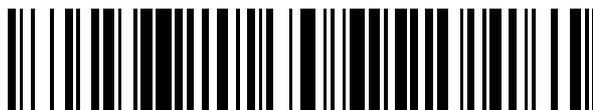


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 670 870**

51 Int. Cl.:

H04R 3/00 (2006.01)

G10L 21/02 (2013.01)

H04R 1/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.12.2011 PCT/JP2011/079978**

87 Fecha y número de publicación internacional: **28.06.2012 WO12086834**

96 Fecha de presentación y número de la solicitud europea: **19.12.2011 E 11852100 (4)**

97 Fecha y número de publicación de la concesión europea: **14.03.2018 EP 2642768**

54 Título: **Método de realce de sonido, dispositivo, programa y medio de grabación**

30 Prioridad:

21.12.2010 JP 2010285175

21.12.2010 JP 2010285181

09.02.2011 JP 2011025784

01.09.2011 JP 2011190768

01.09.2011 JP 2011190807

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

01.06.2018

73 Titular/es:

**NIPPON TELEGRAPH AND TELEPHONE CORPORATION (100.0%)
3-1 Otemachi 2-chome
Chiyoda-ku , Tokyo 100-8116, JP**

72 Inventor/es:

**NIWA, KENTA;
SAKAUCHI, SUMITAKA;
FURUYA, KENICHI y
HANEDA, YOICHI**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 670 870 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método de realce de sonido, dispositivo, programa y medio de grabación

Campo técnico

5 La presente invención se refiere a una técnica capaz de realzar sonidos en un intervalo estrecho deseado (técnica de realce de sonido).

Antecedentes de la técnica

10 Cuando un dispositivo de filmación de películas (cámara de video o videocámara), por ejemplo, equipado con un micrófono se acerca a un sujeto para filmar al sujeto, es preferible para la grabación de video que solamente los sonidos de alrededor del sujeto se deberían realzar en sincronización con el acercamiento de la filmación. Se han estudiado y desarrollado técnicas (técnicas de realce de sonido directivo agudo) para realzar sonidos en un intervalo estrecho que incluye una dirección deseada (una dirección del objetivo). La sensibilidad de un micrófono pertinente a las direcciones alrededor del micrófono se denomina directividad. Cuando la directividad en una dirección particular es aguda, se realzan los sonidos que llegan desde un intervalo estrecho incluyendo la dirección particular y se suprimen los sonidos fuera del intervalo. Aquí se describirán en primer lugar tres técnicas convencionales relacionadas con la técnica de realce de sonido directivo agudo. El término "sonido o sonidos" como se usa en la presente memoria no se limita a la voz humana, sino que se refiere a "sonido o sonidos" en general, tales como música y ruido ambiental, así como llamadas de animales y voz humana.

[1] Técnica de realce de sonido directivo agudo que usa propiedades físicas

20 Ejemplos típicos de esta categoría incluyen micrófonos de cañón y micrófonos parabólicos. El principio de un micrófono 900 de tubo acústico se describirá primero con referencia a la Fig. 1. El micrófono 900 de tubo acústico usa una interferencia de sonido para realzar sonidos que llegan desde una dirección del objetivo. La Fig. 1A ilustra el realce de sonidos que llegan desde una dirección del objetivo mediante el micrófono 900 de tubo acústico. La abertura del tubo 901 acústico del micrófono 900 de tubo acústico se apunta a la dirección del objetivo. Los sonidos que llegan desde la parte delantera (dirección del objetivo) de la abertura del tubo 901 acústico viajan directamente a través del interior del tubo 901 acústico y alcanzan un micrófono 902 del micrófono 900 de tubo acústico con baja pérdida de energía. Por otra parte, los sonidos que llegan desde direcciones distintas de la dirección del objetivo entran en el tubo 901 a través de muchas rendijas 903 provistas en los lados del tubo como se ilustra en la Fig. 1B. Los sonidos que se introducen a través de las rendijas 903 interfieren unos con otros, lo que reduce los niveles de presión de sonido de los sonidos que vinieron de las direcciones distintas a la dirección del objetivo y alcanzan el micrófono 902.

35 El principio de un micrófono 910 parabólico se describirá a continuación con referencia a la Fig. 2. El micrófono 910 parabólico usa la reflexión de sonidos para realzar los sonidos que llegan desde una dirección del objetivo. La Fig. 2A es un diagrama que ilustra el realce de sonidos que llegan desde la dirección del objetivo por el micrófono 910 parabólico. Un reflector 911 parabólico (superficie parabólica) del micrófono 910 parabólico se apunta a la dirección del objetivo de modo que la línea que enlaza entre el vértice del reflector 911 parabólico y el punto focal del reflector 911 parabólico coincide con la dirección del objetivo. Los sonidos que llegan desde la dirección del objetivo se reflejan por el reflector 911 parabólico y se enfocan en el punto focal. Por consiguiente, un micrófono 912 colocado en el punto focal puede realzar y captar señales de sonido incluso con poca energía. Por otra parte, los sonidos que llegan desde las direcciones distintas a la dirección del objetivo y se reflejan por el reflector 911 parabólico no se enfocan en el punto focal, como se ilustra en la Fig. 2B. Por consiguiente, se disminuyen los niveles de presión de sonido de los sonidos que vienen desde la dirección distinta a la dirección del objetivo y que llegaron al micrófono 912.

[2] Técnica de realce de sonido directivo agudo que usa procesamiento de señal

45 Ejemplos típicos de esta categoría incluyen agrupaciones de micrófonos en fase (véase la bibliografía no de patente 1). La Fig. 3 es un diagrama que ilustra que una agrupación de micrófonos en fase que incluye múltiples micrófonos se usa para realzar sonidos de una dirección del objetivo y suprime los sonidos de las otras direcciones distintas de la dirección del objetivo. La agrupación de micrófonos en fase realiza procesamiento de señal para aplicar un filtro que incluye información acerca de las diferencias de fase y/o de amplitud entre los micrófonos a las señales captadas con los micrófonos y superpone las señales resultantes para realzar los sonidos de la dirección del objetivo. A diferencia del micrófono de tubo acústico y el micrófono parabólico descritos en la categoría [1], la agrupación de micrófonos en fase puede realzar los sonidos que llegan desde cualquier dirección debido a que realza los sonidos mediante el procesamiento de señal.

[3] Técnica de realce de sonido directivo agudo mediante captación selectiva de sonidos reflejados

55 Ejemplos típicos de esta categoría incluyen conformación de haces múltiples (véase la bibliografía no de patente 2). La conformación de haces múltiples es una técnica de realce de sonido directivo agudo que recoge sonidos individuales, incluyendo sonidos directos y sonidos reflejados, juntos para captar los sonidos que llegan de una

dirección del objetivo con una relación señal a ruido alta y se ha estudiado más intensivamente en el campo de lo inalámbrico más que en el de la acústica.

El procesamiento de la conformación de haces múltiples en un dominio de frecuencia se describirá a continuación. Los símbolos se definirán antes de la descripción. El índice de una frecuencia se denota por ω y el índice de un número de tiempo de trama se denota por k . Representaciones en el dominio de frecuencia de señales analógicas recibidas en M micrófonos se denotan por $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$, la dirección desde la cual un sonido directo de una fuente de sonido ubicada en una dirección θ_s a ser realizada se denota por θ_{s1} , las direcciones desde las cuales llegan sonidos reflejados se denotan por $\theta_{s2}, \dots, \theta_{sR}$. Aquí, T representa la transposición y $R - 1$ es el número total de sonidos reflejados. Un filtro que realiza un sonido desde una dirección θ_{sr} se denota por $W^{\rightarrow}(\omega, \theta_{sr})$. Aquí, r es un número entero que satisface $1 \leq r \leq R$.

Una precondition para la conformación de haces múltiples es que se conozcan las direcciones a partir de las cuales llegan los sonidos directos y reflejados y sus tiempos de llegada. Es decir, el número de objetos, tales como paredes, suelos, reflectores, que obviamente se espera que reflejen los sonidos es igual a $R - 1$. El número de sonidos reflejados, $R - 1$, a menudo se establece en un valor relativamente pequeño, tal como 3 o 4. Esto se basa en el hecho de que hay una alta correlación entre un sonido directo y un sonido reflejado de bajo orden. Dado que la conformación de haces múltiples realiza los sonidos individualmente y añade sincronamente las señales realizadas, una señal de salida $Y(\omega, k, \theta_s)$ se puede dar por la ecuación (1). Aquí, H representa la transposición hermitiana.

$$Y(\omega, k, \theta_s) = \sum_{r=1}^R \vec{W}^H(\omega, \theta_{sr}) \vec{X}(\omega, k) \quad (1)$$

La conformación de haces de retardo y suma se describirá como método para diseñar un filtro $W^{\rightarrow}(\omega, \theta_{sr})$. Suponiendo que los sonidos directos y reflejados llegan como ondas planas, entonces el filtro $W^{\rightarrow}(\omega, \theta_{sr})$ se puede dar por la ecuación (2).

$$\vec{W}(\omega, \theta_{sr}) = \frac{\vec{h}(\omega, \theta_{sr})}{\vec{h}^H(\omega, \theta_{sr}) \vec{h}(\omega, \theta_{sr})} \quad (2)$$

donde, $\vec{h}^{\rightarrow}(\omega, \theta_{sr}) = [h_1(\omega, \theta_{sr}), \dots, h_M(\omega, \theta_{sr})]^T$ es un vector de propagación de un sonido que llega de una dirección θ_{sr} .

Suponiendo que ondas planas llegan a una agrupación de micrófonos lineales (una agrupación de micrófonos en la que M micrófonos están dispuestos linealmente), entonces los elementos $h_m(\omega, \theta_{sr})$ que componen $\vec{h}^{\rightarrow}(\omega, \theta_{sr})$ se pueden dar por la ecuación (3).

$$h_m(\omega, \theta_{sr}) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos\theta_{sr}\right] \cdot \exp[-j\omega\tau(\theta_{sr})] \quad (3)$$

donde m es un número entero que satisface $1 \leq m \leq M$, c es la velocidad del sonido, u representa la distancia entre micrófonos adyacentes, j es una unidad imaginaria, y $\tau(\theta_{sr})$ representa un retardo de tiempo entre un sonido directo y un sonido reflejado que llega desde la dirección θ_{sr} .

Por último, una señal de salida $Y(\omega, k, \theta_s)$ se transforma a un dominio de tiempo para obtener una señal en la que se realiza un sonido de la fuente de sonido situada en la dirección del objetivo θ_s .

La Fig. 4 ilustra una configuración funcional de la técnica de realce de sonido directivo agudo que usa la conformación de haces múltiples.

Paso 1

Un convertidor 110 AD convierte señales analógicas emitidas desde M micrófonos 100-1, ..., 100- M en señales digitales $x^{\rightarrow}(t) = [x_1(t), \dots, x_M(t)]^T$. Aquí, t representa el índice de un tiempo discreto.

Paso 2

Una sección 120 de transformada en el dominio de frecuencia transforma la señal digital de cada canal en una señal en el dominio de frecuencia mediante un método tal como transformada de Fourier discreta rápida. Por ejemplo, para el micrófono de orden m ($1 \leq m \leq M$), las señales $x_m((k-1)N+1), \dots, x_m(kN)$ en N puntos de muestreo se almacenan en un almacenador temporal. Aquí, N es aproximadamente 512 en el caso de muestreo a 16 KHz. La

transformada de Fourier discreta rápida de las señales analógicas de M canales almacenados en el almacenador temporal se realiza para obtener señales en el dominio de frecuencia $X^T(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$.

