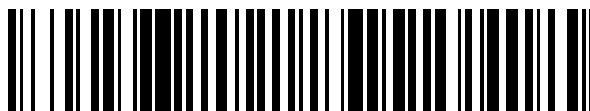


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 602 060**

51 Int. Cl.:

G10L 21/0208 (2013.01)

G10L 21/0216 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.10.2014** **E 14188582 (2)**

97 Fecha y número de publicación de la concesión europea: **17.08.2016** **EP 2863392**

54 Título: **Reducción de ruido en sistemas multi-micrófono**

30 Prioridad:

21.10.2013 GB 201318597

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

17.02.2017

73 Titular/es:

NOKIA TECHNOLOGIES OY (100.0%)
Karaportti 3
02610 Espoo, FI

72 Inventor/es:

NIEMISTÖ, RIITTA y
MYLLYLÄ, VILLE

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 602 060 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

Reducción de ruido en sistemas multi-micrófono

5 Campo

La presente solicitud se refiere a aparatos y métodos para la implementación de la reducción de ruido o de mejora de audio en sistemas multi-micrófono y, específicamente, pero no solo, la aplicación de la reducción de ruido o de la mejora de audio en sistemas multi-micrófono dentro de un aparato móvil.

10 Antecedentes

Los sistemas de grabación de audio pueden hacer uso de más de un micrófono para recoger y grabar audio en el medio ambiente circundante.

Un sistema multi-micrófono a modo de ejemplo se describe en el documento US 2012/0051548 A1.

Estos sistemas multi-micrófono (o sistemas MMic) permiten la implementación de un procesamiento de señal digital tal como la mejora de la voz que se aplica a las salidas de micrófono. La intención en la mejora de la voz es el uso de métodos matemáticos para mejorar la calidad de la voz, presentada como señales digitales. Una implementación de la mejora de la voz se preocupa del enlace ascendente que procesa las señales de audio de tres entradas o micrófonos.

25 Sumario

De acuerdo con un primer aspecto, se proporciona un método que comprende: recibir al menos tres señales de audio de micrófono, comprendiendo las al menos tres señales de audio de micrófono al menos dos señales de audio de micrófono cercanas generadas por al menos dos micrófonos cercanos localizados cerca de una fuente de audio deseada y al menos una señal de audio de micrófono lejano generada por un micrófono lejano localizado más allá de la fuente de audio deseada que los al menos dos micrófonos cercanos; generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono, siendo la primera selección a partir de las señales de audio de micrófono cercanas; generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono, la al menos una selección adicional de las al menos tres señales de audio de micrófono, siendo la segunda selección de todas las señales de micrófono; determinar, a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional, la señal de audio con una mayor supresión de ruido.

La mayor supresión de ruido puede comprender una supresión de ruido mejorada.

Recibir al menos tres señales de audio de micrófono puede comprender: recibir una primera señal de audio de micrófono de un primer micrófono cercano localizado sustancialmente en una parte delantera de un aparato; recibir una segunda señal de audio de micrófono de un segundo micrófono cercano localizado sustancialmente en una parte trasera del aparato; y recibir una tercera señal de audio de micrófono de un micrófono lejano localizado sustancialmente en el extremo opuesto de los micrófonos primero y segundo.

Generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono puede comprender generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda.

Generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono comprende generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono.

El método puede comprender además: generar una señal de audio de haz principal: aplicando un primer filtro de respuesta de impulso finita a la primera señal de audio; aplicando un segundo filtro de respuesta de impulso finita a la segunda señal de audio; y combinando la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio de haz principal; y generar una señal de audio anti-haz: aplicando un tercer filtro de respuesta de impulso finita a la primera señal de audio; aplicando un cuarto filtro de respuesta de impulso finita a la segunda señal de audio; y combinando la salida del tercer filtro de respuesta de impulso finita y del cuarto filtro de respuesta de impulso finita para generar la señal de audio anti-haz.

Generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono puede comprender filtrar la señal de audio de haz principal basada en la tercera señal de audio de micrófono.

Generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda puede comprender filtrar la señal de audio de haz principal basada en la señal de audio anti-haz.

5 Generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono puede comprender: seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la primera señal de audio procesada.

15 Generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono puede comprender: seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional.

25 Filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada puede comprender además filtrar la supresión de ruido de la primera entrada de procesamiento basándose en la segunda entrada de procesamiento.

30 El método puede comprender además conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz.

35 Conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz puede comprender: aplicar un primer filtro de respuesta de impulso finita a una primera de las al menos dos de las al menos tres señales de audio de micrófono; aplicar un segundo filtro de respuesta de impulso finita a una segunda de las al menos dos de las al menos tres señales de audio de micrófono; y combinar la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio conformada en haz.

40 El método puede comprender la supresión de ruido de canal único de la señal de audio con una mayor supresión de ruido, en el que la supresión de ruido de canal único comprende: generar un indicador que muestre si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; estimar y actualizar un ruido de fondo de la señal de audio cuando el indicador muestra que el período de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; procesar la señal de audio basándose en la estimación de ruido de fondo para generar una señal de audio de ruido suprimido.

45 Generar un indicador que muestra si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido puede comprender: normalizar una selección de las al menos tres señales de audio de micrófono, en el que la selección comprende: unas señales de audio conformadas de haz de al menos dos de las al menos tres señales de audio de micrófono; y las señales de audio de micrófono; filtrar las selecciones normalizadas de las al menos tres señales de audio de micrófono; comparar las selecciones normalizadas filtradas para determinar una relación de diferencia de potencia; generar el indicador que muestra un período de la señal de audio que comprende una falta de componentes de voz o es significativamente un ruido donde al menos una comparación de las selecciones normalizadas filtradas tiene una relación de diferencia de potencia mayor que un umbral determinado.

50 Determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido puede comprender al menos uno de: determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de señal más alta; y determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de potencia más alta. De acuerdo con un segundo aspecto, se proporciona un aparato que comprende al menos un procesador y al menos una memoria que incluye un código informático para uno o más programas, estando la al menos una memoria y el código informático configurados para con el al menos un procesador hacer que el aparato: reciba al menos tres señales de audio de micrófono, comprendiendo las al menos tres señales de audio de micrófono al menos dos señales de audio de micrófono cercanas generadas por al menos dos micrófonos cercanos localizados cerca de una fuente de audio

- deseada y al menos una señal de audio micrófono lejano generada por un micrófono lejano localizado más allá de la fuente de audio deseada que los al menos dos micrófonos cercanos; genere una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono, siendo la primera selección a partir de las señales de audio de micrófono cercanas; genere al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono, la al menos una selección adicional de las al menos tres señales de audio de micrófono, siendo la segunda selección de todas las señales de micrófono; determine a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido.
- 10 Recibir al menos tres señales de audio de micrófono puede hacer que el aparato: reciba una primera señal de audio de micrófono de un primer micrófono cercano localizado sustancialmente en una parte delantera de un aparato; reciba una segunda señal de audio de micrófono de un segundo micrófono cercano localizado sustancialmente en una parte trasera del aparato; y reciba una tercera señal de audio de micrófono de un micrófono lejano localizado sustancialmente en el extremo opuesto de los micrófonos primero y segundo.
- 15 Generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono puede hacer que el aparato genere una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda.
- 20 Generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono puede hacer que el aparato genere una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono.
- 25 Puede hacerse además que el aparato: genere una señal de audio de haz principal aplicando un primer filtro de respuesta de impulso finita a la primera señal de audio; aplicando un segundo filtro de respuesta de impulso finita a la segunda señal de audio; y combinando la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio de haz principal; y genere una señal de audio anti-haz: aplicando un tercer filtro de respuesta de impulso finita a la primera señal de audio; aplicando un cuarto filtro de respuesta de impulso finita a la segunda señal de audio; y combinando la salida del tercer filtro de respuesta de impulso finita y del cuarto filtro de respuesta de impulso finita para generar la señal de audio anti-haz.
- 30
- 35 Generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono puede hacer que el aparato filtre la señal de audio de haz principal basándose en la tercera señal de audio de micrófono.
- 40 Generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda puede hacer que el aparato filtre la señal de audio de haz principal basándose en la señal de audio anti-haz.
- 45 Generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono puede hacer que el aparato: seleccione como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; seleccione como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; filtre la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la primera señal de audio procesada.
- 50
- 55 Generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono puede hacer que el aparato: seleccione como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; seleccione como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; filtre la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional.
- 60
- 65 Filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional puede hacer que el aparato filtre la supresión de ruido de la primera entrada de procesamiento basándose en la segunda entrada de procesamiento. Puede hacerse que el aparato conforme el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de

audio conformada en haz.

Conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz puede hacer que el aparato: aplique un primer filtro de respuesta de impulso finita a una primera de las al menos dos de las al menos tres señales de audio de micrófono; aplique un segundo filtro de respuesta de impulso finita a una segunda de las al menos dos de las al menos tres señales de audio de micrófono; y combine la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio conformada en haz.

Puede hacerse que el aparato suprima el ruido de canal único de la señal de audio con una mayor supresión de ruido, en el que la supresión de ruido de canal único puede hacer que el aparato: genere un indicador que muestre si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; estime y actualice un ruido de fondo a partir de la señal de audio cuando el indicador muestra que el periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; procese la señal de audio basándose en la estimación de ruido de fondo para generar una señal de audio de ruido suprimido.

Generar un indicador que muestre si un periodo de la señal de audio que comprende una falta de componentes de voz o es significativamente un ruido puede hacer que el aparato: normalice una selección de las al menos tres señales de audio de micrófono, en el que la selección comprende: unas señales de audio conformadas de haz de al menos dos de las al menos tres señales de audio de micrófono; y las señales de audio de micrófono; filtre las selecciones normalizadas a partir de las al menos tres señales de audio de micrófono; compare las selecciones normalizadas filtradas para determinar una relación de diferencia de potencia; genere el indicador que muestra un período de la señal de audio que comprende una falta de componentes de voz o es significativamente un ruido en la que al menos una comparación de las selecciones normalizadas filtradas tiene una relación de diferencia de potencia mayor que un umbral determinado.

Determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido que puede hacer que el aparato realice al menos uno de: determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de señal más alta; y determinar a partir de la primera señal de audio procesada y la al menos una señal de audio procesada adicional de la señal de audio con la salida de nivel de potencia más alta.

