realización factibles de la presente solicitud, se puede omitir la etapa de clasificar las señales de trama de acuerdo con las varianzas, y las tramas de ruido se pueden determinar directamente en función de las varianzas de las señales originales. Además, después de determinar múltiples tramas de señal de ruido en la presente solicitud, el valor de estimación del espectro de potencia P_{ruido} generalmente se calcula usando algunos de las tramas, para evitar la sobreestimación. Por ejemplo, se pueden capturar los primeros 30 tramas para calcular el valor de estimación del espectro de potencia del P_{ruido} si la señal de ruido determinada incluye 50 tramas. Como tal, se puede mejorar la precisión del valor de estimación del espectro de potencia.

Una realización de la presente solicitud proporciona además un aparato de determinación de señal de ruido correspondiente a la implementación del proceso de más arriba. El aparato se puede implementar a través de software, y también se puede implementar a través de hardware o una combinación de software y hardware. Al usar una manera de implementación de software como ejemplo, se puede formar un aparato en un sentido lógico leyendo un programa de computadora correspondiente a través de una Unidad Central de Procesamiento (CPU) de un servidor en una memoria y ejecutando el programa de computadora. Consulte la Figura 8 para una estructura de hardware del aparato.

La Figura 6 es un diagrama de bloques de un aparato de determinación de señal de ruido de acuerdo con una realización de la presente solicitud. En esta realización, las funciones de las unidades en el aparato pueden corresponder a las funciones de las etapas en el método de determinación de señal de ruido de más arriba. Consulte la realización del método de más arriba para más detalles. El aparato de determinación de señal de ruido
100 incluye:

35 una unidad de adquisición de espectro de potencia 101 configurada para realizar la transformada de Fourier en cada señal de trama en un segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama en el segmento de señal de voz;
una unidad de determinación de varianza 102 configurada para determinar una varianza de valores de potencia de cada señal de trama en el segmento de señal de voz a diversas frecuencias en base a el espectro de potencia de la señal de trama; y
40 una unidad de determinación de ruido 103 configurada para determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza.

Preferentemente, el aparato incluye, además: una unidad de adquisición de segmento configurada para:
45 determinar un segmento de señal de voz con una variación de amplitud menor que un umbral preestablecido en una voz a procesar como el segmento de señal de voz a analizar en función de una variación de amplitud de una señal de dominio de tiempo de la voz a procesar; o
capturar las primeras N señales de voz de trama en una voz a procesar como el segmento de señal de voz a analizar.

50 Preferentemente, la unidad de determinación de ruido 103 está configurada para:
determinar si la varianza correspondiente a cada señal de trama en el segmento de señal de voz es mayor que un primer umbral; y
si no, determinar la señal de trama como una señal de ruido.

55 Preferentemente, la unidad de determinación de varianza 102 está configurada para:
al menos clasificar los valores de potencia de la señal de trama en varias frecuencias en un primer conjunto de valores de potencia correspondiente a un primer intervalo de frecuencia de acuerdo con los intervalos de frecuencia a los que pertenecen las frecuencias correspondientes al espectro de potencia; y
determinar una primera varianza de los valores de potencia incluidos en el primer conjunto de valores de
60 potencia.

Entonces la unidad de determinación de ruido 103 está configurada para:
determinar si la primera varianza es mayor que el primer umbral; y
si no, determinar la señal de trama como una señal de ruido.

65 Preferentemente, la unidad de determinación de varianza 102 está configurada específicamente para:

al menos clasificar los valores de potencia de cada señal de trama en varias frecuencias en un primer conjunto de valores de potencia correspondiente a un primer intervalo de frecuencia y un segundo conjunto de valores de potencia correspondiente a un segundo intervalo de frecuencia de acuerdo con los intervalos de frecuencia a los que pertenecen las frecuencias correspondientes a los valores de potencia de la señal de trama, en el que el primer intervalo de frecuencia es menor que el segundo intervalo de frecuencia; determinar una primera varianza de los valores de potencia incluidos en el primer conjunto de valores de potencia; y determinar una segunda varianza de los valores de potencia incluidos en el segundo conjunto de valores de potencia.