Paso 3

5 Cada una de las secciones de filtrado de realce 130-r ($1 \leq r \leq R$) aplica un filtro $W^H(\omega, \theta_{sr})$ para una dirección θ_{sr} a las señales en el dominio de frecuencia $X^T(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y emite una señal $Z_r(\omega, k)$ en la que se realiza un sonido de la dirección θ_{sr} . Es decir, cada sección de filtrado de realce 130-r ($1 \leq r \leq R$) realiza un procesamiento dado por la ecuación (4):

$$Z_r(\omega, k) = \vec{W}^H(\omega, \theta_{sr}) \vec{X}(\omega, k) \quad (4)$$

10 Un sumador 140 toma entradas de las señales $Z_1(\omega, k), \dots, Z_R(\omega, k)$ y emite una señal de suma $Y(\omega, k)$. La suma se puede dar por la ecuación (5):

$$Y(\omega, k) = \sum_{r=1}^R Z_r(\omega, k) \quad (5)$$

Paso 5

Una sección 150 de transformada en el dominio de tiempo transforma la señal suma $Y(\omega, k)$ a un dominio de tiempo y emite una señal $y(t)$ en el dominio de tiempo en la que se realiza el sonido de la dirección θ_s .

15 En algunas situaciones, por ejemplo en una situación donde hay múltiples fuentes de sonido en aproximadamente la misma dirección a diferentes distancias de un micrófono, se puede desear que los sonidos que llegan de las fuentes de sonido sean realizados selectivamente mediante la técnica de realce de sonido directivo agudo. Consideremos una situación donde un dispositivo de filmación de películas equipado con un micrófono se acerca a un sujeto para filmar al sujeto como en el ejemplo descrito anteriormente. Si hay una fuente de sonido (conocida como la "fuente de sonido trasera") en la parte trasera del sujeto enfocado (conocido como la "fuente de sonido enfocado") en el intervalo de directividad del micrófono, un sonido de la fuente de sonido enfocado y un sonido de la fuente de sonido trasera se mezclan y se realzan, dando a los espectadores una experiencia de escucha no natural. Por lo tanto, se desea una técnica capaz de realzar sonidos en un intervalo estrecho que incluye una dirección deseada según las distancias desde un micrófono (una técnica de realce de punto de sonido). Se describirán a modo de ilustración tres técnicas convencionales con relación a la técnica de realce de punto de sonido.

20

25

(1) La técnica descrita en la bibliografía no de patente 3 es un método de diseño óptimo para una agrupación de retardo y suma en un campo cercano de sonido donde las ondas de sonido son esféricas. La agrupación está diseñada de modo que se maximiza la relación SN entre una señal de objetivo de una posición de fuente de sonido y sonidos indeseados (ruido de fondo y reverberación).

30 (2) La técnica descrita en la bibliografía no de patente 4 requiere dos agrupaciones pequeñas de micrófonos y permite una captura de sonido de punto según las distancias sin necesitar una agrupación grande de micrófonos.

(3) La técnica descrita en la bibliografía no de patente 5 distingue entre distancias a una fuente de sonido con una única agrupación de micrófonos y realza o suprime sonidos solamente de la fuente de sonido en un intervalo de distancia particular, eliminando por ello el ruido de interferencia. Esta técnica tiene la ventaja del hecho de que la potencia de un sonido que llega directamente desde una fuente de sonido y la potencia de un sonido entrante reflejado varían según las distancias para realzar sonidos según las distancias de las fuentes de sonido.

35

Además se hace referencia a la bibliografía no de patente 6 que investiga el efecto de la reflexión de sala en una separación de fuente ciega. Se demuestra que se puede reducir la reflexión de orden más alto usando el método de subespacio. Se demuestra además que la reflexión de orden más bajo tiene poco efecto en el rendimiento de separación. La bibliografía no de patente 6 falla al describir que el filtro que realza los sonidos se obtiene antes de captar los sonidos a ser realzados.

40

Lista de referencias

Bibliografía no de patente

45 Bibliografía no de patente 1: O. L. Frost, "An algorithm for linearly constrained adaptative array processing", Actas del IEEE, vol. 60, páginas 926 - 935, 1972.

Bibliografía no de patente 2: J. L. Flanagan, A. C. Surendran, E. E. Jan, "Spatially selective sound capture for speech and audio processing", Speech Communication, Volumen 13, Número 1-2, páginas 207 - 222, octubre de 1993.

Bibliografía no de patente 3: Hiroaki Nomura, Yutaka Kaneda, Junji Kojima, "Microphone array for near sound field", El Diario de la Sociedad Acústica de Japón, Vol. 53, Nº 2, páginas 110 - 116, 1997.

5 Bibliografía no de patente 4: Yusuke Hioka, Kazunori Kobayashi, Kenichi Furuya y Akitoshi Kataoka, "Enhancement of Sound Sources Located within a Particular Area Using a Pair of Small Microphone arrays", Actas del IEICE sobre Fundamentos, Vol. E91-A, Nº 2, páginas 561 -574, agosto de 2004.

Bibliografía no de patente 5: Yusuke Hioka, Kenta Niwa, Sumitaka Sakauchi, Ken'ichi Furuta y Yoichi Haneda, "A method of separating sound sources located at different distances based on direct-to-reverberation ratio", Actas de la Reunión de Otoño de la Sociedad de Acústica de Japón, páginas 633 - 634, septiembre de 2009.

10 Bibliografía no de patente 6: Futoshi Asano et al., "Blind Source Separation in Reflective Sound Fields", Taller Internacional sobre Comunicación de Habla Manos Libres (HSC2001), Kyoto, Japón, 9-11 de abril de 2001, páginas 51-54.

Compendio de la invención

Problemas a ser resueltos por la invención

15 Según la técnica de realce de sonido directivo agudo descrita en la categoría [1], un sonido que llega de una dirección del objetivo no se puede realzar a menos que el micrófono en sí mismo apunte a la dirección del objetivo, como se puede ver a partir de los ejemplos de micrófonos de tubo acústico y los micrófonos parabólicos. Es decir, cuando la dirección del objetivo pueda variar, se necesitan medios de accionamiento y control para cambiar la orientación del micrófono de tubo acústico o del micrófono parabólico en sí mismo, a menos que se use una acción física humana. Además, mientras que el micrófono parabólico destaca en la captación de sonido de relación SN alta debido a que el micrófono parabólico puede enfocar la energía de los sonidos reflejados por el reflector parabólico en el punto focal, es difícil para el micrófono parabólico, así como para el micrófono de tubo acústico lograr una directividad alta, por ejemplo, un ángulo visual de aproximadamente 5° a 10° (directividad aguda de un ángulo de aproximadamente $\pm 5^\circ$ a $\pm 10^\circ$ con respecto a la dirección del objetivo).

25 Según la técnica de realce de sonido directivo agudo descrita en la categoría [2], con el fin de lograr una directividad más alta, se requieren más micrófonos y un tamaño de la agrupación más grande (una longitud total más grande de la agrupación). No es realista aumentar el tamaño de la agrupación ilimitadamente, debido a un espacio restringido donde se coloca la agrupación de micrófonos en fase, los costes y el número de micrófonos capaces de realizar procesamiento en tiempo real. Por ejemplo, los micrófonos disponibles en el mercado son capaces de procesar en tiempo real hasta aproximadamente 100 señales. La directividad que se puede lograr con una agrupación de micrófonos en fase con alrededor de 100 micrófonos es aproximadamente $\pm 30^\circ$ con respecto a la dirección del objetivo y por lo tanto es difícil para una agrupación de micrófonos en fase realzar el sonido de una dirección del objetivo con una directividad aguda de aproximadamente $\pm 5^\circ$ a $\pm 10^\circ$, por ejemplo. Además, es difícil para la técnica convencional en la categoría [2] captar un sonido de una dirección del objetivo con una relación SN alta de modo que el sonido no se oculte en sonidos de otras direcciones distintas de la dirección del objetivo.

35 Según la técnica de realce de sonido directivo agudo descrita en la categoría [3], mientras que un sonido de una dirección del objetivo se puede captar con una relación SN alta de modo que el sonido no se oculte en sonidos de direcciones distintas de la dirección del objetivo y los sonidos de cualquier dirección se pueden realzar sin necesitar los medios de accionamiento y control mencionados anteriormente, es difícil para la técnica lograr una directividad alta. En particular, la voz humana incluye una proporción alta de componentes de frecuencia en un intervalo de aproximadamente 100 Hz a aproximadamente 2 kHz. Sin embargo, es difícil para la técnica convencional en la categoría [3] lograr una directividad aguda de aproximadamente $\pm 5^\circ$ a $\pm 10^\circ$ en una dirección del objetivo en tal banda de baja frecuencia.

45 La técnica de realce de punto de sonido descrita en (1) no toma ninguna medida para proteger contra fuentes de interferencia debido a que la técnica usa el método de agrupación de retardo y suma. La técnica de realce de punto de sonido descrita en (2) requiere una pluralidad de agrupaciones de micrófonos y, por lo tanto, puede ser desventajosa debido al aumento de tamaño y coste del sistema. El aumento de tamaño de las agrupaciones de micrófonos restringe la instalación y el transporte de las agrupaciones. La información que se refiere a reverberación varía con los cambios de entorno y es difícil para la técnica de realce de punto de sonido descrita en (3) responder de manera robusta a tales cambios de entorno.

50 A la luz de estas circunstancias, un primer objeto de la presente invención es proporcionar una técnica de realce de sonido (una técnica de realce de punto de sonido) que pueda captar un sonido con una relación SN suficientemente alta y seguir un sonido de cualquier dirección sin necesidad de mover físicamente un micrófono, y aún tiene una directividad más aguda en una dirección deseada que las técnicas convencionales y puede realzar sonidos según las distancias de la agrupación de micrófonos. Un segundo objeto de la presente invención es proporcionar una técnica de realce de sonido (una técnica de realce de sonido directivo agudo) que puede captar un sonido con una relación SN suficientemente alta, puede seguir un sonido de cualquier dirección sin necesitar mover físicamente un micrófono y aún tiene una directividad más aguda en una dirección deseada que las técnicas convencionales.

Medios para resolver los problemas

(Técnica de realce de punto de sonido)

5 Una función de transferencia $a_{i,g}$ de un sonido que viene de cada una de una o más posiciones que se supone que son fuentes de sonido (donde i denota la dirección y g denota la distancia para identificar cada posición) y llega a los
 10 micrófonos (el número de micrófonos $M \geq 2$) se usa para obtener un filtro para una posición que es un objetivo de realce de sonido antes de capturar los M sonidos capturados con los M micrófonos [un proceso de diseño de filtro]. Cada función de transferencia $a_{i,g}$ se representa mediante la suma de funciones de transferencia de un sonido directo que viene de una posición determinada por una dirección i y una distancia g y llega directamente a los M
 15 micrófonos y funciones de transferencia de uno o más sonidos reflejados que se producen por reflexión del sonido directo fuera de un objeto reflector y llega a los M micrófonos. El filtro está diseñado para ser aplicado, para cada frecuencia, a una señal en el dominio de frecuencia transformada de cada una de las M señales captadas obtenidas captando sonidos con los M micrófonos. El filtro obtenido como resultado del proceso de diseño de filtro se aplica a una señal en el dominio de frecuencia para cada frecuencia para obtener una señal de salida [un proceso de aplicación de filtro]. La señal de salida es una señal en el dominio de frecuencia en la que se realiza el sonido de la
 posición que es el objetivo del realce de sonido.

Cada función de transferencia $a_{i,g}$ puede ser, por ejemplo, la suma de un vector de dirección de un sonido directo y un vector o vectores de dirección de uno o más sonidos reflejados cuyas descomposiciones debidas a reflexión y a diferencias de tiempo de llegada del sonido directo han sido corregidas o se pueden obtener mediante mediciones en un entorno real.

20 En el proceso de diseño de filtro, se puede obtener un filtro para cada frecuencia de manera que se minimiza la potencia de sonidos de posiciones distintas de la posición que es el objetivo del realce de sonido. Alternativamente, se puede obtener un filtro para cada frecuencia de manera que se maximice la relación de SN de un sonido de la posición que es el objetivo del realce de sonido. Alternativamente, se puede obtener un filtro para cada frecuencia de
 25 manera que se minimiza la potencia de sonidos de posiciones distintas de una o más posiciones que supone que son fuentes de sonido mientras que un coeficiente de filtro para uno de los M micrófonos se mantiene en un valor constante.