De acuerdo con un tercer aspecto, se proporciona un aparato que comprende: una entrada configurada para recibir al menos tres señales de audio de micrófono, comprendiendo las al menos tres señales de audio de micrófono al menos dos señales de audio de micrófono cercanas generadas por al menos dos micrófonos cercanos localizados cerca de una fuente de audio deseada y al menos una señal de audio de micrófono lejana generada por un micrófono lejano localizado más allá de la fuente de audio deseada que los al menos dos micrófonos cercanos; un primer módulo de cancelador de interferencias configurado para generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono, siendo la primera selección a partir de las señales de audio de micrófono cercanas; al menos un módulo de cancelador de interferencias adicional configurado para generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono, la al menos una selección adicional a partir de las al menos tres señales de audio de micrófono, siendo la segunda selección a partir de todas las señales de micrófono; un comparador configurado para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido.

La entrada puede configurarse para: recibir una primera señal de audio de micrófono de un primer micrófono cercano localizado sustancialmente en una parte delantera de un aparato; recibir una segunda señal de audio de micrófono de un segundo micrófono cercano localizado sustancialmente en una parte trasera del aparato; y recibir una tercera señal de audio de micrófono de un micrófono lejano localizado sustancialmente en el extremo opuesto de los micrófonos primero y segundo.

El primer módulo de cancelador de interferencias puede configurarse para generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda.

El al menos un módulo de cancelador de interferencias adicional puede configurarse para generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono.

El aparato puede comprender además: un conformador de haz del haz principal configurado para generar una señal de audio de haz principal que comprende un primer filtro de respuesta de impulso finita configurado para recibir la primera señal de audio; un segundo filtro de respuesta de impulso finita configurado para recibir la segunda señal de audio; y un combinador configurado para combinar la salida del primer filtro de respuesta de impulso finita y del

segundo filtro de respuesta de impulso finita para generar la señal de audio de haz principal; y un conformador de haz anti-haz configurado para generar una señal de audio anti-haz que comprende: un tercer filtro de respuesta de impulso finita configurado para recibir la primera señal de audio; un cuarto filtro de respuesta de impulso finita configurado para recibir la segunda señal de audio; y un combinador configurado para combinar la salida del tercer filtro de respuesta de impulso finita y del cuarto filtro de respuesta de impulso finita para generar la señal de audio anti-haz.

El al menos un módulo de cancelador de interferencias adicional puede comprender un filtro configurado para filtrar la señal de audio de haz principal basándose en la tercera señal de audio de micrófono.

El primer módulo de cancelador de interferencias puede comprender un filtro configurado para filtrar la señal de audio de haz principal basándose en la señal de audio anti-haz.

El primer módulo de cancelador de interferencias puede comprender: un selector configurado para seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; un segundo selector configurado para seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; un filtro configurado para filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la primera señal de audio procesada.

El al menos un generador de interferencias adicional puede comprender: un selector configurado para seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; un segundo selector configurado para seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; un filtro configurado para filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional.

El filtro puede configurarse para filtrar la supresión de ruido de la primera entrada de procesamiento basándose en la segunda entrada de procesamiento.

El aparato puede comprender un conformador de haz configurado para conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz.

El conformador de haz puede comprender: un primer filtro de respuesta de impulso finita configurado para filtrar una primera de las al menos dos de las al menos tres señales de audio de micrófono; un segundo filtro de respuesta finita configurado para filtrar una segunda de las al menos dos de las al menos tres señales de audio de micrófono; y un combinador configurado para combinar la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio conformada en haz.

El aparato puede comprender un supresor de ruido de canal único configurado para suprimir el ruido de la señal de audio con una mayor supresión de ruido, el supresor de ruido de canal único puede comprender: una entrada configurada para recibir un indicador que muestre si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; un estimador configurado para estimar y actualizar un ruido de fondo a partir de la señal de audio cuando el indicador muestra que el período de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; un filtro configurado para procesar la señal de audio con una mayor supresión de ruido basándose en la estimación de ruido de fondo para generar una señal de audio de ruido suprimido.

El aparato puede comprender un detector de actividad de voz configurado para generar un indicador que muestre si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido que comprende: un normalizador configurado para normalizar una selección a partir de las al menos tres señales de audio de micrófono, en el que la selección comprende: unas señales de audio conformadas de haz de al menos dos de las al menos tres señales de audio de micrófono; y las señales de audio de micrófono; un filtro configurado para filtrar las selecciones normalizadas de las al menos tres señales de audio de micrófono; un comparador configurado para comparar las selecciones normalizadas filtradas para determinar una relación de diferencia de potencia; un generador de indicador configurado para generar el indicador que muestra un período de la señal de audio con una mayor supresión de ruido que comprende una falta de componentes de voz o es significativamente un ruido donde al menos una comparación de las selecciones normalizadas filtradas tiene una relación de diferencia de potencia mayor que un umbral determinado.

El comparador configurado para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido puede configurarse para realizar al menos uno de: determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de señal más alta; y determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de potencia más alta.

De acuerdo con un cuarto aspecto, se proporciona un aparato que comprende: unos medios para recibir al menos tres señales de audio de micrófono, comprendiendo las al menos tres señales de audio de micrófono al menos dos señales de audio de micrófono cercanas generadas por al menos dos micrófonos cercanos localizados cerca de una fuente de audio deseada y al menos una señal de audio de micrófono lejana generada por un micrófono lejano localizado más allá de la fuente de audio deseada que los al menos dos micrófonos cercanos; unos medios para generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono, siendo la primera selección a partir de las señales de audio de micrófono cercanas; unos medios para generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono, la al menos una selección adicional de las al menos tres señales de audio de micrófono, siendo la segunda selección a partir de la totalidad de las señales de micrófono; unos medios para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido.

Los medios para recibir al menos tres señales de audio de micrófono pueden comprender: unos medios para recibir una primera señal de audio de micrófono de un primer micrófono cercano localizado sustancialmente en una parte delantera de un aparato; unos medios para recibir una segunda señal de audio de micrófono de un segundo micrófono cercano localizado sustancialmente en una parte trasera del aparato; y unos medios para recibir una tercera señal de audio de micrófono de un micrófono lejano localizado sustancialmente en el extremo opuesto de los micrófonos primero y segundo.

Los medios para generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono pueden comprender unos medios para generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda.

Los medios para generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono pueden comprender unos medios para generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono.

El aparato puede comprender además: unos medios para generar una señal de audio de haz principal que comprende: unos medios para aplicar un primer filtro de respuesta de impulso finita a la primera señal de audio; unos medios para aplicar un segundo filtro de respuesta de impulso finita a la segunda señal de audio; y unos medios para combinar la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio de haz principal; y unos medios para generar una señal de audio anti-haz pueden comprender: unos medios para aplicar un tercer filtro de respuesta de impulso finita a la primera señal de audio; unos medios para aplicar un cuarto filtro de respuesta de impulso finita a la segunda señal de audio; y unos medios para combinar la salida del tercer filtro de respuesta de impulso finita y del cuarto filtro de respuesta de impulso finita para generar la señal de audio anti-haz.

Los medios para generar una señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono pueden comprender unos medios para filtrar la señal de audio de haz principal basándose en la tercera señal de audio de micrófono.

Los medios para generar una primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda pueden comprender unos medios para filtrar la señal de audio de haz principal basándose en la señal de audio anti-haz.

Los medios para generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono pueden comprender: unos medios para seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; unos medios para seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas; unos medios para filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la primera señal de audio

procesada.

Los medios para generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono pueden comprender: unos medios para seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; unos medios para seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono; unos medios para filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional.

Los medios para filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional comprenden, filtrar la supresión de ruido de la primera entrada de procesamiento basándose en la segunda entrada de procesamiento. El aparato puede comprender además unos medios para conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz.

Los medios para conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz pueden comprender: un medio para aplicar un primer filtro de respuesta de impulso finita a una primera de las al menos dos de las al menos tres señales de audio de micrófono; unos medios para aplicar un segundo filtro de respuesta de impulso finita a una segunda de las al menos dos de las al menos tres señales de audio de micrófono; y unos medios para combinar la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio conformada en haz.

El aparato puede comprender además unos medios para la supresión de ruido de canal único de la señal de audio con una mayor supresión de ruido, en el que los medios para la supresión de ruido de canal único pueden comprender: unos medios para generar un indicador que muestre si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; unos medios para estimar y actualizar un ruido de fondo a partir de la señal de audio cuando el indicador muestra que el período de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido; unos medios para procesar la señal de audio basándose en la estimación de ruido de fondo para generar una señal de audio de ruido suprimido.

Los medios para generar un indicador que muestra si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido pueden comprender: unos medios para normalizar una selección de las al menos tres señales de audio de micrófono, en el que la selección comprende: unas señales de audio conformadas de haz de al menos dos de las al menos tres señales de audio de micrófono; y las señales de audio de micrófono; unos medios para filtrar las selecciones normalizadas de las al menos tres señales de audio de micrófono; unos medios para comparar las selecciones normalizadas filtradas para determinar una relación de diferencia de potencia; unos medios para generar el indicador que muestra un período de la señal de audio que comprende una falta de componentes de voz o es significativamente un ruido donde al menos una comparación de las selecciones normalizadas filtradas tiene una relación de diferencia de potencia mayor que un umbral determinado.

Los medios para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido comprenden al menos uno de: unos medios para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de señal más alta; y unos medios para determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional la señal de audio con la salida de nivel de potencia más alta.

Las realizaciones de la presente solicitud tienen el objetivo de abordar los problemas asociados con el estado de la técnica.

Sumario de las figuras

Para una mejor comprensión de la presente solicitud, se hará ahora referencia por medio de un ejemplo a los dibujos adjuntos, en los que:

- la figura 1 muestra esquemáticamente un aparato adecuado para emplearse en algunas realizaciones;
- la figura 2 muestra esquemáticamente un ejemplo de un aparato de tres micrófonos adecuado para emplearse en algunas realizaciones;
- la figura 3 muestra esquemáticamente un procesador de señal para un sistema de multi-micrófono de acuerdo con algunas realizaciones;
- la figura 4 muestra esquemáticamente un diagrama de flujo del funcionamiento del procesador de señal para el

sistema de multi-micrófono como se muestra en la figura 3 de acuerdo con algunas realizaciones;
la figura 5 muestra esquemáticamente unos diagramas de ganancia de ejemplo de los haces de señal de audio de haz principal y de anti-haz de acuerdo con algunas realizaciones;
la figura 6 muestra esquemáticamente un diagrama de flujo de ejemplo del funcionamiento del procesador de
señal basado en una entrada de control de acuerdo con algunas realizaciones; y
la figura 7 muestra un cancelador de interferencias adaptativo de ejemplo de acuerdo con algunas realizaciones.