Entonces la unidad de determinación de ruido 103 está configurada para:
determinar si una diferencia entre la primera varianza y la segunda varianza que corresponde a cada señal de trama es mayor que un segundo umbral; y
si no, determinar la señal de trama como una señal de ruido.

Una realización de la presente solicitud proporciona además un aparato de eliminación de ruido de voz correspondiente a la implementación del proceso de más arriba. El aparato se puede implementar a través de software, y también se puede implementar a través de hardware o una combinación de software y hardware. Al usar una manera de implementación de software como ejemplo, se puede formar un aparato en un sentido lógico leyendo un programa de computadora correspondiente a través de una Unidad Central de Procesamiento (CPU) de un servidor en una memoria y ejecutando el programa de computadora. Consulte la Figura 8 para una estructura de hardware del aparato.

La Figura 7 es un diagrama de bloques de un aparato de eliminación de ruido de voz de acuerdo con una realización de la presente solicitud. En esta realización, las funciones de las unidades en el aparato pueden corresponder a las funciones de las etapas en el método de más arriba de eliminación de ruido de voz. Consulte la realización del método de más arriba para más detalles. En esta realización, el aparato de eliminación de ruido de voz 200 incluye:

- una unidad de determinación de segmento 201 configurada para determinar un segmento de señal de voz a analizar incluido en una voz a procesar;
- una unidad de adquisición de espectro de potencia 202 configurada para realizar la transformada de Fourier en cada señal de trama en el segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama en el segmento de señal de voz;
- una unidad de determinación de varianza 203 configurada para determinar una varianza de valores de potencia de cada señal de trama en el segmento de señal de voz a diversas frecuencias en base a el espectro de potencia de la señal de trama; una unidad de determinación de ruido 205 configurada para determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza, y obtener varias tramas de ruido incluidas en el segmento de señal de voz; y
- una unidad de eliminación de ruido de voz 10 configurada para determinar una potencia media correspondiente a las diversas tramas de ruido incluidas en el segmento de señal de voz, y eliminar el ruido de la voz a procesar en función de la potencia media de las tramas de ruido.

Preferentemente, el aparato incluye, además: una unidad de clasificación 204 configurada para: clasificar las señales de trama en el segmento de señal de voz a analizar de acuerdo con las magnitudes de las varianzas.

Entonces la unidad de determinación de ruido 205 está configurada específicamente para:
determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido en base a la varianza de los valores de potencia de cada señal de trama clasificado en varias frecuencias.

Al realizar la transformada de Fourier en un segmento de señal de voz a analizar para adquirir un espectro de potencia de cada señal de trama, determinar una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a analizar en varias frecuencias, y finalmente determinar si la señal de trama es una señal de ruido en base a la varianza, el método y el aparato de determinación de la señal de ruido, así como el método y aparato de eliminación de ruido de voz proporcionados en las realizaciones de la presente solicitud, pueden determinar con precisión varias tramas de ruido incluidas en el segmento de señal de voz a analizar. Se puede eliminar el ruido de la voz a procesar en función de una potencia promedio de las diversas tramas de ruido determinados en el proceso de eliminación de ruido de voz, y por lo tanto se mejora el efecto de eliminación de ruido de voz.

Para facilitar la descripción, el aparato se divide en varias unidades en términos de funciones para las descripciones respectivas. Definitivamente, cuando se implementa la presente solicitud, las funciones de las unidades pueden implementarse en el mismo componente de software y/o hardware o múltiples componentes de software y/o hardware.