Alternativamente, el filtro se puede obtener para cada frecuencia en el proceso de diseño de filtro de manera que la potencia de sonidos de posiciones distintas a la posición que es el objetivo del realce de sonido y puntos de
 30 supresión se minimiza en condiciones que (1) el filtro pasa sonidos en todas las bandas de frecuencia de la posición que es el objetivo del realce de sonido y que (2) el filtro suprime sonidos en todas las bandas de frecuencia de uno o más puntos de supresión. Alternativamente, el filtro puede ser obtenido para cada frecuencia normalizando una función de transferencia $a_{s,h}$ de un sonido desde la posición en $i = s$, $g = h$ que es el objetivo de realce del sonido. Alternativamente, se puede obtener un filtro para cada frecuencia usando una matriz de correlación espacial representada por funciones de transferencia $a_{i,g}$ correspondientes a posiciones distintas de la posición que es el
 35 objetivo del realce de sonido. Alternativamente, el filtro se puede obtener para cada frecuencia de manera que la potencia de sonidos de posiciones distintas de la posición que es el objetivo de realce de sonido se minimice bajo la condición de que el filtro reduzca la cantidad de descomposición de un sonido de la posición que es el objetivo del realce de sonido a un valor predeterminado o menos. Alternativamente, se puede obtener un filtro para cada frecuencia usando una matriz de correlación espacial representada por señales en el dominio de frecuencia
 40 obtenidas transformando señales obtenidas mediante la observación con una agrupación de micrófonos. Alternativamente, se puede obtener un filtro para cada frecuencia usando una matriz de correlación espacial representada por funciones de transferencia $a_{i,g}$ correspondientes a cada una de una o más posiciones que se supone que son fuentes de sonido.

(Técnica de realce de sonido directivo agudo)

45 Una función de transferencia a_{θ} de un sonido que viene de cada una de una o más direcciones a partir de las cuales se supone que vienen sonidos y llegan a los micrófonos (el número de micrófonos $M \geq 2$) se usa para obtener un filtro para una posición que es el objetivo de realce de sonido antes de captar los M sonidos captados con los M
 50 micrófonos [un proceso de diseño de filtro]. Cada función de transferencia a_{θ} se representa por la suma de funciones de transferencia de un sonido directo que viene de una dirección θ y llega directamente a los M micrófonos y las funciones de transferencia de uno o más sonidos reflejados que se producen por reflexión del sonido directo fuera de un objeto reflector y llega a los M micrófonos. El filtro está diseñado para ser aplicado, para cada frecuencia, a una señal en el dominio de frecuencia transformada a partir de cada una de las M señales captadas obtenidas captando sonidos con los M micrófonos. El filtro obtenido como resultado del proceso de diseño de filtro se aplica a una señal en el dominio de frecuencia para cada frecuencia para obtener una señal de salida [un proceso de aplicación de filtro]. La señal de salida es una señal en el dominio de frecuencia en la que se realiza el sonido de la
 55 posición que es el objetivo de realce de sonido.

Cada función de transferencia a_{θ} puede ser, por ejemplo, la suma de un vector de dirección de un sonido directo y un vector o vectores de dirección de uno o más sonidos reflejados cuyas descomposiciones debido a reflexión y a

diferencias de tiempo de llegada del sonido directo han sido corregidas o se pueden obtener mediante mediciones en un entorno real.

5 En el proceso de diseño de filtro, se puede obtener un filtro para cada frecuencia de manera que se minimiza la potencia de sonidos de direcciones distintas de la dirección que es el objetivo realce de sonido. Alternativamente, se puede obtener un filtro para cada frecuencia de manera que se maximice la relación SN de un sonido de la dirección que es el objetivo del realce de sonido. Alternativamente, se puede obtener un filtro para cada frecuencia de manera que se minimice la potencia de sonidos de direcciones desde las cuales es probable que lleguen sonidos mientras que un coeficiente de filtro para uno de los M micrófonos se mantiene en un valor constante.

10 Alternativamente, el filtro se puede obtener para cada frecuencia en el proceso de diseño de filtro de manera que se minimice la potencia de los sonidos de direcciones distintas a la dirección que es el objetivo de realce de sonido y direcciones nulas en condiciones que (1) el filtro pasa sonidos en todas las bandas de frecuencia de la dirección que es el objetivo de realce de sonido y que (2) el filtro suprime sonidos en todas las bandas de frecuencia de una o más direcciones nulas. Alternativamente, el filtro se puede obtener para cada frecuencia normalizando una función de transferencia a_s de un sonido de la dirección $\theta = s$ que es el objetivo de realce de sonido. Alternativamente, se puede obtener un filtro para cada frecuencia usando una matriz de correlación espacial representada por funciones de transferencia a_θ correspondientes a direcciones distintas a la dirección que es el objetivo de realce de sonido. Alternativamente, el filtro se puede obtener para cada frecuencia de manera que se minimice la potencia de sonidos de direcciones distintas de la dirección que es el objetivo de realce de sonido bajo la condición de que el filtro reduzca la cantidad de descomposición de un sonido de la dirección que es el objetivo de realce de sonido a un valor predeterminado o menos. Alternativamente, se puede obtener un filtro para cada frecuencia usando una matriz de correlación espacial representada por señales en el dominio de frecuencia obtenidas transformando señales obtenidas mediante la observación con una agrupación de micrófonos-

15

20

Efectos de la invención

(Técnica de realce de punto de sonido)

25 Dado que la técnica de realce de punto de sonido de la presente invención usa no solamente un sonido directo de una dirección deseada, sino también sonidos reflejados, la técnica de realce de punto de sonido es capaz de captar sonidos con una relación SN suficientemente alta desde la dirección. Además, la técnica de realce de punto de sonido de la presente invención es capaz de seguir un sonido en cualquier dirección sin necesidad de mover físicamente el micrófono debido a que el realce de del sonido se logra mediante procesamiento de señal. Además,

30 dado que cada función de transferencia $a_{i,g}$ está representada por la suma de la función de transferencia de un sonido directo que viene de la posición determinada por una dirección i y una distancia g y llega directamente a M micrófonos y la función o funciones de transferencia de uno o más sonidos reflejados que se producen por reflexión del sonido fuera de un objeto reflector y llegan a los M micrófonos, un filtro que aumenta el grado de supresión de coherencia que determina el grado de directividad en una dirección deseada, se pueden diseñar según los criterios

35 típicos de diseño de filtro, como se describirá más tarde en mayor detalle en la sección <<Principio de la técnica de realce de punto de sonido>>. Es decir, se puede lograr una directividad más aguda en la dirección deseada de lo que era posible anteriormente. Dado que los sonidos reflejados se usan como se describirá más tarde en la sección <<Principio de la técnica de realce del punto de sonido>>, hay diferencias significativas en la función de transferencia entre sonidos de diferentes posiciones a diferentes distancias en aproximadamente la misma dirección como se ve desde la agrupación de micrófonos. Extrayendo las diferencias entre las funciones de transferencia mediante la conformación de haces, sonidos en un intervalo estrecho que incluye una dirección deseada se pueden realzar según las distancias de la agrupación de micrófonos.

40

(Técnica de realce de sonido directivo agudo)

45 Dado que la técnica de realce de sonido directivo agudo de la presente invención usa no solamente un sonido directo de una dirección deseada sino también sonidos reflejados, la técnica de realce de sonido directivo agudo es capaz de captar sonidos con una relación SN suficientemente alta desde la dirección. Además, la técnica de realce de sonido directivo agudo de la presente invención es capaz de seguir un sonido en cualquier dirección sin necesidad de mover físicamente el micrófono debido a que el realce de sonido se logra mediante procesamiento de señal. Además, dado que cada función de transferencia a_θ se representa por la suma de la función de transferencia de un sonido directo que viene de una dirección Φ y llega directamente a M micrófonos y la función o funciones de transferencia de uno o más sonidos reflejados que se producen por reflexión del sonido fuera de un objeto reflector y llegan a los M micrófonos, un filtro que aumenta el grado de supresión de coherencia que determina el grado de directividad en una dirección deseada se puede diseñar con criterios típicos de diseño de filtro, como se describirá más tarde en más detalle en la sección <<Principio de realce de sonido directivo agudo>>. Es decir, se puede lograr

50

55 una directividad más aguda en una dirección deseada de lo que era posible anteriormente.

Breve descripción de los dibujos

La Fig. 1A es un diagrama que ilustra que sonidos que llegan de una dirección del objetivo se realzan mediante un micrófono de tubo acústico;

- La Fig. 1B es un diagrama que ilustra que sonidos que llegan de direcciones distintas de una dirección del objetivo se suprimen por un micrófono de tubo acústico;
- La Fig. 2A es un diagrama que ilustra que sonidos que llegan de una dirección objetivo se realzan mediante un micrófono parabólico;
- 5 La Fig. 2B es un diagrama que ilustra que sonidos que llegan de direcciones distintas de una dirección del objetivo se suprimen por un micrófono parabólico;
- La Fig. 3 es un diagrama que ilustra que se realiza un sonido de una dirección del objetivo y se suprime un sonido de una dirección distinta de la dirección del objetivo usando una agrupación de micrófonos en fase que incluye una pluralidad de micrófonos;
- 10 La Fig. 4 es un diagrama que ilustra una configuración funcional de una técnica de realce de sonido directivo agudo que usa conformación de haces múltiples como ejemplo de técnicas convencionales.
- La Fig. 5A es un diagrama que muestra esquemáticamente que una directividad suficientemente alta no se puede lograr teniendo en cuenta solamente sonidos directos;
- 15 La Fig. 5B es un diagrama que muestra esquemáticamente que se puede lograr una directividad suficientemente alta teniendo en cuenta tanto los sonidos directos como los reflejados;
- La Fig. 6 es un diagrama que muestra las dependencias de dirección de coherencias de una técnica convencional y un principio de la presente invención;
- La Fig. 7 es un diagrama que ilustra una configuración funcional de un aparato de realce de sonido directivo agudo (primera realización);
- 20 La Fig. 8 es un diagrama que ilustra un procedimiento de un método de realce de sonido directivo agudo (primera realización);
- La Fig. 9 es un diagrama que ilustra una configuración de un primer ejemplo;
- La Fig. 10 es un diagrama que ilustra una configuración funcional de un aparato de realce de sonido directivo agudo (segunda realización);
- 25 La Fig. 11 es un diagrama que ilustra un procedimiento de un método de realce de sonido directivo agudo (segunda realización);
- La Fig. 12 es un diagrama que muestra los resultados de un experimento en un primer ejemplo;
- La Fig. 13 es un diagrama que muestra los resultados de un experimento en el primer ejemplo;
- La Fig. 14 es un diagrama que muestra la directividad con un filtro $W^-(\omega, \theta)$ en el primer ejemplo;
- 30 La Fig. 15 es un diagrama que ilustra una configuración de un segundo ejemplo;
- La Fig. 16 es un diagrama que muestra los resultados de un experimento en un ejemplo experimental;
- La Fig. 17 es un diagrama que ilustra los resultados de un experimento en un ejemplo experimental;
- La Fig. 18A es un diagrama que ilustra sonidos directos que llegan a una agrupación de micrófonos de dos fuentes de sonido A y B;
- 35 La Fig. 18B es un diagrama que ilustra sonidos directos que llegan a una agrupación de micrófonos de dos fuentes de sonido A y B y sonidos reflejados que llegan a la agrupación de micrófonos de dos fuentes de sonido virtuales $A(\xi)$ y $B(\xi)$;
- La Fig. 19 es un diagrama que ilustra una configuración funcional de un aparato de realce de punto de sonido (primera realización);
- 40 La Fig. 20 es un diagrama que ilustra un procedimiento de un método de realce de punto de sonido (primera realización);
- La Fig. 21 es un diagrama que ilustra una configuración funcional de un aparato de realce de punto de sonido (segunda realización);
- 45 La Fig. 22 es un diagrama que ilustra un procedimiento de un método de realce de punto de sonido (segunda realización);

La Fig. 23A ilustra la directividad (en un dominio bidimensional) de un conformador de haces de varianza mínima sin reflector;

La Fig. 23B ilustra la directividad (en un dominio bidimensional) de un conformador de haces de varianza mínima con reflector;

5 La Fig. 24A es una vista en planta que ilustra una configuración ejemplar de una implementación de la presente invención;

La Fig. 24B es una vista frontal que ilustra la configuración ejemplar de la implementación de la presente invención;

La Fig. 24C es una vista lateral que ilustra la configuración ejemplar de la implementación de la presente invención;

10 La Fig. 25A es una vista lateral que ilustra otra configuración ejemplar de una implementación de la presente invención;

La Fig. 25B es una vista lateral que ilustra otra configuración ejemplar de una implementación de la presente invención;

La Fig. 26 es un diagrama que ilustra una forma en uso de la configuración ejemplar de la implementación ilustrada en la Fig. 25B;

15 La Fig. 27A es una vista en planta que ilustra una configuración ejemplar de una implementación de la presente invención;

La Fig. 27B es una vista frontal que ilustra la configuración ejemplar de la implementación de la presente invención;

La Fig. 27C es una vista lateral que ilustra la configuración ejemplar de la implementación de la presente invención;

20 La Fig. 28 es una vista lateral que ilustra una configuración ejemplar de una implementación de la presente invención.