Realizaciones

- 10 A continuación se describe con más detalle un aparato adecuado y los posibles mecanismos para el suministro del procesamiento de señal dentro de los sistemas multi-micrófono. Algunas implementaciones de mejora de voz que procesan señales digitales usan tres señales de micrófono (del número disponible de micrófonos en el aparato o acoplados al aparato). Dos de los micrófonos o las señales de entrada originadas por los 'nearmics', (en otros palabras, unos micrófonos que se localizan cerca uno del otro tal como en la parte inferior del dispositivo) y un tercer
- 15 micrófono, 'farmic', localizado más allá en el otro extremo del aparato o dispositivo. Un ejemplo de un aparato 10 de este tipo se muestra en la figura 2 que muestra el aparato con un primer micrófono (mic1) 101, un 'nearmics' delantero, localizado hacia la parte inferior del aparato y frente a la pantalla o parte delantera del aparato, un segundo micrófono (mic2) 103, un 'nearmics' trasero, mostrado por el óvalo discontinuo y localizado hacia la parte inferior del aparato y en la cara opuesta a la pantalla (o de otra manera en la parte trasera del aparato) y un tercer
- 20 micrófono (mic3) 105, un 'farmic', localizado en la "parte superior" del aparato 10. Aunque los ejemplos siguientes se describen con respecto a una configuración de sistema de 3 micrófonos, debería entenderse que en algunas realizaciones el sistema puede comprender más de 3 micrófonos a partir de lo cual puede hacerse una selección adecuada de 3 micrófonos.
- 25 Con dos o más nearmics es posible formar dos haces direccionales de las señales de audio generadas a partir de los micrófonos. Estos pueden ser, por ejemplo, como se muestra en la figura 5, un 'haz principal' 401 y un 'anti-haz' 403. En la voz local de 'haz principal' se pasa sustancialmente mientras se atenúa significativamente el ruido procedente de la dirección opuesta. En la voz local de 'anti-haz' se atenúa sustancialmente mientras que el ruido de otras direcciones se pasa sustancialmente. En tales situaciones, el nivel de ruido ambiente es casi el mismo en
- 30 ambos haces.
- Estos haces (el principal y los anti-haces) pueden usarse, en algunas realizaciones, en su procesamiento de las señales digitales adicionales para reducir aún más el ruido de fondo que queda de la señal de audio de haz principal usando un cancelador de interferencias adaptativo (AIC) y la sustracción espectral.
- 35 El cancelador de interferencias adaptativo (AIC) con dos señales de audio de micrófono cercanas puede realizar un primer método para cancelar aún más el ruido del haz principal. Aunque con una señal de audio nearmic y una conformación de haz de señal de audio farmic no es posible, un AIC puede usarse directamente con señales de micrófono. Además, el ruido puede reducirse aún más usando la sustracción espectral.
- 40 El primer método usa la conformación de haz de las señales de audio de micrófono para reducir el ruido, se entiende que para proporcionar reducciones de ruido eficaces, pero es sensible a cómo se mantiene el dispositivo. El segundo método que usa las señales de audio de micrófono directas es más robusto de orientación, pero no proporciona tanta eficacia como una reducción de ruido.
- 45 En ambos métodos, puede usarse un detector de actividad de voz espacial (VAD) para mejorar la supresión de ruido en comparación con el caso de un canal único sin información direccional disponible. Los VAD espaciales pueden por ejemplo combinarse con otros VAD en el procesamiento de señales y la estimación de ruido de fondo puede actualizarse cuando el detector de actividad de voz determina que la señal de audio no contiene componentes sonoros. En otras palabras, la estimación de ruido de fondo puede actualizarse cuando el método VAD marca un ruido. Un ejemplo de la detección de actividad de voz no espacial para mejorar la supresión de ruido se muestra en la patente estadounidense número 8.244.528.
- 50 En el caso del método de señal de audio de conformación de haz, la salida VAD espacial es normalmente la relación entre las potencias de haz principal y de anti-haz determinadas o estimadas. En el caso del método de señal de audio de micrófono directa, la salida VAD espacial es normalmente la relación entre las señales de entrada.
- Por lo tanto, en tales situaciones, el VAD espacial y el AIC son ambos sensibles a la colocación del aparato o dispositivo. Por ejemplo, cuando hay fugas de voz del anti-haz o segundo micrófono, el cancelador de interferencias adaptativo (AIC) o el supresor de ruido pueden considerarlo como un ruido y atenuar la voz local. Se entiende que el problema es más grave con los métodos de señal de audio de conformación de haz, pero también existe con los métodos de señal de audio de micrófono directo.
- 60 El concepto de la invención como se describe en las realizaciones en el presente documento implementa el procesamiento de señales de audio empleando un tercer o más micrófonos y aborda el problema de proporcionar la reducción de ruido que es a la vez eficaz y robusta en orientación.
- 65

En tales realizaciones como se describe en el presente documento el tercer o más micrófonos se emplean con el fin de lograr una reducción de ruido eficaz a pesar de la posición del aparato, por ejemplo, un teléfono colocado en la proximidad o en el oído del usuario. En el modo portable de mano, el altavoz se localiza normalmente cerca del oído del propio usuario (de lo contrario el usuario no puede escuchar nada), pero el micrófono puede estar localizado lejos de la boca del usuario. En tales circunstancias donde la reducción de ruido no es robusta en orientación el usuario en el otro extremo puede no escuchar nada.

Como se describe en el presente documento y se muestra con respecto a la figura 2, el aparato comprende al menos tres micrófonos, dos 'nearmics' y un 'farmic'.

En las realizaciones que se describen en el presente documento el concepto robusto direccional se implementa por un procesador de señal que comprende dos canceladores de interferencias de audio (AIC) que funcionan en paralelo. El primer AIC, primario o principal se configura para recibir las señales de haz principal y anti-haz como las entradas del primer AIC o principal. El segundo AIC o secundario se configura para recibir las señales de haz principal y del farmic como las entradas al segundo AIC o secundario. Por lo tanto, debería entenderse que el segundo AIC o secundario está configurado para recibir información de los tres micrófonos.

En tales realizaciones, pueden compararse los niveles de señal de salida de los AIC paralelos y donde hay una diferencia considerable (por ejemplo, un valor de diferencia predeterminado de 2 dB) en los niveles de salida, la señal que tiene el nivel más alto se usa como salida.

Una diferencia más pequeña en los niveles de salida puede explicarse por las diferentes capacidades de reducción de ruido de los dos AIC, mientras que una diferencia más grande sería indicativa de que el AIC atenúa la voz local cuyo nivel de señal de salida es menor. La excepción a esto sería cuando el ruido del viento causa problemas. Por lo tanto, en algunas realizaciones puede emplearse un detector de ruido del viento y cuando el detector de ruido del viento marca la detección de viento, se usa el primer AIC o principal.

En las realizaciones que se describen en el presente documento, el detector de actividad de voz espacial (VAD) puede configurarse para recibir como entrada cuatro señales: la señal de micrófono principal (o primer nearmic), la señal farmic, la señal de haz principal y la señal anti-haz. A continuación, como se describe en el presente documento estas señales pueden normalizarse de manera que sus niveles de ruido estacionarios son sustancialmente los mismos. Esta normalización se realiza para eliminar la posibilidad de variabilidad de micrófono debido a que las señales de micrófono pueden tener diferentes sensibilidades. A continuación, como se muestra en las realizaciones descritas en el presente documento los niveles de señal normalizados se comparan con intervalos de frecuencia predefinidos. Estos intervalos de frecuencia predefinidos o determinados pueden ser frecuencias bajas o más bajas para las señales de micrófono y determinadas basadas en el diseño del haz para las señales de audio de haz.

Cuando hay una diferencia considerable entre el nivel del haz principal y del anti-haz para las comparaciones de región de frecuencia, o unas diferencias considerables entre los niveles del micrófono principal y de la señal 'farmic', o unas diferencias considerables entre los niveles del haz principal y de la señal 'farmic', entonces como se describe en el presente documento el detector de actividad de voz espacial puede configurarse para emitir un indicador adecuado tal como un marcador espacial VAD para indicar que una estimación de ruido de la voz y del fondo usada en la supresión de ruido no debe actualizarse. Sin embargo, cuando los niveles de señal son los mismos (que, como se describe en el presente documento, se determina por la diferencia que es inferior a un umbral determinado) en todos estos pares de señales, entonces la señal grabada es más probable que sea un ruido de fondo (o que la colocación del aparato sea muy inusual) y puede actualizarse la estimación de ruido de fondo.

En los ejemplos siguientes se muestra el aparato funcionando en el modo portable de mano (en otras palabras, el aparato o el teléfono se encuentra en o cerca del oído o del usuario en general). Sin embargo, en algunas circunstancias las realizaciones pueden implementarse mientras el usuario está haciendo funcionar el aparato en un modo de altavoz (tal como colocándose lejos del usuario, pero de una manera que el usuario es todavía la fuente de audio más fuerte en el entorno).

La figura 1 muestra una visión general de un sistema adecuado dentro del cual pueden implementarse las realizaciones de la solicitud. La figura 1 muestra un ejemplo de un aparato o dispositivo electrónico 10. El aparato 10 puede usarse para capturar, grabar o escuchar las señales de audio y puede funcionar como un aparato de captura.

El aparato 10 puede ser, por ejemplo, un terminal móvil o equipo de usuario de un sistema de comunicación inalámbrica cuando funciona como un aparato de captura o grabación. En algunas realizaciones, el aparato puede ser un grabador de audio, tal como un reproductor de MP3, un medio de grabación/reproducción (también conocido como un reproductor MP4), o cualquier aparato portable adecuado conveniente para grabar audio o una videocámara de audio/vídeo o una grabadora de audio o vídeo en memoria.