Los expertos en la técnica deben entender que las realizaciones de la presente invención pueden proporcionarse como un método, un sistema, o un producto de programa informático. Por lo tanto, la presente invención puede

5 implementarse como una realización de hardware completa, una realización de software completa o una realización que combina software y hardware. Además, la presente invención puede estar en forma de un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por una computadora (que incluyen, pero no se limitan a, una memoria de disco magnético, un CD-ROM, una memoria óptica, y similares) que incluyen código de programa utilizable por una computadora.

10 La presente invención se describe con referencia a diagramas de flujo y/o diagramas de bloques de acuerdo con el método, el dispositivo (sistema) y el producto de programa informático de acuerdo con las realizaciones de la presente invención. Debe entenderse que las instrucciones del programa informático pueden usarse para implementar cada proceso y/o bloque y una combinación de procesos y/o bloques en los diagramas de flujo y/o diagramas de bloque. Las instrucciones del programa de computadora se pueden proporcionar a una computadora de propósito general, una computadora de propósito especial, un procesador incorporado u otro dispositivo de procesamiento de datos programable para generar una máquina, de modo que la computadora o un procesador de otro dispositivo de procesamiento de datos programable ejecute una instrucción para generar un aparato configurado para implementar funciones designadas en uno o más procesos en un diagrama de flujo y/o uno o más bloques en un diagrama de bloques.

20 Las instrucciones del programa de computadora también pueden almacenarse en un almacenamiento de lectura de la computadora que puede guiar a una computadora u otro dispositivo de procesamiento de datos programable para trabajar de una manera específica, de modo que la instrucción almacenada en el almacenamiento de lectura de la computadora genere una fabricación que incluya un aparato de instrucción que implemente funciones designadas por uno o más procesos en un diagrama de flujo y/o uno o más bloques en un diagrama de bloques.

25 Las instrucciones del programa informático también pueden cargarse en una computadora u otro dispositivo de procesamiento de datos programable, de modo que se ejecuten una serie de etapas de operación en la computadora u otro dispositivo programable, para generar un procesamiento implementado por la computadora. Por lo tanto, la instrucción ejecutada en la computadora u otro dispositivo programable proporciona etapas para implementar funciones designadas en uno o más procesos en un diagrama de flujo y/o uno o más bloques en un diagrama de bloques.

30 Cabe señalar además que el término "incluir", "comprender" u otras variaciones de estos están destinados a cubrir no excluyentes, de modo que un proceso, método, producto o dispositivo que incluye una serie de elementos no solo incluye los elementos, sino que también incluye otros elementos que no están claramente enumerados, o incluye elementos inherentes del proceso, método, producto o dispositivo. En un caso sin más limitaciones, un elemento definido por "que incluye un ..." no excluye que el proceso, método, producto o dispositivo que incluye el elemento tenga además otros elementos idénticos.

40 Los expertos en la técnica deben entender que las realizaciones de la presente solicitud pueden proporcionarse como un método, un sistema, o un producto de programa informático. Por lo tanto, la presente solicitud puede implementarse en forma de una realización de hardware completa, una realización de software completa o una realización que combina software y hardware. Además, la presente solicitud puede estar en la forma de un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por la computadora (que incluyen, pero no se limitan a, una memoria de disco magnético, un CD-ROM, una memoria óptica, y similares) que incluyen código de programa utilizable por una computadora.

45 La presente solicitud puede describirse en un contexto común de una instrucción ejecutable por computadora ejecutada por una computadora, por ejemplo, un módulo de programa. Generalmente, el módulo del programa incluye una rutina, un programa, un objeto, un ensamblaje, una estructura de datos y similares utilizados para ejecutar una tarea específica o implementar un tipo de datos abstracto específico. La presente solicitud también se puede implementar en entornos informáticos distribuidos, en los que se ejecuta una tarea utilizando dispositivos de procesamiento remoto conectados a través de una red de comunicaciones. En los entornos informáticos distribuidos, el módulo del programa puede ubicarse en medios de almacenamiento informáticos locales y remotos, incluido un dispositivo de almacenamiento.