Descripción detallada de las realizaciones

Se describirá en primer lugar una técnica de realce de sonido directivo agudo y luego se describirá una técnica de realce de punto de sonido.

25 <<Técnica de realce de sonido directivo agudo>>

Se describirá un principio de una técnica de realce de sonido directivo agudo de la presente invención. La técnica de realce de sonido directivo agudo de la presente invención se basa en la naturaleza de una técnica de agrupación de micrófonos que es capaz de seguir los sonidos de cualquier dirección sobre la base de procesamiento de señal y usa positivamente los sonidos reflejados para captar sonidos con una relación SN alta. Una característica de la presente invención es un uso combinado de los sonidos reflejados y una técnica de procesamiento de señal que permite una directividad aguda.

30 Antes de la descripción, se definirán de nuevo los símbolos. El índice de una frecuencia discreta se denota por ω (El índice ω de una frecuencia discreta se puede considerar que es una frecuencia angular ω debido a que una frecuencia f y una frecuencia angular ω satisfacen la relación $\omega = 2\pi f$. Con respecto a ω , el "índice de una frecuencia discreta" también se puede conocer algunas veces simplemente como "frecuencia") y el índice del número de tiempo de trama se denota por k . Una representación en el dominio de frecuencia de una trama de orden k de una señal analógica recibida en M micrófonos se denota por $X^T(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y un filtro que realza una señal en el dominio de frecuencia $X^T(\omega, k)$ de un sonido de una dirección del objetivo θ_s como se ve desde el centro de una agrupación de micrófonos con una frecuencia ω se denota por $W^T(\omega, \theta_s)$, donde M es un número entero mayor o igual a 2 y T representa la transposición. Entonces, una señal en el dominio de frecuencia $Y(\omega, k, \theta_s)$ resultante del realce de la señal en el dominio de frecuencia $X^T(\omega, k)$ del sonido de la dirección del objetivo θ_s con la frecuencia ω (en lo sucesivo, la señal resultante se conoce como señal de salida) se puede dar por la ecuación (6):

$$Y(\omega, k, \theta_s) = \vec{W}^H(\omega, \theta_s) \vec{X}(\omega, k) \quad (6)$$

45 donde H representa la transposición hermitiana.

Mientras que el "centro de una agrupación de micrófonos" se puede determinar arbitrariamente, típicamente el centro geométrico de la agrupación de los M micrófonos se trata como el "centro de una agrupación de micrófonos". En el caso de una agrupación de micrófonos lineales, por ejemplo, el punto equidistante de los micrófonos en ambos extremos de la agrupación se trata como el "centro de la agrupación de micrófonos". En el caso de una agrupación

de micrófonos planos en la que los micrófonos están dispuestos en una matriz cuadrada de $m \times m$ ($m^2 = M$), por ejemplo, la posición en la que las diagonales que unen los micrófonos en las esquinas se cruzan, se trata como el "centro de la agrupación de micrófonos".

Un filtro $W^{-}(\omega, \theta_s)$ se puede diseñar de varias formas. Aquí se describirá un diseño que usa el método de respuesta sin distorsión de varianza mínima (MVDR). En el método MVDR, un filtro $W^{-}(\omega, \theta_s)$ está diseñado de modo que la potencia de sonidos de direcciones distintas de una dirección del objetivo θ_s (en lo sucesivo, sonidos de direcciones distintas de la dirección del objetivo θ_s también se conocerán como "ruido") se minimiza a la frecuencia ω (véase la ecuación (7)) usando una matriz de correlación espacial $Q(\omega)$ bajo la condición de restricción de la ecuación (8). Las funciones de transferencia a una frecuencia ω entre una fuente de sonido y los M micrófonos se denotan por $\vec{a}^{-}(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$, donde la fuente de sonido se supone que está en una dirección θ_s . En otras palabras, $\vec{a}^{-}(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ representa funciones de transferencia de un sonido de la dirección θ_s a los micrófonos incluidos en la agrupación de micrófonos a la frecuencia ω . La matriz de correlación espacial $Q(\omega)$ representa la correlación entre los componentes $X_1(\omega, k), \dots, X_M(\omega, k)$ de una señal en el dominio de frecuencia $X^{-}(\omega, k)$ a la frecuencia ω y tiene $E[X_i(\omega, k)X_j^*(\omega, k)]$ ($1 \leq i \leq M, 1 \leq j \leq M$) como sus elementos (i, j) . El operador $E[\cdot]$ representa una operación de promediado estadístico y el símbolo $*$ es un operador conjugado complejo. La matriz de correlación espacial $Q(\omega)$ se puede expresar usando valores estadísticos de $X_1(\omega, k), \dots, X_M(\omega, k)$ obtenidos a partir de la observación o se pueden expresar usando funciones de transferencia. Este último caso, donde la matriz de correlación espacial $Q(\omega)$ se expresa usando funciones de transferencia, se describirá momentáneamente en lo sucesivo.

$$\min_{\vec{W}(\omega, \theta_s)} \left(\vec{W}^H(\omega, \theta_s) Q(\omega) \vec{W}(\omega, \theta_s) \right) \quad (7)$$

$$\vec{W}^H(\omega, \theta_s) \vec{a}(\omega, \theta_s) = 1.0 \quad (8)$$

Es conocido que el filtro $W^{-}(\omega, \theta_s)$ que es una solución óptima de la ecuación (7) se puede dar por la ecuación (9) (véase la Referencia 1 enumerada a continuación).

$$\vec{W}(\omega, \theta_s) = \frac{Q^{-1}(\omega) \vec{a}(\omega, \theta_s)}{\vec{a}^H(\omega, \theta_s) Q^{-1}(\omega) \vec{a}(\omega, \theta_s)} \quad (9)$$

Como se apreciará a partir del hecho de que la matriz inversa de la matriz de correlación espacial $Q(\omega)$ se incluye en la ecuación (9), la estructura de la matriz de correlación espacial $Q(\omega)$ es importante para lograr una directividad aguda. Se apreciará a partir de la ecuación (7) que la potencia de ruido depende de la estructura de la matriz de correlación espacial $Q(\omega)$.

Un conjunto de índices p de direcciones desde las cuales llega el ruido se denota por $\{1, 2, \dots, P-1\}$. Se supone que el índice s de la dirección del objetivo θ_s no pertenece al conjunto $\{1, 2, \dots, P-1\}$. Suponiendo que $P-1$ ruidos vienen de direcciones arbitrarias, la matriz de correlación espacial $Q(\omega)$ se puede dar por la ecuación (10a). Con el fin de diseñar un filtro que funcione suficientemente en presencia de muchos ruidos, es preferible que P sea un valor relativamente grande. Se supone aquí que P es un número entero del orden de M . Mientras que la descripción se da como si la dirección del objetivo θ_s es una dirección constante (y, por lo tanto, direcciones distintas de la dirección del objetivo θ_s se describen como direcciones desde las cuales llega el ruido) por claridad de la explicación del principio de la técnica de realce de sonido directivo agudo de la presente invención, la dirección del objetivo θ_s en realidad puede ser cualquier dirección que pueda ser un objetivo de realce de sonido. Normalmente, una pluralidad de direcciones pueden ser direcciones del objetivo θ_s . En este sentido, la diferenciación entre la dirección del objetivo θ_s y las direcciones de ruido es subjetiva. Es más correcto considerar que una dirección seleccionada de P direcciones diferentes que están predeterminadas como una pluralidad de direcciones posibles a partir de las cuales puede llegar cualquier sonido, incluyendo un sonido o ruido del objetivo, es la dirección del objetivo y las otras direcciones son direcciones de ruido. Por lo tanto, la matriz de correlación espacial $Q(\omega)$ se puede representar mediante funciones de transferencia $\vec{a}^{-}(\omega, \theta_\phi) = [a_1(\omega, \theta_\phi), \dots, a_M(\omega, \theta_\phi)]^T$ ($\phi \in \Phi$) de sonidos que vienen de direcciones θ_ϕ incluidas en una pluralidad de direcciones posibles desde las cuales pueden llegar los sonidos a los micrófonos y se pueden escribir como la ecuación (10b), donde Φ es la unión del conjunto $\{1, 2, \dots, P-1\}$ y un conjunto $\{s\}$. Obsérvese que $|\Phi| = P$ y $|\Phi|$ representa el número de elementos del conjunto Φ .

$$Q(\omega) = \vec{a}(\omega, \theta_s) \vec{a}^H(\omega, \theta_s) + \sum_{p \in \{1, \dots, P-1\}} \vec{a}(\omega, \theta_p) \vec{a}^H(\omega, \theta_p) \quad (10a)$$

$$Q(\omega) = \sum_{\phi \in \Phi} \vec{a}(\omega, \theta_\phi) \vec{a}^H(\omega, \theta_\phi) \quad (10b)$$

Aquí, se supone que la función de transferencia $a^{\rightarrow}(\omega, \theta_s)$ de un sonido de la dirección del objetivo θ_s y las funciones de transferencia $a^{\rightarrow}(\omega, \theta_p) = [a_1(\omega, \theta_p), \dots, a_M(\omega, \theta_p)]^T$ de los sonidos de las direcciones $p \in \{1, 2, \dots, P - 1\}$ son ortogonales entre sí. Es decir, se supone que hay P sistemas base ortogonales que satisfacen la condición dada por la ecuación (11). El símbolo \perp representa ortogonalidad. Si $A^{\rightarrow} \perp B^{\rightarrow}$, el producto interno de los vectores A^{\rightarrow} y B^{\rightarrow} es cero. Se supone aquí que $P \leq M$. Obsérvese que si la condición dada por la ecuación (11) se puede relajar para suponer que hay P sistemas base que se pueden considerar aproximadamente como sistemas base ortogonales, P es preferiblemente un valor del orden de M o un valor relativamente grande mayor o igual a M.

$$\vec{a}(\omega, \theta_s) \perp \vec{a}(\omega, \theta_1) \perp \dots \perp \vec{a}(\omega, \theta_{P-1}) \quad (11)$$

Entonces, la matriz de correlación espacial $Q(\omega)$ se puede expandir como la ecuación (12). La ecuación (12) significa que la matriz de correlación espacial $Q(\omega)$ se puede descomponer en una matriz $V(\omega) = [a^{\rightarrow}(\omega, \theta_s), a^{\rightarrow}(\omega, \theta_1), \dots, a^{\rightarrow}(\omega, \theta_{P-1})]^T$ compuesta de P funciones de transferencia que satisfacen la ortogonalidad y una matriz unidad $\Lambda(\omega)$. Aquí, ρ es un valor propio de una función de transferencia $a^{\rightarrow}(\omega, \theta_p)$ que satisface la ecuación (11) para la matriz de correlación espacial $Q(\omega)$ y es un valor real.

$$Q(\omega) = \rho \vec{V}(\omega) \vec{\Lambda}(\omega) \vec{V}^H(\omega) \quad (12)$$

Entonces, la matriz inversa de la matriz de correlación espacial $Q(\omega)$ se puede dar por la ecuación (13).

$$Q^{-1}(\omega) = \frac{1}{\rho} \vec{V}^H(\omega) \vec{\Lambda}^{-1}(\omega) \vec{V}(\omega) \quad (13)$$

La sustitución de la ecuación (13) en la ecuación (7) muestra que se minimiza la potencia de ruido. Si se minimiza la potencia de ruido, significa que se logra la directividad en la dirección del objetivo θ_s . Por lo tanto, la ortogonalidad entre las funciones de transferencia de sonidos de diferentes direcciones es una condición importante para lograr la directividad en la dirección del objetivo θ_s .