El aparato 10 puede comprender en algunas realizaciones un subsistema de audio. Por ejemplo, el subsistema de audio puede comprender en algunas realizaciones, al menos, tres micrófonos o una serie de micrófonos 11 para

- capturar la señal de audio. En algunas realizaciones, los al menos tres micrófonos o serie de micrófonos pueden ser un micrófono de estado sólido, en otras palabras capaz de capturar señales de audio y emitir una señal de formato digital adecuada. En algunas otras realizaciones, los al menos tres micrófonos o serie de micrófonos 11 pueden comprender cualquier micrófono adecuado o medios de captura de audio, por ejemplo, un micrófono de condensador, un micrófono capacitor, un micrófono electrostático, un micrófono de condensador Electret, un micrófono dinámico, un micrófono de cinta, un micrófono de carbono, un micrófono piezoeléctrico, o un micrófono de sistema micro eléctrico-mecánico (MEMS). En algunas realizaciones, los micrófonos 11 son micrófonos digitales, en otras palabras, configurados para generar una salida de señal digital (no necesitando de este modo un convertidor analógico a digital). Los micrófonos 11 o la serie de micrófonos pueden, en algunas realizaciones emitir la señal capturada de audio a un convertidor analógico a digital (CAD) 14.
- En algunas realizaciones, el aparato puede comprender además, un convertidor analógico a digital (CAD) 14 configurado para recibir la señal de audio capturada analógica desde los micrófonos y emitir la señal capturada de audio en un formato digital adecuado. El convertidor analógico a digital 14 puede ser cualquier medio de conversión o de procesamiento analógico a digital adecuado. En algunas realizaciones, los micrófonos son micrófonos "integrados" que contienen tanto la generación de señal de audio como la capacidad de conversión analógica a digital.
- En algunas realizaciones, los subsistemas de audio del aparato 10 comprenden además un convertidor digital a analógico 32 para convertir las señales de audio digitales de un procesador 21 a un formato analógico adecuado. El convertidor digital a analógico (CDA) o los medios de procesamiento de señales 32 pueden en algunas realizaciones ser cualquier tecnología CDA adecuada.
- Además, el subsistema de audio puede comprender en algunas realizaciones un altavoz 33. El altavoz 33 puede en algunas realizaciones recibir la salida del convertidor digital a analógico 32 y presentar la señal de audio analógica al usuario. En algunas realizaciones, el altavoz 33 puede ser representativo de una disposición de múltiples altavoces, auriculares, por ejemplo, un conjunto de auriculares o auriculares inalámbricos.
- Aunque se muestra el aparato 10 teniendo ambos componentes de captura de audio (voz) y de presentación de audio, debería entenderse que en algunas realizaciones, el aparato 10 puede comprender solamente la parte de captura de audio (voz) del subsistema de audio de tal manera que en algunas realizaciones del aparato los micrófonos (para la captura de voz) están presentes.
- En algunas realizaciones, el aparato 10 comprende un procesador 21. El procesador 21 está acoplado al subsistema de audio y, específicamente, en algunos ejemplos, al convertidor analógico a digital 14 para recibir las señales digitales que representan las señales de audio del micrófono 11, y el convertidor digital a analógico (CDA) 12 configurado para emitir las señales de audio digitales procesadas. El procesador 21 puede configurarse para ejecutar diversos códigos de programa. Los códigos de programa implementados pueden comprender, por ejemplo, unas rutinas de grabación de audio y de procesamiento de señales de audio.
- En algunas realizaciones, el aparato comprende además una memoria 22. En algunas realizaciones, el procesador está acoplado a la memoria 22. La memoria puede ser cualquier medio de almacenamiento adecuado. En algunas realizaciones, la memoria 22 comprende una sección de código de programa 23 para almacenar códigos de programa que pueden implementarse en el procesador 21. Además, en algunas realizaciones, la memoria 22 puede comprender, además, una sección de datos almacenados 24 para almacenar datos, por ejemplo, datos que se han grabado o analizado de acuerdo con la aplicación. El código de programa implementado almacenado dentro de la sección de código de programa 23, y los datos almacenados dentro de la sección de datos almacenados 24 pueden recuperarse por el procesador 21 cuando sea necesario a través del acoplamiento de memoria - procesador.
- En algunas realizaciones adicionales, el aparato 10 puede comprender una interfaz de usuario 15. La interfaz de usuario 15 puede acoplarse en algunas realizaciones al procesador 21. En algunas realizaciones, el procesador puede controlar el funcionamiento de la interfaz de usuario y recibir entradas desde la interfaz de usuario 15. En algunas realizaciones, la interfaz de usuario 15 puede permitir a un usuario introducir comandos en el dispositivo o aparato electrónico 10, por ejemplo, a través de un teclado, y/o para obtener información del aparato 10, por ejemplo, a través de una pantalla que sea parte de la interfaz de usuario 15. La interfaz de usuario 15 puede en algunas realizaciones comprender una pantalla táctil o una interfaz táctil capaz tanto de permitir introducir información en el aparato 10 como de mostrar información adicional al usuario del aparato 10.
- En algunas realizaciones, el aparato comprende además un transceptor 13, el transceptor en tales realizaciones puede acoplarse al procesador y configurarse para permitir una comunicación con otros aparatos o dispositivos electrónicos, por ejemplo, a través de una red de comunicaciones inalámbricas. El transceptor 13 o cualquier transceptor o medio transmisor y/o receptor adecuados pueden configurarse en algunas realizaciones para comunicarse con otros dispositivos o aparatos electrónicos a través de un cable o de un acoplamiento cableado.
- El acoplamiento puede ser cualquier protocolo de comunicaciones conocido adecuado, por ejemplo, en algunas realizaciones, el transceptor 13 o los medios de transceptor pueden usar un protocolo de sistema universal de

telecomunicaciones móviles adecuado (UMTS) o GSM, un protocolo de red de área local inalámbrica (WLAN) como por ejemplo IEEE 802.X, un protocolo de comunicación de radio frecuencia de corto alcance adecuado, tal como Bluetooth, o una ruta de comunicación de datos por infrarrojos (IRDA).

- 5 Debería entenderse de nuevo que la estructura del dispositivo electrónico 10 podría complementarse y variarse de muchas maneras.

10 Tal como se describe en el presente documento el concepto de las realizaciones descritas en el presente documento es la capacidad de implementar un procesamiento de señal de audio robusto direccional/posicional usando al menos tres entradas de micrófono.

Con respecto a la figura 3, se muestra un aparato procesador de señales de audio de ejemplo de acuerdo con algunas realizaciones. Con respecto a la figura 4, se describe con más detalle el funcionamiento del aparato de procesamiento de señal de audio mostrado en la figura 3.

15 El aparato procesador de señal de audio en algunas realizaciones comprende un pre-procesador 201. El pre-procesador 201 puede configurarse para recibir las señales de audio de los micrófonos, mostrados en la figura 3 como los micrófonos cercanos 103, 105 y el micrófono lejano 101. La localización de los micrófonos cercanos y lejanos puede ser como se muestra en la configuración de ejemplo, como se muestra en la figura 2, sin embargo, debería entenderse que pueden usarse en algunas realizaciones estas otras configuraciones y/o el número de micrófonos.

25 Aunque las realizaciones que se describen en el presente documento presentan las señales de audio recibidas directamente de los micrófonos como las señales de entrada, debería entenderse que en algunas realizaciones las señales de audio de entrada pueden ser señales de audio pre-almacenadas o almacenadas. Por ejemplo, en algunas realizaciones las señales de audio de entrada son señales de audio recuperadas de la memoria. Estas señales de audio recuperadas pueden, en algunas realizaciones, ser señales de audio de micrófono grabadas.

30 El funcionamiento de recibir la entrada de audio/micrófono se muestra en la figura 4 por la etapa 301.

El preprocesador 201 puede configurarse en algunas realizaciones para realizar cualquier funcionamiento de pre-procesamiento adecuado. Por ejemplo, en algunas realizaciones, el pre-procesador puede configurarse para realizar una funcionamiento tal como: calibrar las señales de audio de micrófono; determinar si los micrófonos están libres de cualquier deterioro; corregir las señales de audio donde se determina el deterioro; determinar si alguno de los micrófonos están funcionando con viento fuerte; y determinar cuál de las entradas de micrófono es el micrófono principal. Por ejemplo, en algunas realizaciones los micrófonos pueden compararse para determinar cual tiene la señal de entrada más alta y por lo tanto se determina que es la dirigida hacia el usuario. En el ejemplo mostrado en el presente documento, se determina que el micrófono cercano 103 sea el micrófono principal y por lo tanto la salida del pre-procesador determina la salida del micrófono principal como la señal de audio de entrada del micrófono cercano 103.

El funcionamiento de pre-procesamiento, tal como una determinación de la entrada de micrófono principal se muestra en la figura 4 por la etapa 303.

- 45 En algunas realizaciones, la señal de audio de micrófono principal y otras señales de audio de micrófono cercanas determinadas pueden pasarse a continuación al conformador de haz 203.

50 En algunas realizaciones, el procesador de señales de audio comprende un conformador de haz 203. El conformador de haz 203 puede configurarse para recibir las entradas de micrófono cercano tal como se muestra en la figura 3 por el acoplamiento de micrófono principal (MAINM) y el otro acoplamiento de micrófono cercano del pre-procesador. A continuación, el conformador de haz 203 puede configurarse para generar al menos dos señales de audio de haz. Por ejemplo, como se muestra en la figura 3, el conformador de haz 203 puede configurarse para generar unas señales de audio de haz principal (MAINB) y de anti-haz (ANTIB).

55 El conformador de haz 203 puede configurarse para generar cualquier señal de audio conformada en haz adecuada desde el micrófono principal y otras entradas de micrófono cercanas. Como se describe en el presente documento en algunas realizaciones, la señal de audio de haz principal es una donde la voz local se pasa sustancialmente sin procesamiento, mientras que el ruido procedente de la dirección opuesta se atenúa sustancialmente, y la señal de audio anti-haz es una donde la voz local está fuertemente atenuada o sustancialmente atenuada, mientras que el ruido de las otras direcciones no se atenúa.

60 El conformador de haz 203 pueden configurarse en algunas realizaciones para emitir las señales de audio de haz, por ejemplo, las señales de audio de haz principal y de anti-haz, al cancelador de interferencias adaptativo (AIC) 205 y al detector de actividad de voz espacial 207.

65

En algunas realizaciones, el conformador de haz funciona en el dominio de tiempo y emplea filtros de respuesta de impulso finita (FIR) para atenuar algunas direcciones.

Debería entenderse que en las realizaciones con dos nearmics y un farmic hay en total cuatro filtros FIR. (Aunque debería entenderse que en algunas realizaciones podrían implementarse otros tipos de procesamiento). Por ejemplo, los cuatro filtros FIR pueden emplearse de la siguiente forma:

1. El haz principal emplea dos filtros FIR, un primer FIR para la primera señal de audio nearmic y un segundo FIR para la segunda señal de audio nearmic. Estas señales filtradas se combinan a continuación.

2. El anti-haz emplea otros dos filtros FIR, el tercer FIR para la primera señal de audio nearmic y un cuarto FIR para la segunda señal de audio nearmic. Estas señales filtradas se combinan a continuación.

3. Farmic: ningún procesamiento en el conformador de haz.

El funcionamiento de conformar el haz de las señales de audio de micrófono cercanas para generar unas señales de audio de haz principal y de anti-haz se muestra en la figura 4 por la etapa 305.

En algunas realizaciones, el procesador de audio comprende un cancelador de interferencias adaptativo (AIC) 205. El cancelador de interferencias adaptativo (AIC) 205, en algunas realizaciones, comprende al menos dos módulos de cancelador de interferencias de audio. Cada uno de los módulos de cancelador de audio está configurado para proporcionar una salida de procesamiento de audio adecuada para varias combinaciones de entradas de micrófonos.

En algunas realizaciones, el cancelador de interferencias de audio 205 comprende un módulo de cancelador de interferencias de audio (AIC) primario (o primero o principal) 211, un módulo AIC secundario (o segundo) 213 y un comparador 215 configurado para recibir las salidas del módulo AIC primario 211 y del módulo AIC secundario 213.