55 Las realizaciones en la especificación se describen progresivamente, se pueden obtener partes idénticas o similares de las realizaciones con referencia entre sí, y cada realización enfatiza una parte diferente de otras realizaciones. Especialmente, la realización del sistema es básicamente similar a la realización del método, por lo que se describe simplemente. Para partes relacionadas, consulte las descripciones de las partes en la realización del método.

60 Las descripciones de más arriba son meramente realizaciones de la presente solicitud, y no pretenden limitar la presente solicitud. Varias modificaciones y variaciones de la presente solicitud son posibles para los expertos en la técnica. El alcance de la presente invención está definido por las reivindicaciones adjuntas.

REIVINDICACIONES

- 5 1. Un método para determinar una señal de ruido en un segmento de señal de voz, el método que comprende:
 10 determinar un segmento de señal de voz con una variación de amplitud menor que un umbral preestablecido en una voz a procesar como el segmento de señal de voz con base en una variación de amplitud de una señal de dominio de tiempo de la voz a procesar;
 realizar (S101) una transformada de Fourier en cada señal de trama en el segmento de señal de voz para adquirir un espectro de potencia de cada señal de trama en el segmento de señal de voz, en donde el
 15 espectro de potencia comprende múltiples valores de potencia correspondientes a diferentes frecuencias;
 determinar (S102) una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a diversas frecuencias en función del espectro de potencia de la señal de trama; y
 determinar (S103) si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en la varianza.
- 15 2. El método de la reivindicación 1, en donde la etapa de determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en la varianza comprende:
 determinar (S1031) si la varianza correspondiente a cada señal de trama en el segmento de señal de voz es mayor que un primer umbral; y
 20 en caso negativo, determinar (S1032) la señal de trama como señal de ruido.
- 25 3. El método de la reivindicación 2, en donde la etapa de determinar una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a varias frecuencias con base en el espectro de potencia de la señal de trama comprende:
 al menos clasificar los valores de potencia de la señal de trama en varias frecuencias en un primer conjunto de valores de potencia correspondiente a un primer intervalo de frecuencia de acuerdo con los intervalos de
 30 frecuencia a los que pertenecen las frecuencias correspondientes al espectro de potencia; y
 determinar (S1022) una primera varianza de valores de potencia comprendidos en el primer conjunto de valores de potencia;
 entonces la etapa de determinar si la varianza es mayor que un primer umbral comprende determinar si la primera varianza es mayor que el primer umbral.
- 35 4. El método de la reivindicación 1, en donde la etapa de determinar una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a varias frecuencias en base a el espectro de potencia de la señal de trama comprende:
 al menos clasificar los valores de potencia (S1021) de cada señal de trama en varias frecuencias en un primer conjunto de valores de potencia correspondiente a un primer intervalo de frecuencia y un segundo
 40 conjunto de valores de potencia correspondiente a un segundo intervalo de frecuencia de acuerdo con los intervalos de frecuencia a los que pertenecen las frecuencias correspondientes a los valores de potencia de la señal de trama, en donde el primer intervalo de frecuencia es menor que el segundo intervalo de frecuencia;
 determinar (S1022) una primera varianza de valores de potencia comprendidos en el primer conjunto de valores de potencia; y
 45 determinar (S1023) una segunda varianza de valores de potencia comprendidos en el segundo conjunto de valores de potencia;
 entonces la etapa de determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en las varianzas comprende:
 determinar si una diferencia entre la primera varianza y la segunda varianza que corresponde a cada
 50 señal de trama es mayor que un segundo umbral; y
 en caso negativo, determinar la señal de trama como señal de ruido.
- 55 5. El método de la reivindicación 1, en donde después de la etapa de determinar una varianza de los valores de potencia de cada señal de trama en el segmento de señal de voz a varias frecuencias con base en el espectro de potencia de la señal de trama y antes de la etapa de determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en la varianza, el método comprende, además:
 clasificar las señales de trama en el segmento de señal de voz de acuerdo con las magnitudes de las
 60 varianzas;
 entonces la etapa de determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en la varianza comprende determinar si cada señal de trama en el segmento de señal de voz es una señal de ruido con base en la varianza de los valores de potencia de cada señal de trama clasificada

en varias frecuencias.