La razón por la cual es difícil para las técnicas convencionales lograr una directividad aguda en una dirección del objetivo θ_s se tratará a continuación.

Las técnicas convencionales asumieron en el diseño de filtros que las funciones de transferencia estaban compuestas de esos sonidos directos. En realidad, hay sonidos reflejados que se producen por la reflexión de sonidos de la misma fuente de sonido fuera de superficies tales como las paredes y el techo, y llegan a los micrófonos. Sin embargo, las técnicas convencionales consideraban los sonidos reflejados como un factor que degrada la directividad e ignoraban la presencia de los sonidos reflejados. En las técnicas convencionales, las funciones de transferencia $a^{\rightarrow}_{conv}(\omega, \theta) = [a_1(\omega, \theta), \dots, a_M(\omega, \theta)]^T$ se trataron como $a^{\rightarrow}_{conv}(\omega, \theta) = h^{\rightarrow}_d(\omega, \theta)$, donde $h^{\rightarrow}_d(\omega, \theta) = [h_{d1}(\omega, \theta), \dots, h_{dM}(\omega, \theta)]^T$ representa vectores de dirección solamente de un sonido directo que llega de una dirección θ . Obsérvese que un vector de dirección es un vector complejo donde las características de respuesta de fase de los micrófonos a una frecuencia ω con respecto a un punto de referencia están dispuestas para una onda de sonido de una dirección θ vista desde el centro de la agrupación de micrófonos.

Suponiendo que los sonidos llegan a una agrupación de micrófonos lineales como ondas planas, un elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h^{\rightarrow}_d(\omega, \theta)$ de un sonido directo se da, por ejemplo, por la ecuación (14a), donde m es un número entero que satisface $1 \leq m \leq M$, c representa la velocidad del sonido, u representa la distancia entre micrófonos adyacentes, j es una unidad imaginaria. El punto de referencia es el punto medio de la longitud completa de la agrupación de micrófonos lineales (el centro de la agrupación de micrófonos lineales). La dirección θ se define como el ángulo formado por la dirección desde la cual llega un sonido directo y la dirección en la cual los micrófonos incluidos en la agrupación de micrófonos lineales, como se ve desde el centro de la agrupación de micrófonos lineales (véase la Fig. 9). Obsérvese que un vector de dirección se puede expresar de varias formas. Por ejemplo, suponiendo que el punto de referencia es la posición del micrófono en un extremo de la agrupación de micrófonos lineales, un elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h^{\rightarrow}_d(\omega, \theta)$ de un sonido directo se puede dar por la ecuación (14b). En la siguiente descripción, la suposición es que el elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h^{\rightarrow}_d(\omega, \theta)$ de un sonido directo se puede escribir como la ecuación (14a).

$$h_{dm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos\theta\right] \quad (14a)$$

$$h_{dm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos\theta\right] \quad (14b)$$

45

El producto interno $\gamma_{conv}(\omega, \theta)$ de una función de transferencia de una dirección θ y una función de transferencia de una dirección del objetivo θ_s se pueden dar por la ecuación (15), donde $\theta \neq \theta_s$.

$$\begin{aligned} \gamma_{conv}(\omega, \theta) &= \vec{a}_{conv}^H(\omega, \theta_s) \vec{a}_{conv}(\omega, \theta) \\ &= \vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\ &= \sum_{m=1}^M \exp\left[-\frac{j\omega u}{c} \left(m - \frac{M+1}{2}\right) (\cos\theta - \cos\theta_s)\right] \end{aligned} \quad (15)$$

5 En lo sucesivo, $\gamma_{conv}(\omega, \theta)$ se conoce como coherencia. La dirección θ en el cual la coherencia $\gamma_{conv}(\omega, \theta)$ es 0 se puede dar por la ecuación (16), donde q es un número entero arbitrario, excepto 0. Dado que $0 < \theta < \pi/2$, el intervalo de q está limitado para cada banda de frecuencia.

$$\theta = \arccos\left(\frac{2q\pi c}{M\omega u} + \cos\theta_s\right) \quad (16)$$

10 Dado que solamente se pueden cambiar los parámetros relacionados con el tamaño de la agrupación de micrófonos (M y u) en la ecuación (16), es difícil reducir la coherencia $\gamma_{conv}(\omega, \theta)$ sin cambiar ninguno de los parámetros relacionados con el tamaño de la agrupación de micrófonos si la diferencia (diferencia angular) $|\theta - \theta_s|$ entre direcciones es pequeña. Si este es el caso, la potencia de ruido no se reduce a un valor suficientemente pequeño y resultará una directividad que tiene una anchura de haz amplia en la dirección del objetivo θ_s como se ilustra esquemáticamente en la Fig. 5A.

15 La técnica de realce de sonido directivo agudo de la presente invención se basa en la consideración descrita anteriormente y se caracteriza por tener en cuenta positivamente los sonidos reflejados, a diferencia de la técnica convencional, sobre la base de un entendimiento de que con el fin de diseñar un filtro que proporciona una directividad aguda en la dirección del objetivo θ_s , es importante permitir que la coherencia sea reducida a un valor suficientemente pequeño, incluso cuando la diferencia (diferencia angular) $|\theta - \theta_s|$ entre direcciones es pequeña, el filtro que se calcula antes de captar los sonidos a ser realizados con los M micrófonos. Dos tipos de ondas planas, esto es, sonidos directos de una fuente de sonido y sonidos reflejados producidos por la reflexión de ese sonido fuera de un objeto 300 reflector, juntos entran en los micrófonos de una agrupación de micrófonos. Permitamos que el número de sonidos reflejados sea denotado por Ξ . Aquí, Ξ es un número entero predeterminado mayor o igual a 1. Entonces, una función de transferencia $\vec{a}^T(\omega, \theta) = [a_1(\omega, \theta), \dots, a_M(\omega, \theta)]^T$ se puede expresar por la suma de una función de transferencia de un sonido directo que viene de una dirección que puede ser un objetivo de realce de sonido y llega directamente a la agrupación de micrófonos y la función o funciones de transferencia de uno o más sonidos reflejados que se producen por reflexión de ese sonido fuera de un objeto reflector y llegan a la agrupación de micrófonos. Específicamente, la función de transferencia se puede representar como la suma del vector de dirección del sonido directo y del vector de dirección de Ξ sonidos reflejados cuyas descomposiciones debidas a la reflexión y a las diferencias de tiempo de llegada del sonido directo se corrigen, como se muestra en la ecuación (17a), donde $\tau_\xi(\theta)$ es la diferencia de tiempo de llegada entre el sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y α_ξ ($1 \leq \xi \leq \Xi$) es un coeficiente para tener en cuenta descomposiciones de sonidos debidas a la reflexión. Aquí, $\vec{h}_{r\xi}^T(\omega, \theta) = [h_{r1\xi}(\omega, \theta), \dots, h_{rM\xi}(\omega, \theta)]^T$ representa los vectores de dirección de los sonidos reflejados correspondientes al sonido directo de la dirección θ . Típicamente, α_ξ ($1 \leq \xi \leq \Xi$) es menor o igual a 1 ($1 \leq \xi \leq \Xi$). Para cada sonido reflejado, si el número de reflexiones en el camino desde la fuente de sonido a los micrófonos es 1, α_ξ ($1 \leq \xi \leq \Xi$) se puede considerar que representa la reflectancia acústica del objeto desde el cual se reflejó el sonido reflejado de orden ξ .

$$\vec{a}(\omega, \theta) = \vec{h}_d(\omega, \theta) + \sum_{\xi=1}^{\Xi} \alpha_\xi \exp[-j\omega\tau_\xi(\theta)] \cdot \vec{h}_{r\xi}(\omega, \theta) \quad (17a)$$

40 Dado que se proporcionan uno o más sonidos reflejados a la agrupación de micrófonos compuesta por M micrófonos, son necesarios uno o más objetos reflectores. Desde este punto de vista, una fuente de sonido, la agrupación de micrófonos y uno o más objetos reflectores están preferiblemente en tal relación de posición que un

sonido de la fuente de sonido se refleja fuera al menos un objeto reflector antes de llegar a la agrupación de micrófonos, suponiendo que la fuente de sonido se sitúa en la dirección del objetivo. Cada uno de los objetos reflectores tiene una forma bidimensional (por ejemplo, una placa plana) o una forma tridimensional (por ejemplo, una forma parabólica). Cada objeto reflector tiene preferiblemente aproximadamente el tamaño de la agrupación de micrófonos o mayor (mayor en un factor de 1 a 2). Con el fin de usar eficazmente los sonidos reflejados, la reflectancia α_ξ ($1 \leq \xi \leq \Xi$) de cada objeto reflector es preferiblemente al menos mayor que 0, y más preferiblemente, la amplitud de un sonido reflejado que llega a la agrupación de micrófonos es mayor que la amplitud del sonido directo en un factor de 0,2 o mayor. Por ejemplo, cada objeto reflector es un sólido rígido. Cada objeto reflector puede ser un objeto móvil (por ejemplo, un reflector) o un objeto inamovible (como un suelo, una pared o un techo). Obsérvese que si un objeto inamovible se establece como un objeto reflector, el vector de dirección para el objeto reflector necesita ser cambiado a medida que la agrupación de micrófonos se reubica (véanse las funciones $\psi(\theta)$ y $\psi_\xi(\theta)$ descritas más tarde) y, en consecuencia, el filtro necesita ser recalculado (restablecido). Por lo tanto, los objetos reflectores son preferiblemente accesorios de la agrupación de micrófonos en aras de la robustez frente a los cambios de entorno (en este caso, Ξ sonidos reflejados supuestos se considera que son sonidos reflejados fuera de los objetos reflectores). Aquí los “accesorios de la agrupación de micrófonos” son “objetos tangibles capaces de seguir los cambios de posición y orientación de la agrupación de micrófonos mientras se mantiene la relación de posición (relación geométrica) con la agrupación de micrófonos). Un ejemplo simple puede ser una configuración donde los objetos reflectores se fijan a la agrupación de micrófonos.