El módulo de cancelador de interferencias de audio primario 211 puede configurarse para recibir las señales de audio de las señales de audio de haz principal y de anti-haz y determinar una primera salida del módulo de cancelador de interferencias de audio usando el haz principal como una entrada de voz y de ruido y el anti-haz como referencia de ruido y una entrada de voz "filtrada". El módulo de cancelador de interferencias de audio primario 211 puede configurarse para a continuación pasar la salida del módulo procesado a un comparador 215.

El funcionamiento de determinar una primera salida de cancelación de interferencias adaptativa se muestra en la figura 4 por la etapa 307.

El módulo AIC secundario 213 está configurado para recibir como entradas la señal de audio de haz principal y la señal de audio de micrófono lejano (en otras palabras, la información de audio de los tres micrófonos). El módulo AIC secundario 213 puede configurarse para generar una salida de cancelación de interferencias adaptativa usando la señal de audio de haz principal como una entrada de voz y de ruido y la señal de audio de micrófono lejano como una referencia de ruido y una entrada de voz "filtrada". A continuación, el módulo de cancelador de interferencias de audio secundario 213 puede configurarse para emitir una salida de cancelación de interferencias adaptativa secundaria al comparador 215.

El funcionamiento de determinar una salida del módulo AIC secundario se muestra en la figura 4 por la etapa 309.

El cancelador de interferencias adaptativo 205 como se describe en el presente documento comprende, además, un comparador 215 configurado para recibir las salidas de los al menos dos módulos AIC. En la figura 3 estas salidas del módulo AIC son el módulo AIC primario 211 y el módulo AIC secundario 213, sin embargo, debería entenderse que en algunas realizaciones puede usarse cualquier número de módulos AIC y por lo tanto el comparador 215 recibe cualquier número de señales de módulo. A continuación, el comparador 215 puede configurarse para comparar las salidas de módulo AIC y la salida que tiene el nivel de señal de salida más alta.

En algunas realizaciones, el comparador 215 puede configurarse además para tener una salida preferida o por defecto y solo cambiar a una salida del módulo diferente, donde haya una diferencia considerable. Por ejemplo, el comparador 215 puede configurarse para determinar si la diferencia de nivel de señal entre dos módulos AIC es mayor que un valor umbral (por ejemplo, 2 dB) y solo cambiar cuando se pasa el valor umbral. Por ejemplo, en algunas realizaciones, el comparador 215 puede configurarse para emitir la salida del módulo AIC primario 211 mientras que la salida del módulo AIC primario sea igual o mayor que la salida del módulo AIC secundario y solo cambiar a la salida del módulo AIC secundario cuando la salida del módulo AIC secundario 213 sea 2 dB superior que la salida del módulo AIC primario.

El funcionamiento de comparar las salidas AIC primaria y secundaria y emitir la mayor se muestra en la figura 4 por la etapa 313.

El AIC 205 que, como se muestra en este ejemplo comprende dos módulos AIC paralelos, funciona en el dominio del tiempo empleando unos filtros adaptativos tal como se muestra en el presente documento en la figura 7. Sin embargo, cualquier implementación adecuada puede emplearse en algunas realizaciones, tales como las implementaciones AIC en serie o en serie paralelas híbridas.

5 En algunas realizaciones, el AIC 205 puede configurarse para recibir entradas de control. Estas entradas de control pueden usarse para controlar el comportamiento del AIC basándose en factores ambientales, tales como determinar si el micrófono está funcionando con viento (y por lo tanto al menos un micrófono está generando grandes cantidades de ruido del viento) o funcionando en una sombra de viento. Además, en algunas realizaciones, el
10 procesador de audio está configurado para optimizarse para el procesamiento de voz y por lo tanto se produce un proceso de detección de actividad de voz con el fin de que el cancelador de interferencias de audio funcione para optimizar la señal de voz para el ruido de fondo. Debería entenderse que en algunas realizaciones se normalicen las entradas a los módulos AIC.

15 En algunas realizaciones, la salida AIC puede pasarse a un supresor de ruido de canal único. Un supresor de ruido de canal único es un componente conocido que basándose en una estimación de ruido puede realizar una supresión de ruido adicional. El supresor de ruido único y el funcionamiento del supresor de ruido de canal único no se describen con más detalle en el presente documento, pero debería entenderse que el supresor de ruido de canal único recibe una entrada de una señal de voz ruidosa, y a partir de la señal de voz ruidosa estima el ruido de fondo.
20 La estimación del ruido de fondo se usa entonces para mejorar la señal de voz ruidosa, por ejemplo, aplicando un filtro de Weiner u otro método conocido. La estimación del ruido se hace desde la señal de voz ruidosa cuando se determina que la señal de voz ruidosa es solo ruido, por ejemplo, basándose en una salida de un detector de actividad de voz y/o, como se describe en el presente documento, un detector de actividad de voz espacial (VAD espacial). El supresor de ruido de canal único funciona normalmente dentro del dominio de frecuencia, sin embargo,
25 debería entenderse que en algunas realizaciones podría emplearse un supresor de ruido de canal único de dominio de tiempo.

El supresor de ruido de canal único puede usar, por lo tanto, la información VAD espacial para atenuar el ruido de fondo no estacionario, como un balbuceo, clics, radio, altavoces de la competencia, y los niños que tratan de llamar su atención durante las llamadas telefónicas.
30

Así, por ejemplo, el procesador de audio en algunas realizaciones puede comprender un detector de actividad de voz espacial 207. El detector de actividad de voz espacial 207 puede configurarse en algunas realizaciones para recibir como entradas las señales de audio de haz principal, el anti-haz, el micrófono principal y el micrófono lejano.
35 El funcionamiento del detector de actividad de voz espacial es forzar al supresor de ruido de canal único para actualizar solamente la estimación de ruido cuando la señal de audio comprende un ruido (o en otras palabras para no actualizar la estimación de ruido cuando la señal de audio comprende la voz desde la dirección esperada)

40 En algunas realizaciones, el detector de seguridad de voz espacial 207 comprende un normalizador 221. El normalizador 221 puede configurarse en algunas realizaciones para recibir las señales de audio del micrófono principal, el micrófono lejano, el haz principal y el anti-haz y realizar un proceso de normalización en estas señales de audio. El proceso de normalización se realiza de tal manera que los niveles de las señales de audio durante el ruido estacionario son sustancialmente los mismos. Este proceso de normalización se realiza con el fin de evitar cualquier desviación debida a las variaciones de sensibilidad del micrófono o a las variaciones de sensibilidad del haz.
45

En algunas realizaciones el normalizador está configurado para realizar una determinación mínima de señal suavizada en las señales de audio. En tales realizaciones, el normalizador puede entonces determinar una relación entre la mínima de las entradas para determinar un factor de ganancia de normalización que se aplicará a cada
50 entrada para normalizar el ruido estacionario. En algunas realizaciones el normalizador puede configurarse además para determinar el ruido estacionario espacial (por ejemplo, la carretera en un lado y el bosque en el otro lado del aparato) y en tales realizaciones adaptar la normalización a los niveles de ruido y evitar el marcado del ruido como voz. Similar o igual normalización puede realizarse para controlar los bloques de filtrado adaptativo en el AIC 205. Como tal, en algunas realizaciones un normalizador común puede emplearse tanto para el AIC (y por lo tanto, en
55 algunas realizaciones los módulos AIC) como para el VAD espacial de tal manera que los módulos AIC y el VAD espacial reciben entradas de las entradas de audio normalizadas.

En algunas realizaciones las señales de audio Nearmics se calibran antes de cualquier procesamiento, por ejemplo, la conformación del haz, (de manera que solo se permiten pequeñas diferencias en la sensibilidad del micrófono)
60 con el fin de tener haces adecuados que apunten a donde deberían hacerlo (en estos ejemplos hacia la boca del usuario y en la dirección opuesta).

Debería entenderse que el nivel de ruido en la señal de audio de haz principal es normalmente menor que en la señal de audio farmic, debido a que la conformación del haz reduce el ruido de fondo. Antes de comparar los niveles
65 de señal para el control interno del VAD espacial y del AIC, estas señales tienen que normalizarse. Esta normalización puede realizarse después de la conformación del haz.

Además, debería entenderse que, si bien los niveles de ruido en las señales de audio de haz principal y de anti-haz son los mismos para el ruido ambiental (por ejemplo, dentro de un automóvil), los niveles de ruido no deberían ser necesariamente los mismos para el ruido estacionario direccional (por ejemplo, cuando un usuario está de pie en un lado de una calle). Por lo tanto, en algunas realizaciones, las señales de audio de haz principal y de anti-haz tienen que normalizarse después de conformar el haz para el control interno del VAD espacial y del AIC.

Los niveles de ruido en las primeras señales de audio nearmic y farmic son en general aproximadamente los mismos, pero ya que estas señales no necesitan calibrarse contra las diferencias de sensibilidad del micrófono, en algunas realizaciones las primeras señales de audio nearmic y farmic se normalizan para el VAD espacial (no se usan en el AIC como un par de señales de entrada en los ejemplos mostrados en el presente documento).

El funcionamiento de normalizar las entradas se muestra en la figura 4 por la etapa 311.

En algunas realizaciones, el detector de actividad de voz espacial 207 comprende un filtro de frecuencia 223. El filtro de frecuencia 223 puede configurarse para recibir las entradas de señal de audio normalizadas y el filtro de frecuencia de las señales de audio. En algunas realizaciones, las señales de audio de micrófono y/o conformadas de haz tal como el micrófono principal, y las señales de audio de micrófono lejano se filtran a una frecuencia de paso bajo. En algunas realizaciones, a la comparación haz principal-'farmic' de las señales del micrófono (o señales de audio conformadas de haz) y también a la comparación micrófono principal (primer nearmic)-farmic (en otras palabras, la comparación de las señales del micrófono) puede implementarse un filtro de paso bajo con una banda de paso de por ejemplo aproximadamente 0-800 Hz. Las señales de audio de haz, por ejemplo, las señales de audio de haz principal y de anti-haz también se filtran a una frecuencia. El filtrado de frecuencia de las señales de audio de haz puede determinarse basándose en el diseño del haz del conformador de haz 203. Esto es debido a que los haces están diseñados de tal manera que la mayor separación está sobre un cierto intervalo de frecuencia. Un ejemplo de la banda de paso de frecuencia para la comparación de las señales de audio de haz principal y de anti-haz sería de aproximadamente 500 Hz a 2500 Hz. las señales de audio filtradas pueden pasarse, a continuación, a un comparador de relación 225.

El funcionamiento de filtrar las entradas para generar las bandas de frecuencia se muestra en la figura 4 por la etapa 315.