- 5 6. Un aparato (100) para determinar señales de ruido en un segmento de señal de voz, el aparato que comprende múltiples unidades (101, 102, 103) configuradas para realizar el método de cualquiera de las reivindicaciones 1 a 5.

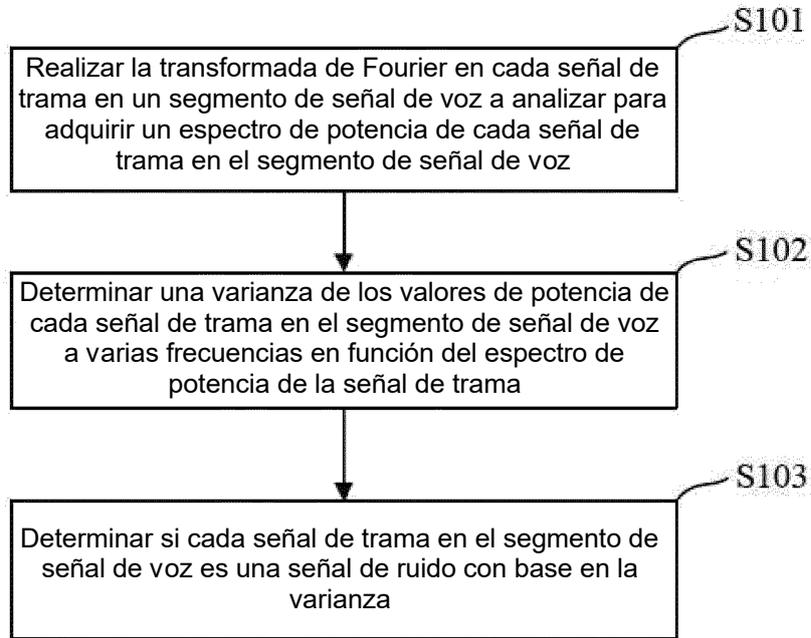


Figura 1

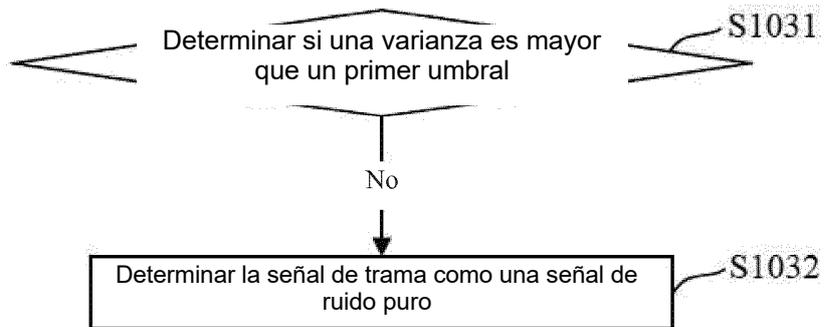


Figura 2

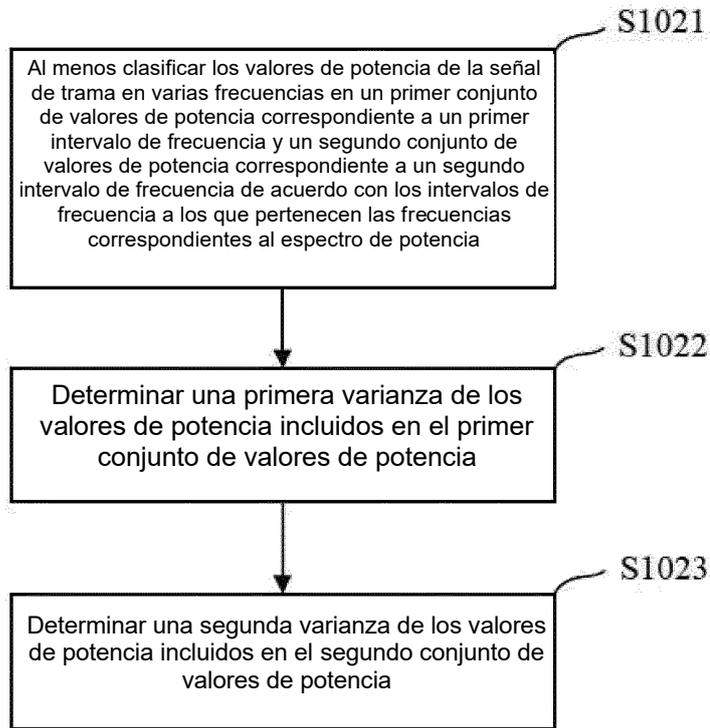


Figura 3

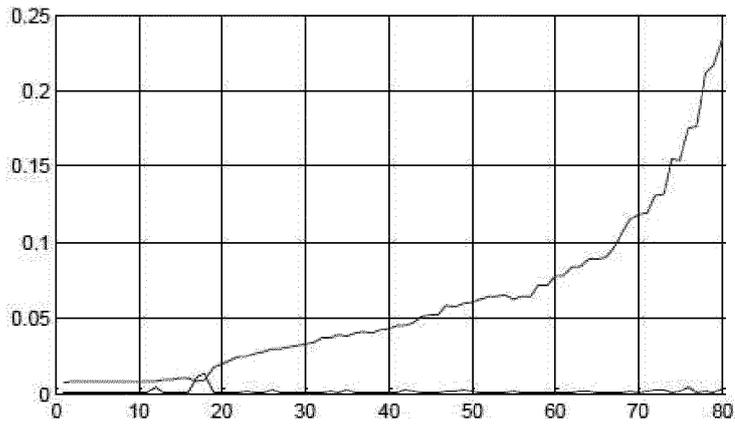