Con el fin de describir concretamente las ventajas de la técnica de realce de sonido directivo agudo de la presente invención, se supone en lo siguiente que $\Xi = 1$, los sonidos se reflejan una vez y existe un objeto reflector a una distancia de L metros del centro de la agrupación de micrófonos. El objeto reflector es un objeto grueso y rígido. Dado que $\Xi = 1$ en este caso, el símbolo que representa esto se omite y, por lo tanto, la ecuación (17a) se puede reescribir como la ecuación (17b):

$$\bar{a}(\omega, \theta) = \bar{h}_d(\omega, \theta) + \alpha \exp[-j\omega\tau(\theta)] \cdot \bar{h}_r(\omega, \theta) \quad (17b)$$

Un elemento de orden m del vector de dirección $\bar{h}_r(\omega, \theta) = [h_{r1}(\omega, \theta), \dots, h_{rm}(\omega, \theta)]^T$ de un sonido reflejado se puede dar por la ecuación (18a) de la misma forma que se representa el vector de dirección de un sonido directo (véase la ecuación (14a)). La función $\psi(\theta)$ emite la dirección desde la cual llega un sonido reflejado. Obsérvese que si el vector de dirección de un sonido directo se escribe como la ecuación (14b), un elemento de orden m del vector de dirección $\bar{h}_r(\omega, \theta) = [h_{r\xi}(\omega, \theta), \dots, h_{rM}(\omega, \theta)]^T$ de un sonido reflejado se da por la ecuación (18b). Típicamente, un elemento de orden m de un vector de dirección de orden ξ ($1 \leq \xi \leq \Xi$) $\bar{h}_{r\xi}(\omega, \theta) = [h_{r1\xi}(\omega, \theta), \dots, h_{rM\xi}(\omega, \theta)]^T$ se da por la ecuación (18c) o la ecuación (18d). La función $\psi_\xi(\theta)$ emite la dirección desde la cual llega el sonido reflejado de orden ξ .

$$h_{rm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos(\Psi(\theta))\right] \quad (18a)$$

$$h_{rm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos(\Psi(\theta))\right] \quad (18b)$$

$$h_{r\xi}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos(\Psi_\xi(\theta))\right] \quad (18c)$$

$$h_{r\xi}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos(\Psi_\xi(\theta))\right] \quad (18d)$$

Dado que la ubicación de un objeto reflector se puede establecer según sea adecuado, la dirección desde la cual llega un sonido reflejado se puede tratar como un parámetro variable.

Suponiendo que un objeto reflector de placa plana está cerca de la agrupación de micrófonos (la distancia L no es extremadamente grande en comparación con el tamaño de la agrupación de micrófonos), la coherencia $\gamma(\omega, \theta)$ se da por la ecuación (19), donde $\theta \neq \theta_s$.

$$\begin{aligned}
 \gamma(\omega, \theta) &= \vec{a}^H(\omega, \theta_s) \vec{a}(\omega, \theta) \\
 &= \vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\
 &\quad + \alpha \exp[-j\omega\tau(\theta)] \cdot \vec{h}_d^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) \\
 &\quad + \alpha \exp[j\omega\tau(\theta_s)] \cdot \vec{h}_r^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\
 &\quad + \alpha^2 \exp[-j\omega(\tau(\theta) - \tau(\theta_s))] \cdot \vec{h}_r^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) \quad (19)
 \end{aligned}$$

5 Será evidente a partir de la ecuación (19) que la coherencia $\gamma(\omega, \theta)$ de la ecuación (19) puede ser menor que la coherencia $\gamma_{conv}(\omega, \theta)$ de la técnica convencional de la ecuación (15). Dado que los parámetros ($\psi(\theta)$ y L) que se pueden cambiar reubicando o reorientando el objeto reflector se incluyen en el segundo al cuarto términos de la ecuación (19), hay una posibilidad de que se pueda eliminar el primer término, $\vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta)$.

10 Por ejemplo, si se coloca un reflector plano en tal posición que la dirección a lo largo de la cual se disponen los micrófonos en una agrupación de micrófonos lineales es normal al reflector, $\psi(\theta) = \pi - \theta$ se mantiene para la función $\psi(\theta)$ y la ecuación (20) recoge la diferencia $\tau(\theta)$ en el tiempo de llegada entre un sonido directo y un sonido reflejado. Por lo tanto, las condiciones de las ecuaciones (21) y (22) se generan para los elementos de la ecuación (19). Aquí, el símbolo * es un operador complejo conjugado:

$$\tau(\theta) = \begin{cases} (2L \cos \theta)/c & \left(0 < \theta \leq \frac{\pi}{4} \right) \\ (2L \sin \theta \tan \theta)/c & \left(\frac{\pi}{4} < \theta < \frac{\pi}{2} \right) \end{cases} \quad (20)$$

$$\vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) = \vec{h}_r^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) \quad (21)$$

$$\vec{h}_d^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) = \left[\vec{h}_r^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \right]^* \quad (22)$$

Dado que el valor absoluto de $\vec{h}_d^H(\omega, \theta_s) \vec{h}_r(\omega, \theta)$ es suficientemente menor que $\vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta)$, el segundo y tercer términos de la ecuación (19) son despreciables. Entonces la coherencia $\gamma(\omega, \theta)$ se puede aproximar como la ecuación (23):

$$15 \quad \tilde{\gamma}(\omega, \theta) \approx \left\{ 1 + \alpha^2 \exp[-j\omega(\tau(\theta) - \tau(\theta_s))] \right\} \vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \quad (23)$$

Incluso si $\vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \neq 0$, una coherencia aproximada $\tilde{\gamma}(\omega, \theta)$ tiene una solución mínima θ de la ecuación (24), donde q es un número entero positivo arbitrario. El intervalo de q está restringido para cada frecuencia.

$$\theta = \begin{cases} \arccos\left(\frac{(2q+1)\pi c}{2\omega L} + \cos \theta_s\right) & \left(0 < \theta \leq \frac{\pi}{4} \right) \\ \frac{(2q+1)\pi c}{4\omega L} + \frac{1}{2} \sqrt{\left(\frac{(2q+1)\pi c}{4\omega L}\right)^2 + 4} & \left(\frac{\pi}{4} < \theta < \frac{\pi}{2} \right) \end{cases} \quad (24)$$

20 Es decir, no solamente se puede suprimir la coherencia en una dirección dada por la ecuación (16), sino también la coherencia en una dirección dada por la ecuación (24). Dado que la supresión de la coherencia puede reducir la potencia de ruido, se puede lograr una directividad aguda como se muestra esquemáticamente en la Fig. 5B.

Mientras que las Fig. 5A y 5B muestran esquemáticamente la diferencia entre la directividad lograda por el principio de la técnica de realce de sonido directivo agudo de la presente invención y la directividad lograda por una técnica convencional, la Fig. 6 muestra específicamente la diferencia entre θ dada por la ecuación (16) y θ dada por la ecuación (24). Aquí, $\omega = 2\pi \times 1000$ [rad/s], $L = 0,70$ [m], y $\theta_s = \pi/4$ [rad]. La dependencia de la dirección de la coherencia normalizada se muestra en la Fig. 6 para la comparación entre las técnicas. La dirección indicada por un círculo es θ dada por la ecuación (16) y las direcciones indicadas por el símbolo + son θ dadas por la ecuación (24). Como se puede ver a partir de la Fig. 6, según la técnica convencional, θ que produce una coherencia de 0 para $\theta_s = \pi/4$ [rad] existe solamente en la dirección indicada por el círculo, mientras que según el principio de la mejora de sonido directivo agudo de la presente invención, θ que produce una coherencia de 0 para $\theta_s = \pi/4$ [rad] existe en muchas direcciones indicadas por el símbolo +. Especialmente, las direcciones indicadas por el símbolo + existen mucho más cerca de $\theta_s = \pi/4$ [rad] que la dirección indicada por el círculo. Por lo tanto, se entenderá que la técnica de la presente invención logra una directividad más aguda que la técnica convencional.

Como es evidente a partir de la descripción precedente, la esencia de la técnica de realce de sonido directivo agudo de la presente invención es que la función de transferencia $\vec{a}^-(\omega, \theta) = [a_1(\omega, \theta), \dots, a_M(\omega, \theta)]^T$ se representa por la suma del vector de dirección de un sonido directo y los vectores de dirección de M sonidos reflejados, como se muestra en la Ecuación (17a), por ejemplo. Dado que esto no afecta al concepto de diseño de filtro, los filtros $\vec{W}^-(\omega, \theta_s)$ se pueden diseñar mediante un método distinto de la respuesta sin distorsión de varianza mínima (MVDR).

Se describirán métodos distintos del método de MVDR descrito anteriormente. Son: <1> un método de diseño de filtro basado en el criterio de maximización de SNR, <2> un método de diseño de filtro basado en inversión de potencia, <3> un método de diseño de filtro que usa MVDR con una o más direcciones nulas (direcciones en las que se suprime la ganancia de ruido) como una condición de restricción, <4> un método de diseño de filtro usando conformación de haces de retardo y suma, <5> un método de diseño de filtro que usa el método de máxima verosimilitud y <6> un método de diseño de filtro usando la agrupación adaptativa de micrófonos para reducción de ruido (AMNOR). Para <1> el método de diseño de filtro basado en el criterio de maximización de SNR y <2> el método de diseño de filtro basado en inversión de potencia, se hace referencia a la Referencia 2 enumerada a continuación. Para <3> el método de diseño de filtro que usa MVDR con una o más direcciones nulas (direcciones en las que se suprime la ganancia de ruido) como condición de restricción, se hace referencia a la Referencia 3 enumerada a continuación. Para <6> el método de diseño de filtro que usa el método de la agrupación adaptativa de micrófonos para reducción de ruido (AMNOR), se hace referencia a la Referencia 4 enumerada a continuación.

<1> Método de diseño de filtro basado en el criterio de maximización de SNR

En el método de diseño de filtro basado en el criterio de maximización de SNR, se determina un filtro $\vec{W}^-(\omega, \theta_s)$ sobre la base de un criterio de maximización de la relación SN (SNR) en una dirección del objetivo θ_s . La matriz de correlación espacial para un sonido de la dirección del objetivo θ_s se denota por $R_{ss}(\omega)$ y la matriz de correlación espacial para un sonido desde una dirección distinta de la dirección del objetivo θ_s se denota por $R_{nn}(\omega)$. Entonces la SNR se puede dar por la ecuación (25). Aquí, $R_{ss}(\omega)$ se puede dar por la ecuación (26) y $R_{nn}(\omega)$ se puede dar por la ecuación (27). Las funciones de transferencia $\vec{a}^-(\omega, \theta) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ se pueden dar por la ecuación (17a) (para ser precisos, la ecuación (17a) donde θ se sustituye por θ_s).

$$SNR = \frac{\vec{W}^H(\omega, \theta_s) R_{ss}(\omega) \vec{W}(\omega, \theta_s)}{\vec{W}^H(\omega, \theta_s) R_{nn}(\omega) \vec{W}(\omega, \theta_s)} \quad (25)$$

$$R_{ss}(\omega) = \vec{a}(\omega, \theta_s) \vec{a}^H(\omega, \theta_s) \quad (26)$$

$$R_{nn}(\omega) = \sum_{p \in \{1, \dots, P-1\}} \vec{a}(\omega, \theta_p) \vec{a}^H(\omega, \theta_p) \quad (27)$$

El filtro $\vec{W}^-(\omega, \theta_s)$ que maximiza la SNR de la ecuación (25) se puede obtener estableciendo el gradiente relacionado con el filtro $\vec{W}^-(\omega, \theta_s)$ en cero, es decir, por la ecuación (28).

$$\nabla_{\vec{W}(\omega, \theta_s)} [SNR] = 0 \quad (28)$$

donde

$$\nabla_{\vec{W}(\omega, \theta_s)}[SNR] = \frac{2R_{ss}(\omega)\vec{W}(\omega, \theta_s)(\vec{W}^H(\omega, \theta_s)R_{nn}(\omega)\vec{W}(\omega, \theta_s)) - 2R_{nn}(\omega)\vec{W}(\omega, \theta_s)(\vec{W}^H(\omega, \theta_s)R_{ss}(\omega)\vec{W}(\omega, \theta_s))}{(\vec{W}^H(\omega, \theta_s)R_{nn}(\omega)\vec{W}(\omega, \theta_s))^2}$$