En algunas realizaciones, el detector de actividad de voz espacial 207 comprende un comparador de relación 225. El comparador de relación 225 puede configurarse para recibir la frecuencia de las señales de audio normalizadas filtradas y generar pares de comparación para determinar si las señales de audio comprenden una información de voz orientada espacialmente. En algunas realizaciones, los pares de comparación son:

- los niveles de señal de audio normalizada filtrada de haz principal y de anti- haz (por ejemplo, 500-2500 Hz)
- los niveles de señal de audio normalizada filtrada de micrófono y de micrófono cercano (por ejemplo, 0- 800 Hz)
- los niveles de señal de audio normalizada filtrada de haz principal y de micrófono lejano (por ejemplo, 0-800 Hz)

Cuando la comparación de los pares produce una relación que es mayor que un valor umbral determinado por cualquiera de las comparaciones, entonces no se determina que sea una actividad de voz significativa en una dirección espacial. En otras palabras, solo cuando el nivel de la señal es el mismo para los micrófonos y los haces se determina que las señales de audio son el ruido de fondo.

Una forma de voz de este tipo puede detectarse incluso cuando la colocación del aparato no es óptima.

El funcionamiento de la comparación de relación para determinar una marca de detección de actividad de voz espacial (para las actualizaciones de referencia de ruido) se muestra en la figura 4 por la etapa 317.

En algunas realizaciones, la salida de VAD espacial 207 puede emplearse como una entrada de control en un supresor de ruido de canal único como se ha tratado en el presente documento u otro supresor de ruido adecuado, de tal manera que cuando el VAD espacial 207 determina que cada una de las relaciones es similar o sustancialmente similar entonces, el supresor de ruido de canal único u otro supresor de ruido adecuado puede usar la estimación de ruido de fondo considerando que donde el nivel de señal difiere entre alguna de las comparaciones entonces, la estimación de ruido de fondo no se usa (y en algunas realizaciones se usa una estimación más antigua).

Con respecto a la figura 6, se muestra con más detalle un diagrama de flujo de ejemplo que muestra el funcionamiento del procesador de audio, y especialmente del AIC, basándose en las entradas de control como se ha descrito en el presente documento.

El AIC y, específicamente, en las realizaciones descritas en el presente documento determina si la salida AIC secundaria es más fuerte que la salida AIC primaria.

El funcionamiento de determinar si la salida AIC secundaria es más fuerte que la salida AIC primaria se muestra en la figura 6 por la etapa 503.

5 Cuando la salida de la AIC secundaria es más fuerte que la salida AIC primaria entonces, se determina una prueba adicional de que el sistema está funcionando con viento suave.

El funcionamiento de determinar si el sistema está funcionando con viento suave se muestra en la figura 6 por la etapa 507.

10 Cuando el sistema no está funcionando con viento suave entonces, se usa el funcionamiento de procesamiento de tres micrófonos, en otras palabras el AIC secundario se emite por el comparador.

El funcionamiento de usar la salida de procesamiento del AIC secundario (tres micrófonos) se muestra en la figura 6 por la etapa 509.

15 Cuando el sistema está funcionando con viento suave o la salida AIC secundaria no es más fuerte que la salida AIC primaria entonces se usa la salida AIC primaria.

20 El uso de la salida AIC primaria se muestra en la figura 6 por la etapa 511.

Además, con respecto a la figura 7 se usa un ejemplo AIC en el que un primer micrófono o haz para la referencia de ruido y la voz filtrada se pasa como una entrada positiva a un primer sumador 601. El primer sumador 601 emite a un primer filtro adaptativo 603 la entrada de control y a un segundo filtro adaptativo 605 la entrada de datos. El primer sumador 601 recibe además como una entrada negativa la salida del primer filtro adaptativo 603. El primer
25 filtro adaptativo 603 recibe como una entrada de datos, la señal de audio de micrófono o de haz de voz y ruido. La señal de audio de micrófono o de haz de voz y ruido se pasa además a un retardo 607. La salida del retardo 607 se pasa como entrada positiva a un segundo sumador 609. El segundo sumador 609 recibe como una entrada negativa la salida del segundo filtro adaptativo 605. La salida del segundo sumador 609 se emite a continuación como la salida de señal y se usa como la entrada de control para el segundo filtro adaptativo 605.

30 De esta manera el filtrado de Wiener funciona como un método de supresión que puede realizarse a la señal de audio de canal único $s(k)$. Aunque el ejemplo mostrado en la figura 7 aparecería para permitir que el AIC elimine todo el ruido, esto no se logra en situaciones prácticas ya que normalmente hay un ruido de fondo de salida que se reduce aún más en algunas realizaciones por el supresor de ruido de canal único.

35 En otras palabras la figura 7 muestra un módulo AIC de ejemplo que comprende dos filtros adaptativos: un AF de reducción de voz (configurado para reducir la voz filtrada de la entrada secundaria = ruido + voz filtrada) y un AF de reducción de ruido (configurado para reducir el ruido de la entrada primaria = voz + ruido). Aunque en esta realización mostrada hay una estructura de filtrado adaptativo doble configurada para proporcionar una mejor robustez de colocación reduciendo la voz filtrada de la entrada secundaria antes de que se use en el AF de
40 reducción de ruido como una referencia de ruido, debería entenderse que puede aplicarse cualquier filtro y filtrado adecuado.

Se apreciará que el dispositivo electrónico 10 puede ser cualquier dispositivo que incorpore un sistema de grabación de audio, por ejemplo, un tipo de equipo de usuario inalámbrico, tales como los teléfonos móviles, los dispositivos de procesamiento de datos portables o los navegadores de web portables, así como los dispositivos llevables.

En general, las diversas realizaciones de la invención pueden implementarse en hardware o circuitos de fin especial, software, lógica o cualquier combinación de los mismos. Por ejemplo, algunos aspectos pueden implementarse en hardware, mientras que otros aspectos pueden implementarse en firmware o software que puede ejecutarse por un controlador, un microprocesador u otro dispositivo informático, aunque la invención no está limitada a los mismos. Aunque diversos aspectos de la invención pueden ilustrarse y describirse como diagramas de bloques, diagramas de flujo, o usando alguna otra representación gráfica, se entiende bien que estos bloques, aparatos, sistemas, técnicas o métodos descritos en el presente documento pueden implementarse en, como ejemplos no limitativos,
55 hardware, software, firmware, circuitos de fin especial o lógica, hardware de fin general o un controlador u otros dispositivos informáticos, o alguna combinación de los mismos.

Las realizaciones de esta invención pueden implementarse por software informático ejecutable por un procesador de datos del dispositivo móvil, tal como en la entidad de procesador, o por hardware, o por una combinación de software y hardware. Además, a este respecto, debería observarse que cualquiera de los bloques del flujo lógico como en las figuras pueden representar etapas de programa, o circuitos lógicos interconectados, bloques y funciones, o una combinación de etapas de programa y circuitos lógicos, bloques y funciones. El software puede almacenarse en tales medios físicos como chips de memoria, o bloques de memoria implementados dentro del procesador, medios magnéticos tales como un disco duro o disquetes, y medios ópticos tales como por ejemplo un
60 DVD y los variantes de datos de los mismos, un CD.

- La memoria puede ser de cualquier tipo adecuado para el entorno técnico local y puede implementarse usando cualquier tecnología de almacenamiento de datos adecuada, tal como los dispositivos de memoria basados en semiconductores, dispositivos y sistemas de memoria magnética, dispositivos y sistemas de memoria óptica, memoria fija y memoria extraíble. Los procesadores de datos pueden ser de cualquier tipo adecuado para el entorno técnico local, y pueden incluir uno o más de entre ordenadores de fin general, ordenadores de fin especial, microprocesadores, procesadores de señal digital (DSP), circuitos integrados de aplicación específica (ASIC), circuitos de nivel de puertas y procesadores basados en la arquitectura de procesadores de múltiples núcleos, como ejemplos no limitativos.
- 10 Las realizaciones de las invenciones pueden ponerse en práctica en diversos componentes, tales como unos módulos de circuitos integrados. El diseño de los circuitos integrados es por lo general un proceso altamente automatizado. Unas herramientas de software complejas y potentes están disponibles para convertir un diseño de nivel lógico en un diseño de circuito de semiconductores listo para grabarse y formarse sobre un sustrato semiconductor.
- 15 Los programas, tales como los proporcionados por Synopsys, Inc., de Mountain View, California, y Cadence Design, de San José, California encaminan los conductores y localizan los componentes automáticamente en un chip semiconductor usando reglas bien establecidas de diseño, así como las bibliotecas de los módulos de diseño pre-almacenadas. Una vez que el diseño de un circuito semiconductor se ha completado, el diseño resultante, en un
- 20 formato electrónico normalizado (por ejemplo, Opus, GDSII, o similares) pueden transmitirse a una instalación de fabricación de semiconductores o "fab" para su fabricación.
- La descripción anterior ha proporcionado, por medio de unos ejemplos no limitativos y a modo de ejemplo, una descripción completa e informativa de la realización a modo de ejemplo de esta invención. Sin embargo, diversas modificaciones y adaptaciones pueden ser evidentes para los expertos en la materia a la vista de la descripción anterior, cuando se lee junto con los dibujos adjuntos y las reivindicaciones adjuntas. Sin embargo, todas estas y similares modificaciones de las enseñanzas de esta invención pueden aún estar dentro del alcance de esta invención como se define en las reivindicaciones adjuntas.
- 25

REIVINDICACIONES

1. Un método que comprende:

- 5 recibir al menos tres señales de audio de micrófono, comprendiendo las al menos tres señales de audio de micrófono al menos dos señales de audio de micrófono cercanas generadas por al menos dos micrófonos cercanos localizados cerca de una fuente de audio deseada y al menos una señal de audio de micrófono lejana generada por un micrófono lejano localizado más allá de la fuente de audio deseada que los al menos dos micrófonos cercanos;
- 10 generar una primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono, siendo la primera selección a partir de las señales de audio de micrófono cercanas;
- generar al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono, la al menos una selección adicional de las al menos tres señales de audio de micrófono, siendo la segunda selección de todas las señales de audio de micrófono;
- 15 determinar a partir de la primera señal de audio procesada y de la al menos una señal de audio procesada adicional, la señal de audio con una mayor supresión de ruido.

2. El método de acuerdo con la reivindicación 1, en el que recibir las al menos tres señales de audio de micrófono comprende:

- 20 recibir una primera señal de audio de micrófono de un primer micrófono cercano localizado sustancialmente en una parte delantera de un aparato;
- recibir una segunda señal de audio de micrófono de un segundo micrófono cercano localizado sustancialmente en una parte trasera del aparato; y
- 25 recibir una tercera señal de audio de micrófono de un micrófono lejano localizado sustancialmente en el extremo opuesto de los micrófonos primero y segundo.

3. El método de acuerdo con la reivindicación 2, en el que generar la primera señal de audio procesada basándose en una primera selección de las al menos tres señales de audio de micrófono comprende generar la primera señal de audio procesada basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y una señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda.