Figura 4

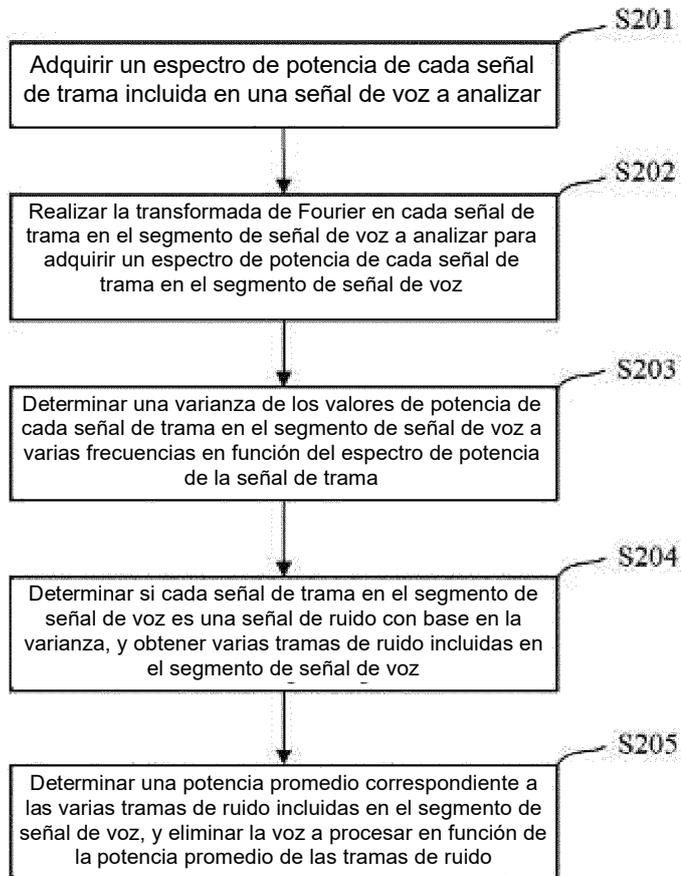


Figura 5

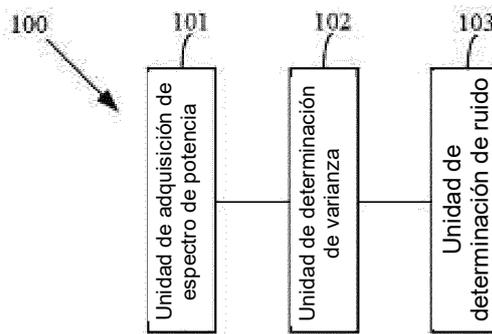


Figura 6

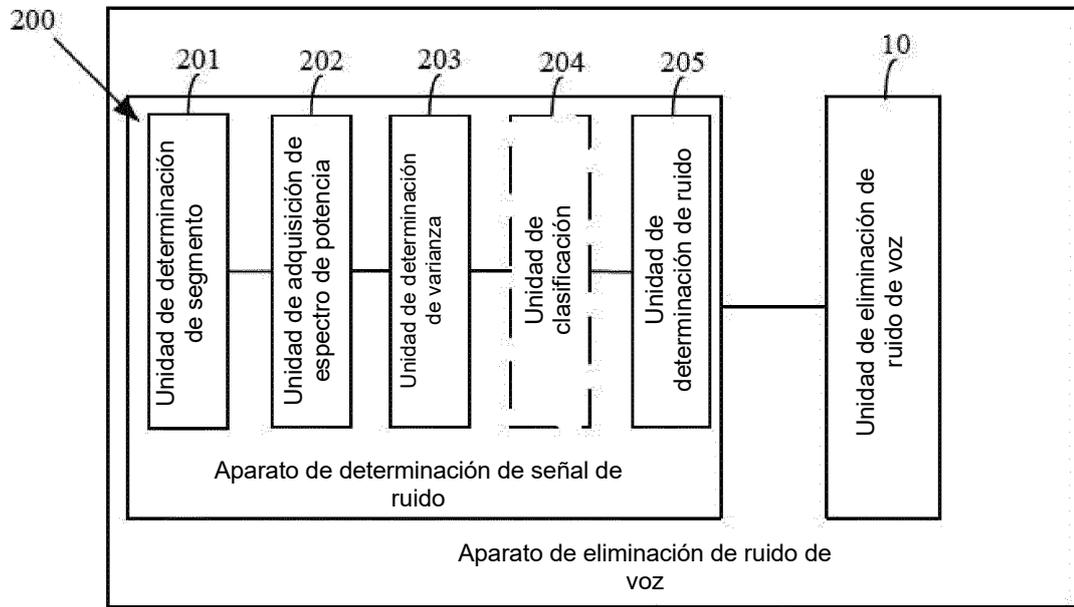


Figura 7

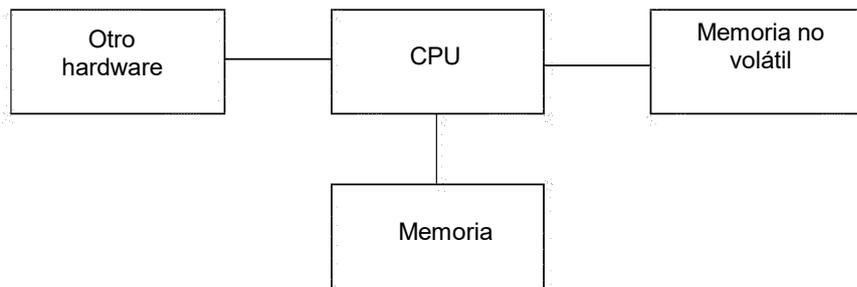


Figura 8