De esta manera, el filtro $W^{\rightarrow}(\omega, \theta_s)$ que maximiza la SNR de la ecuación (25) se puede dar por la ecuación (29):

$$\vec{W}(\omega, \theta_s) = R_{nn}^{-1}(\omega)\vec{a}(\omega, \theta_s) \quad (29)$$

5 La ecuación (29) incluye la matriz inversa de la matriz de correlación espacial $R_{nn}(\omega)$ de un sonido de una dirección distinta de la dirección del objetivo θ_s . Es conocido que la matriz inversa de $R_{nn}(\omega)$ se puede sustituir por la matriz inversa de una matriz de correlación espacial $R_{xx}(\omega)$ de una entrada completa incluyendo sonidos de la dirección del objetivo θ_s y otras direcciones distintas de la dirección del objetivo θ_s . Obsérvese que $R_{xx}(\omega) = R_{ss}(\omega) + R_{nn}(\omega) = Q(\omega)$ (véanse las ecuaciones (10a), (26) y (27)). Es decir, el filtro $W^{\rightarrow}(\omega, \theta_s)$ que maximiza la SNR de la ecuación (25) se puede obtener por la ecuación (30):

$$\vec{W}(\omega, \theta_s) = R_{xx}^{-1}(\omega)\vec{a}(\omega, \theta_s) \quad (30)$$

<2> Método de diseño de filtro basado en inversión de potencia

En el método de diseño de filtro basado en inversión de potencia, un filtro $W^{\rightarrow}(\omega, \theta_s)$ se determina sobre la base de un criterio de minimización de la potencia media de salida de un conformador de haces mientras que un coeficiente de filtro para un micrófono se fija a un valor constante. Aquí, se describirá un ejemplo donde se fija el coeficiente de filtro para el primer micrófono entre M micrófonos. En este método de diseño, se diseña un filtro $W^{\rightarrow}(\omega, \theta_s)$ que minimiza la potencia de los sonidos de todas las direcciones (todas las direcciones desde las cuales pueden llegar sonidos) usando una matriz de correlación espacial $R_{xx}(\omega)$ (véase la ecuación (31)) bajo la condición de restricción de la ecuación (32). Las funciones de transferencia $\vec{a}^{\rightarrow}(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ se pueden dar por la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_s). Aquí, $R_{xx}(\omega) = Q(\omega)$ (véanse las ecuaciones (10a), (26) y (27)).

$$\min_{\vec{W}(\omega, \theta_s)} (\vec{W}^H(\omega, \theta_s)R_{xx}(\omega)\vec{W}(\omega, \theta_s)) \quad (31)$$

$$\vec{W}^H(\omega, \theta_s)\vec{G} = \vec{G}^H R_{xx}^{-1}(\omega)\vec{G} \quad (32)$$

donde

$$\vec{G} = [1, 0, \dots, 0]^T$$

Es conocido que el filtro $W^{\rightarrow}(\omega, \theta_s)$ que es una solución óptima de la ecuación (31) se puede dar por la ecuación (33):

$$\vec{W}(\omega, \theta_s) = R_{xx}^{-1}(\omega)\vec{G} \quad (33)$$

<3> Método de diseño de filtro que usa MVDR con una o más direcciones nulas como condición de restricción

En el método MVDR descrito anteriormente, se ha diseñado un filtro $W^{\rightarrow}(\omega, \theta_s)$ bajo la condición de restricción única que se obtiene un filtro que minimiza la potencia media de salida de un conformador de haces dada por la ecuación (7) (es decir, la potencia de ruido que es sonidos de direcciones distintas de la dirección del objetivo) bajo la condición de restricción que el filtro pasa sonidos de una dirección del objetivo θ_s en todas las bandas de frecuencia como se expresa por la ecuación (8). Según el método, la potencia de ruido se puede suprimir de manera general. Sin embargo, el método no es necesariamente preferible si se conoce previamente que hay una fuente o fuentes de ruido que tienen una potencia fuerte en una o más direcciones particulares. Si este es el caso, se requiere un filtro que suprima intensamente una o más direcciones conocidas particulares (es decir, direcciones nulas) en las que existe la fuente o fuentes de ruido. Por lo tanto, el método de diseño de filtro descrito aquí obtiene un filtro que minimiza la potencia media de salida del conformador de haces dada por la ecuación (7) (es decir, minimiza la potencia media de salida de sonidos de direcciones distintas a una dirección del objetivo y las direcciones nulas) bajo las condiciones de restricción que (1) el filtro pasa sonidos desde la dirección del objetivo θ_s en todas las bandas de frecuencia y que (2) el filtro suprime sonidos de B direcciones nulas conocidas $\theta_{N1}, \theta_{N2}, \dots, \theta_{NB}$ (B es un número entero predeterminado mayor o igual a 1) en todas las bandas de frecuencia. Permitamos que un conjunto

de índices Φ de direcciones desde las cuales llega el sonido sean denotados por $\{1, 2, \dots, P\}$, entonces $N_j \in \{1, 2, \dots, P\}$ (donde $j \in \{1, 2, \dots, B\}$) y $B \leq P-1$, como se ha descrito anteriormente.

5 Permitamos que $\vec{a}^-(\omega, \theta_i) = [a_1(\omega, \theta_i), \dots, a_M(\omega, \theta_i)]^T$ sean funciones de transferencia entre una fuente de sonido que se supone que está situada en una dirección θ_i y los M micrófonos a una frecuencia ω , en otras palabras, funciones de transferencia de un sonido de una dirección θ_i a una frecuencia ω que llega a los micrófonos de una agrupación de micrófonos, entonces una condición de restricción se puede dar por la ecuación (34). Aquí, los índices $i \in \{s, N1, N2, \dots, NB\}$, las funciones de transferencia $\vec{a}^-(\omega, \theta_i) = [a_1(\omega, \theta_i), \dots, a_M(\omega, \theta_i)]^T$ se puede dar por la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_i), y $f_i(\omega)$ representa una característica de paso a una frecuencia ω para una dirección θ_i .

$$10 \quad \vec{W}^H(\omega, \theta_s) \vec{a}(\omega, \theta_i) = f_i(\omega) \quad i \in \{s, N1, N2, \dots, NB\} \quad (34)$$

La ecuación (34) se puede representar como una matriz, por ejemplo como la ecuación (35). Aquí $\vec{A}^-(\omega, \theta_s) = [\vec{a}^-(\omega, \theta_s), \vec{a}^-(\omega, \theta_{N1}), \dots, \vec{a}^-(\omega, \theta_{NB})]$

$$\vec{W}^H(\omega, \theta_s) \vec{A}(\omega, \theta_s) = \vec{F} \quad (35)$$

donde

$$15 \quad \vec{F} = [f_s(\omega), f_{N1}(\omega), \dots, f_{NB}(\omega)]$$

Teniendo en consideración las condiciones de restricción que (1) el filtro pasa sonidos de la dirección del objetivo θ_s en todas las bandas de frecuencia y que (2) el filtro suprime sonidos de B direcciones nulas conocidas $\theta_{N1}, \theta_{N2}, \dots, \theta_{NB}$ en todas las bandas de frecuencia, se deberían establecer idealmente $f_s(\omega) = 1,0$ y $f_i(\omega) = 0,0$ ($i \in \{N1, N2, \dots, NB\}$). Esto significa que el filtro pasa por completo sonidos en todas las bandas de frecuencia de la dirección del objetivo θ_s y bloquea por completo sonidos en todas las bandas de frecuencia de B direcciones nulas conocidas $\theta_{N1}, \theta_{N2}, \dots, \theta_{NB}$. En realidad, sin embargo, es difícil en algunas situaciones efectuar tal control como que pasa por completo todas las bandas de frecuencia o que bloquea por completo todas las bandas de frecuencia. En tal caso, el valor absoluto de $f_s(\omega)$ se establece en un valor cercano a 1,0 y el valor absoluto de $f_i(\omega)$ ($i \in \{N1, N2, \dots, NB\}$) se establece en un valor cerca de 0,0. Por supuesto, $f_i(\omega)$ y $f_j(\omega)$ ($i \neq j$; i y $j \in \{N1, N2, \dots, NB\}$) pueden ser iguales o diferentes.

Según el método de diseño de filtro descrito aquí, el filtro $\vec{W}^-(\omega, \theta_s)$ que es una solución óptima de la ecuación (7) bajo la condición de restricción dada por la ecuación (35) se puede dar por la ecuación (36) (véase la Referencia 3 enumerada a continuación).

$$25 \quad \vec{W}^-(\omega, \theta_s) = Q^{-1}(\omega) \vec{A}(\omega, \theta_s) \left(\vec{A}^H(\omega, \theta_s) Q^{-1}(\omega) \vec{A}(\omega, \theta_s) \right)^{-1} \vec{F} \quad (36)$$

30 <4> Método de diseño de filtro que usa conformación de haces de retardo y suma

Como es evidente a partir de la ecuación (2), suponiendo que los sonidos directos y reflejados que llegan son ondas planas, entonces un filtro $\vec{W}^-(\omega, \theta_s)$ se puede dar por la ecuación (37). Es decir, el filtro $\vec{W}^-(\omega, \theta_s)$ se puede obtener normalizando una función de transferencia $\vec{a}^-(\omega, \theta_s)$. La función de transferencia $\vec{a}^-(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ se puede dar por la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_s). El método de diseño de filtro no logra necesariamente una alta precisión de filtrado sino que requiere solamente una pequeña cantidad de cálculo.

$$35 \quad \vec{W}^-(\omega, \theta_s) = \frac{\vec{a}(\omega, \theta_s)}{\vec{a}^H(\omega, \theta_s) \vec{a}(\omega, \theta_s)} \quad (37)$$

<5> Método de diseño de filtro que usa el método de máxima verosimilitud

Excluyendo la información espacial que se refiere a los sonidos de una dirección del objetivo a partir de una matriz de correlación espacial $Q(\omega)$ en el método MVDR descrito anteriormente, se puede mejorar la flexibilidad de supresión de ruido y se puede suprimir además la potencia de ruido. Por lo tanto, en el método de diseño de filtro descrito aquí, la matriz de correlación espacial $Q(\omega)$ se escribe como el segundo término del lado derecho de la ecuación (10a), es decir, la ecuación (10c). Un filtro $\vec{W}^-(\omega, \theta_s)$ se puede dar por la ecuación (9) o (36). Aquí, $Q(\omega)$ incluida en la ecuación (9) y (36) o $R_{xx}(\omega) = Q(\omega)$ incluida en las ecuaciones (30) y (33) es una matriz de correlación espacial dada por la ecuación (10c).

$$Q(\omega) = \sum_{p \in \{1, \dots, P-1\}} \vec{a}(\omega, \theta_p) \vec{a}^H(\omega, \theta_p) \quad (10c)$$

<6> Método de diseño de filtro que usa el método AMNOR

El método AMNOR obtiene un filtro que permite alguna cantidad de descomposición D de un sonido de una dirección del objetivo mediante una solución de compromiso de la cantidad de descomposición D del sonido de la dirección del objetivo frente a la potencia de ruido restante en una señal de salida de filtro (por ejemplo, la cantidad de descomposición D se mantiene en un cierto umbral D^\wedge o menos) y, cuando se introduce una señal mixta de [a] una señal producida aplicando las funciones de transferencia entre una fuente de sonido y micrófonos a una señal virtual de una dirección del objetivo (en lo sucesivo, conocida como la señal virtual del objetivo) y [b] ruido (obtenido mediante la observación con M micrófonos en un entorno ruidoso sin un sonido de la dirección del objetivo), emite una señal de salida de filtro que reproduce mejor la señal virtual del objetivo en términos de error cuadrático mínimo (es decir, se minimiza la potencia de ruido contenida en una señal de salida de filtro). Según el método AMNOR, un filtro $W^\rightarrow(\omega, \theta_s)$ se puede dar por la ecuación (38) (véase la Referencia 4 enumerada a continuación). Aquí, $R_{ss}(\omega)$ se puede dar por la ecuación (26) y $R_{nn}(\omega)$ se puede dar por la ecuación (27). Las funciones de transferencia $\vec{a}^\rightarrow(\omega, \theta) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ se pueden dar por la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_s).