- 30 4. El método de acuerdo con la reivindicación 3, en el que generar la al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono comprende generar la señal de audio procesada adicional basándose en una señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono.

5. El método de acuerdo con cualquiera de las reivindicaciones 3 o 4, que comprende además:

- 40 generar la señal de audio de haz principal: aplicando un primer filtro de respuesta de impulso finita a la primera señal de audio de micrófono; aplicando un segundo filtro de respuesta de impulso finita a la segunda señal de audio de micrófono; y combinando una salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio de haz principal; y
- 45 generar la señal de audio anti-haz: aplicando un tercer filtro de respuesta de impulso finita a la primera señal de audio de micrófono; aplicando un cuarto filtro de respuesta de impulso finita a la segunda señal de audio de micrófono; y combinando una salida del tercer filtro de respuesta de impulso finita y del cuarto filtro de respuesta de impulso finita para generar la señal de audio anti-haz.

6. El método de acuerdo con cualquiera de las reivindicaciones 4 o 5, en el que generar la señal de audio procesada adicional basándose en la señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la tercera señal de audio de micrófono comprende filtrar la señal de audio de haz principal basándose en la tercera señal de audio de micrófono.

- 50 7. El método de acuerdo con cualquiera de las reivindicaciones 3 a 5, en el que generar la primera señal de audio procesada basándose en la señal de audio de haz principal basada en las señales de audio de micrófono primera y segunda y la señal de audio anti-haz basada en las señales de audio de micrófono primera y segunda comprende filtrar la señal de audio de haz principal basándose en la señal de audio anti-haz.

8. El método de acuerdo con la reivindicación 1, en el que generar la primera señal de audio procesada basándose en la primera selección de las al menos tres señales de audio de micrófono comprende:

- 60 seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas;
- 65 seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en las al menos tres señales de audio de

micrófono, siendo las selecciones a partir de las señales de audio de micrófono cercanas;
filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la primera señal de audio procesada.

5 9. El método de acuerdo con la reivindicación 1, en el que generar la al menos una señal de audio procesada adicional basándose en al menos una selección adicional de las al menos tres señales de audio de micrófono comprende:

10 seleccionar como una primera entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y una señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono;
seleccionar como una segunda entrada de procesamiento al menos una de: una de las al menos tres señales de audio de micrófono; y la señal de audio conformada en haz basada en al menos dos de las al menos tres señales de audio de micrófono, siendo las selecciones a partir de todas las señales de micrófono;
15 filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional.

10. El método de acuerdo con cualquiera de las reivindicaciones 8 o 9, en el que filtrar la primera entrada de procesamiento basándose en la segunda entrada de procesamiento para generar la al menos una señal de audio procesada adicional comprende al menos uno de:

20 filtrar una supresión de ruido de la primera entrada de procesamiento basándose en la segunda entrada de procesamiento; y
conformar el haz de al menos dos de las al menos tres señales de audio de micrófono para generar una señal de audio conformada en haz.
25

11. El método de acuerdo con la reivindicación 10, en el que cuando se genera la señal de audio conformada en haz, el método comprende además:

30 aplicar un primer filtro de respuesta de impulso finita a una primera de las al menos dos de las al menos tres señales de audio de micrófono;
aplicar un segundo filtro de respuesta de impulso finita a una segunda de las al menos dos de las al menos tres señales de audio de micrófono; y
combinar la salida del primer filtro de respuesta de impulso finita y del segundo filtro de respuesta de impulso finita para generar la señal de audio conformada en haz.
35

12. El método de acuerdo con cualquier reivindicación anterior, que comprende además la supresión de ruido de canal único de la señal de audio con una mayor supresión de ruido, en donde la supresión de ruido de canal único comprende:

40 generar un indicador que muestra si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido;
estimar y actualizar un ruido de fondo de la señal de audio cuando el indicador muestra que el periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido;
45 procesar la señal de audio basándose en la estimación de ruido de fondo para generar una señal de audio de ruido suprimido.

13. El método de acuerdo con la reivindicación 12, en el que generar el indicador que muestra si un periodo de la señal de audio comprende una falta de componentes de voz o es significativamente un ruido comprende:

50 normalizar una selección de las al menos tres señales de audio de micrófono, en donde la selección comprende: unas señales de audio conformadas de haz de al menos dos de las al menos tres señales de audio de micrófono; y unas señales de audio de micrófono;
filtrar las selecciones normalizadas de las al menos tres señales de audio de micrófono;
55 comparar las selecciones normalizadas filtradas para determinar una relación de diferencia de potencia;
generar el indicador que muestre un periodo de la señal de audio que comprende una falta de componentes de voz o es significativamente un ruido donde al menos una comparación de las selecciones normalizadas filtradas tiene una relación de diferencia de potencia mayor que un umbral determinado.

60 14. El método de acuerdo con cualquier reivindicación anterior, en el que determinar, a partir de la primera señal de audio procesada y la al menos una señal de audio procesada adicional la señal de audio con una mayor supresión de ruido, comprende al menos uno de:

65 determinar, a partir de la primera señal de audio procesada y la al menos una señal de audio procesada adicional, la señal de audio con una salida de nivel de señal más alta; y
determinar, a partir de la primera señal de audio procesada y la al menos una señal de audio procesada

adicional, la señal de audio con una salida de nivel de potencia más alta.

15. Un aparato configurado para realizar las acciones del método de cualquiera de las reivindicaciones 1 a 14.

Figura 1

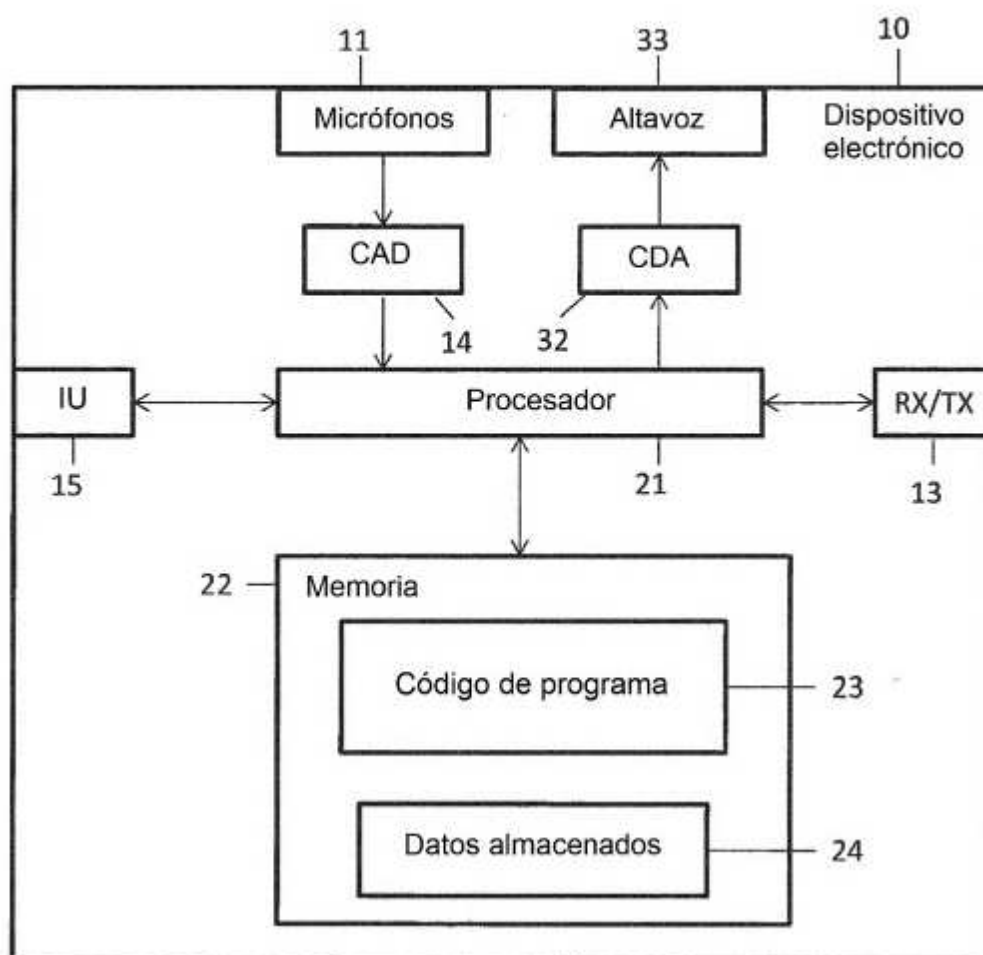


Figura 2

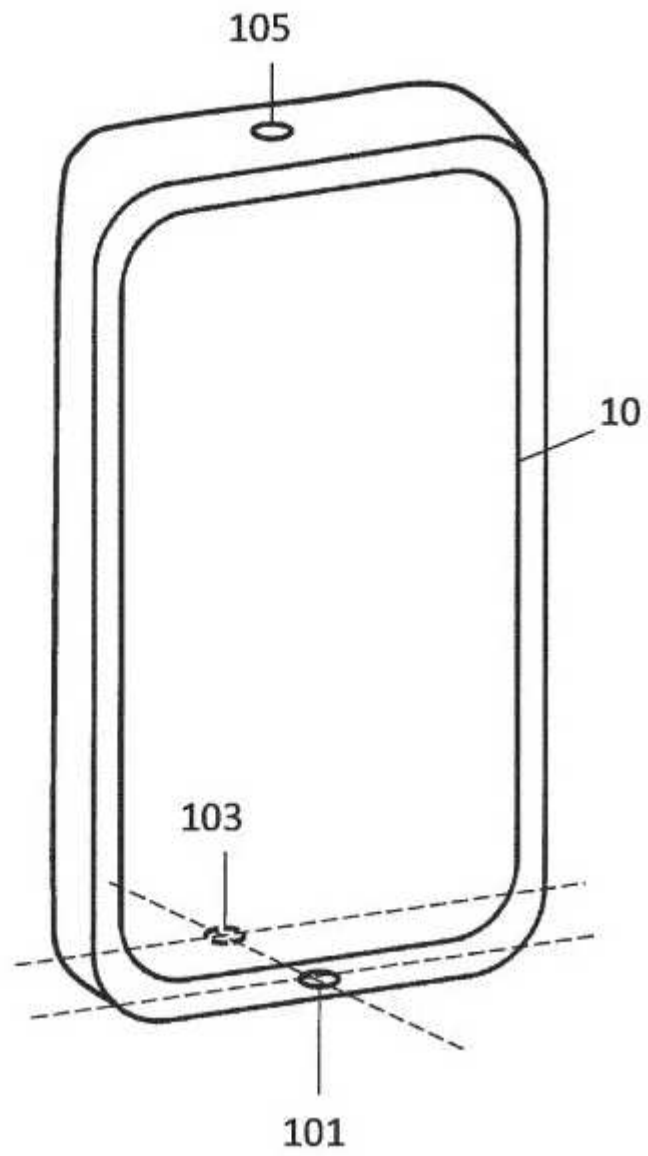


Figura 3

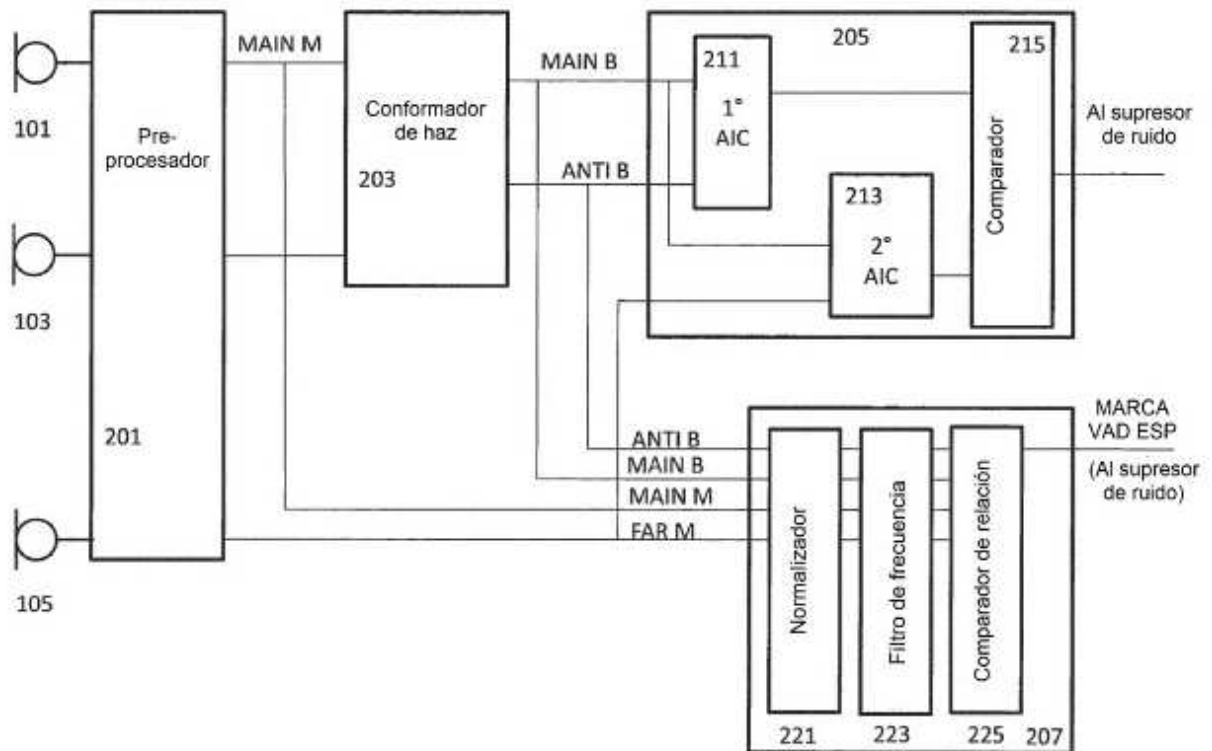


Figura 4

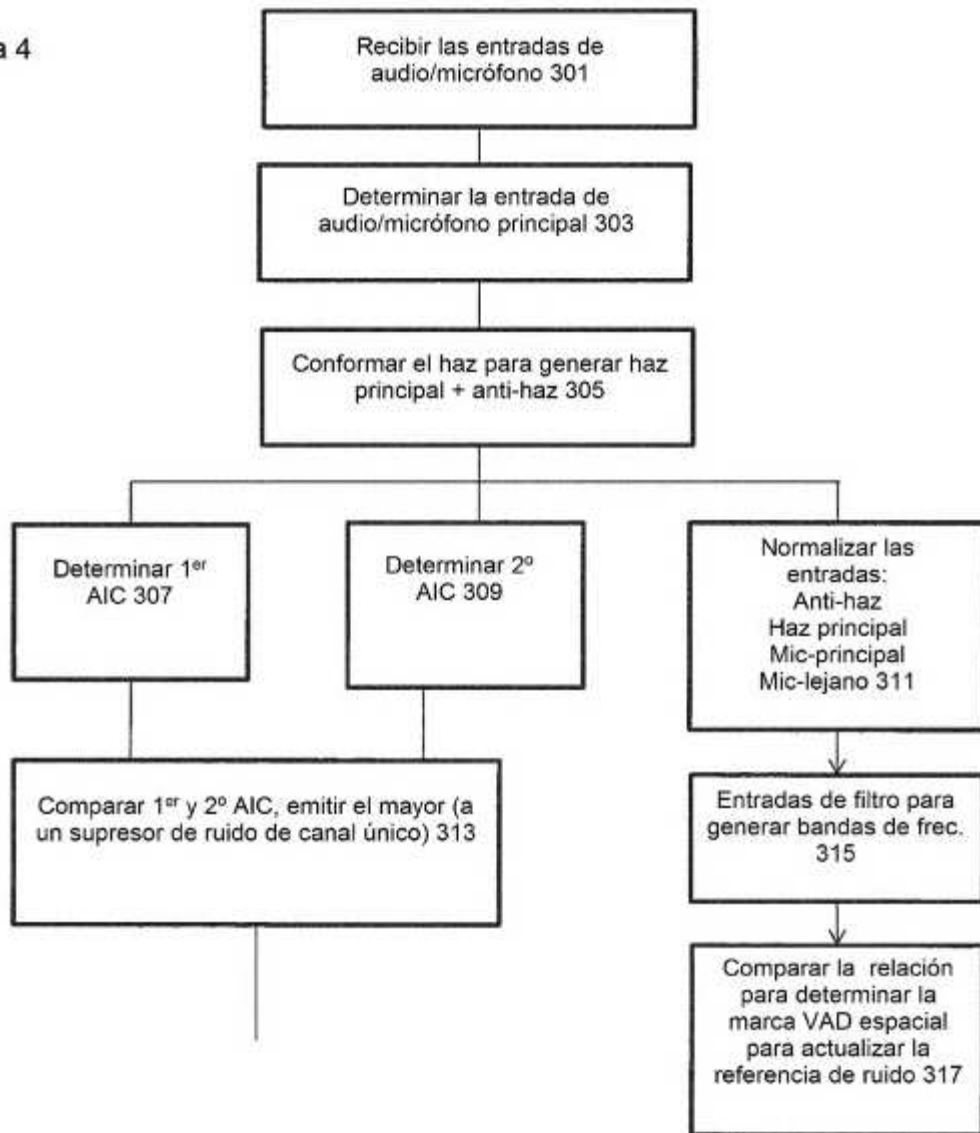


Figura 5

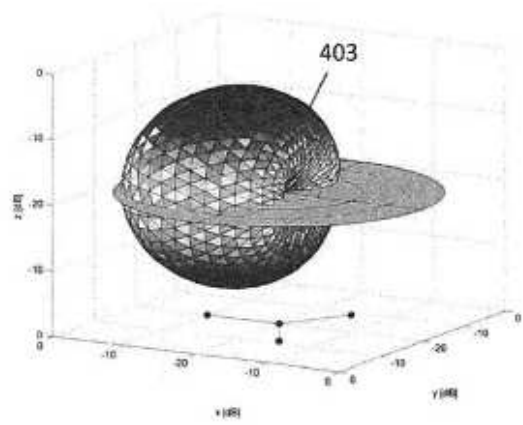
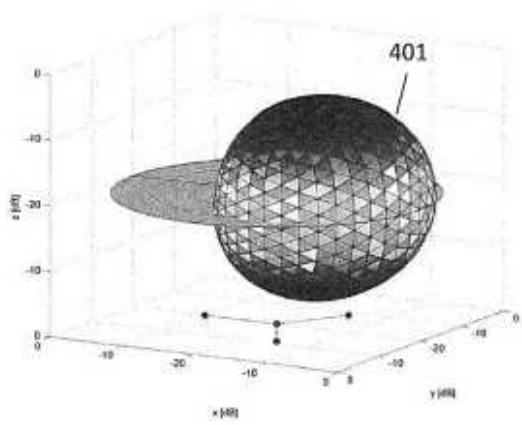


Figura 6

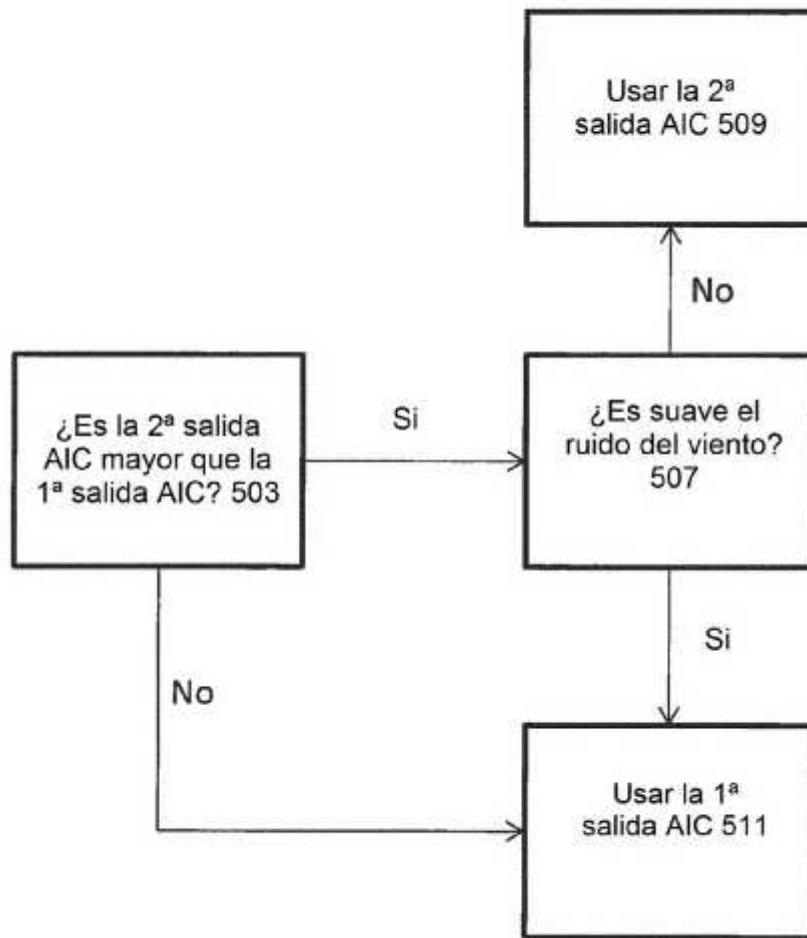


Figura 7