$$\vec{W}(\omega, \theta_s) = P_s \vec{a}(\omega, \theta_s) (R_{nn}(\omega) + P_s R_{ss}(\omega))^{-1} \quad (38)$$

P_s es un coeficiente que asigna una ponderación al nivel de la señal virtual del objetivo y denominado el nivel de señal virtual del objetivo. El nivel de señal virtual del objetivo P_s es una constante que no es dependiente de las frecuencias. El nivel de señal virtual del objetivo P_s se puede determinar empíricamente o se puede determinar de modo que la diferencia entre la cantidad de descomposición D de un sonido de la dirección del objetivo y el umbral D^\wedge esté dentro de un margen de error predeterminado arbitrariamente. Se describirá este último caso. La respuesta de frecuencia $F(\omega)$ del filtro $W^\rightarrow(\omega, \theta_s)$ a un sonido de una dirección del objetivo θ_s en el método AMNOR se puede dar por la ecuación (39). Permitamos que la cantidad de descomposición $D(P_s)$ cuando se usa el filtro $W^\rightarrow(\omega, \theta_s)$ dado por la ecuación (38) sea denotada por $D(P_s)$, entonces la cantidad de descomposición $D(P_s)$ se puede definir por la ecuación (40). Aquí, ω_0 representa el límite superior de la frecuencia ω (típicamente, una frecuencia más alta adyacente a una frecuencia discreta ω). La cantidad de descomposición $D(P_s)$ es una función monótonamente decreciente de P_s . Por lo tanto, un nivel de señal virtual del objetivo P_s tal que la diferencia entre la cantidad de descomposición $D(P_s)$ y el umbral D^\wedge está dentro de un margen de error predeterminado arbitrariamente se puede obtener obteniendo repetidamente la cantidad de descomposición $D(P_s)$ mientras se cambia P_s con la monotonía de $D(P_s)$.

$$F(\omega) = \vec{W}^H(\omega, \theta_s) \vec{a}(\omega, \theta_s) \quad (39)$$

$$D(P_s) = \frac{1}{2\omega_0} \int_{-\omega_0}^{\omega_0} |1 - F(\omega)|^2 d\omega \quad (40)$$

<Variación>

En la descripción precedente, las matrices de correlación espacial $Q(\omega)$, $R_{ss}(\omega)$ y $R_{nn}(\omega)$ se expresan usando funciones de transferencia. Sin embargo, las matrices de correlación espacial $Q(\omega)$, $R_{ss}(\omega)$ y $R_{nn}(\omega)$ también se pueden expresar usando las señales en el dominio de frecuencia $X^\rightarrow(\omega, k)$ descritas anteriormente. Mientras que la matriz de correlación espacial $Q(\omega)$ se describirá a continuación, la siguiente descripción también se aplica a $R_{ss}(\omega)$ y $R_{nn}(\omega)$. ($Q(\omega)$ se puede sustituir por $R_{ss}(\omega)$ y $R_{nn}(\omega)$). La matriz de correlación espacial $R_{ss}(\omega)$ se puede obtener usando representaciones en el dominio de frecuencia de señales analógicas obtenidas mediante observación con una agrupación de micrófonos (que incluye M micrófonos) en un entorno donde solamente existen sonidos de una dirección del objetivo. La matriz de correlación espacial $R_{nn}(\omega)$ se puede obtener usando representaciones en el dominio de frecuencia de una señal analógica obtenida mediante observación con una agrupación de micrófonos (que incluye M micrófonos) en un entorno donde no existen sonidos de la dirección del objetivo (es decir, un entorno ruidoso).

La matriz de correlación espacial $Q(\omega)$ que usa señales en el dominio de frecuencia $X^\rightarrow(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ se puede dar por la ecuación (41). Aquí, el operador $E[\cdot]$ representa una operación de promediado estadístico. Cuando se ve una serie de tiempo discreta de una señal analógica recibida con una agrupación de micrófonos (que incluye M micrófonos) como un proceso estocástico, el operador $E[\cdot]$ representa una operación de valor de media aritmética (valor esperado) si el proceso estocástico es un denominado proceso estacionario en sentido amplio o un proceso estacionario de segundo orden. En este caso, la matriz de correlación espacial $Q(\omega)$ se puede dar por la

ecuación (42) usando señales en el dominio de frecuencia $X^{\rightarrow}(\omega, k - i)$ ($i = 0, 1, \dots, \zeta - 1$) de un total de ζ tramas actuales y pasadas almacenadas en una memoria, por ejemplo. Cuando $i = 0$, una trama de orden k es la trama actual. Obsérvese que la matriz de correlación espacial $Q(\omega)$ dada por la ecuación (41) o (42) se puede recalcul para cada trama o se puede calcular a intervalos regulares o irregulares, o se puede calcular antes de la implementación de una realización, que se describirá más tarde (especialmente cuando se usa $R_{ss}(\omega)$ o $R_{nn}(\omega)$ en el diseño de filtro, la matriz de correlación espacial $Q(\omega)$ se calcula preferiblemente de antemano usando señales en el dominio de frecuencia obtenidas antes de la implementación de la realización). Si la matriz de correlación espacial $Q(\omega)$ se recalcula para cada trama, la matriz de correlación espacial $Q(\omega)$ depende de las tramas actuales y pasadas y, por lo tanto, la matriz de correlación espacial se representará explícitamente como $Q(\omega, k)$ como en las ecuaciones (41a) y (42a).

$$Q(\omega) = E[\vec{X}(\omega, k)\vec{X}^H(\omega, k)] \quad (41)$$

$$Q(\omega) = \sum_{i=0}^{\zeta-1} \vec{X}(\omega, k-i)\vec{X}^H(\omega, k-i) \quad (42)$$

$$Q(\omega, k) = E[\vec{X}(\omega, k)\vec{X}^H(\omega, k)] \quad (41a)$$

$$Q(\omega, k) = \sum_{i=0}^{\zeta-1} \vec{X}(\omega, k-i)\vec{X}^H(\omega, k-i) \quad (42a)$$

Si se usa la matriz de correlación espacial $Q(\omega, k)$ representada por la ecuación (41a) o (42a), el filtro $W^{\rightarrow}(\omega, \theta_s)$ también depende de las tramas actuales y pasadas y, por lo tanto, se representa explícitamente como $W^{\rightarrow}(\omega, \theta_s, k)$. Entonces, un filtro $W^{\rightarrow}(\omega, \theta_s)$ representado por cualquiera de las ecuaciones (9), (29), (30), (33), (36) y (38) descritas con los métodos de diseño de filtro descritos anteriormente se reescribe como las ecuaciones (9m), (29m), (30m), (33m), (36m) o (38m).

$$\vec{W}(\omega, \theta_s, k) = \frac{Q^{-1}(\omega, k)\vec{a}(\omega, \theta_s)}{\vec{a}^H(\omega, \theta_s)Q^{-1}(\omega, k)\vec{a}(\omega, \theta_s)} \quad (9m)$$

$$\vec{W}(\omega, \theta_s, k) = R_{nn}^{-1}(\omega, k)\vec{a}(\omega, \theta_s) \quad (29m)$$

$$\vec{W}(\omega, \theta_s, k) = R_{xx}^{-1}(\omega, k)\vec{a}(\omega, \theta_s) \quad (30m)$$

$$\vec{W}(\omega, \theta_s, k) = R_{xx}^{-1}(\omega, k)\vec{G} \quad (33m)$$

$$\vec{W}(\omega, \theta_s, k) = Q^{-1}(\omega, k)\vec{A}(\omega, \theta_s)\left(\vec{A}^H(\omega, \theta_s)Q^{-1}(\omega, k)\vec{A}(\omega, \theta_s)\right)^{-1}\vec{F} \quad (36m)$$

$$\vec{W}(\omega, \theta_s, k) = P_s\vec{a}(\omega, \theta_s)(R_{nn}(\omega, k) + P_sR_{ss}(\omega, k))^{-1} \quad (38m)$$

<<Primera realización de la técnica de realce de sonido directivo agudo>>

Las Fig. 7 y 8 ilustran una configuración funcional y un flujo de proceso de una primera realización de una técnica de realce de sonido directivo agudo de la presente invención. Un aparato 1 de realce de sonido de la primera realización (en lo sucesivo conocido como el aparato de realce de sonido directivo agudo) incluye un convertidor 210 AD, un generador 220 de tramas, una sección 230 de transformación en el dominio de frecuencia, una sección 240 de aplicación de filtro, una sección 250 de transformación en el dominio de tiempo, una sección 260 de diseño de filtro y almacenamiento 290.

[Paso S1]

La sección 260 de diseño de filtro calcula de antemano un filtro $W^{\rightarrow}(\omega, \theta_i)$ para cada frecuencia para cada una de las direcciones discretas a partir de las cuales pueden llegar los sonidos a ser realzados. La sección 260 de diseño de filtro calcula los filtros $W^{\rightarrow}(\omega, \theta_1), \dots, W^{\rightarrow}(\omega, \theta_i), \dots, W^{\rightarrow}(\omega, \theta_l)$ ($1 \leq i \leq l, \omega \in \Omega, i$ es un número entero y Ω es un conjunto de frecuencias ω), donde l es el número total de direcciones discretas a partir de las cuales pueden llegar los sonidos a ser realzados (l es un número entero predeterminado mayor o igual a 1 y satisface $l \leq P$).

Para hacerlo así, las funciones de transferencia $a^{-}(\omega, \theta_i) = [a_1(\omega, \theta_i), \dots, a_M(\omega, \theta_i)]^T$ ($1 \leq i \leq I, \omega \in \Omega$) necesitan ser obtenidas excepto para el caso de <Variación> descrito anteriormente. La función de transferencia $a^{-}(\omega, \theta_i) = [a_1(\omega, \theta_i), \dots, a_M(\omega, \theta_i)]^T$ se puede calcular prácticamente según la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_i) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos y la información de entorno tal como la relación de posición de los objetos reflectores tales como un reflector, el suelo, las paredes o el techo, a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y el sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y la reflectancia acústico del objeto reflector. Obsérvese que si se usa <3> el método de diseño de filtro que usa MVDR con una o más direcciones nulas como condición de restricción, los índices i de las direcciones usadas para calcular las funciones de transferencia $a^{-}(\omega, \theta_i)$ ($1 \leq i \leq I, \omega \in \Omega$) preferiblemente cubren todos los índices N1, N2, ..., NB de direcciones de al menos B direcciones nulas. En otras palabras, los índices N1, N2, ..., NB de las direcciones de B direcciones nulas se establecen en cualquiera de los diferentes números enteros mayores o iguales a 1 y menores o iguales a I.

El número Ξ de sonidos reflejados se establece en un número entero que satisface $1 \leq \Xi$. El número Ξ no está limitado y se puede establecer en un valor adecuado según la capacidad de cálculo y otros factores. Si se coloca un reflector cerca de la agrupación de micrófonos, las funciones de transferencia $a^{-}(\omega, \theta_i)$ se pueden calcular prácticamente según la ecuación (17b) (para ser precisos, por la ecuación (17b) donde θ se sustituye por θ_i).

Para calcular vectores de dirección, se pueden usar las ecuaciones (14a), (14b), (18a), (18b), (18d) o (18d), por ejemplo. Obsérvese que las funciones de transferencia obtenidas mediante mediciones reales en un entorno real, por ejemplo, se pueden usar para diseñar los filtros en lugar de usar las ecuaciones (17a) y (17b).

Entonces, $W^{-}(\omega, \theta_i)$ ($1 \leq i \leq I$) se obtiene según cualquiera de las ecuaciones (9), (29), (30), (33), (36), (37) y (38), por ejemplo, usando las funciones de transferencia $a^{-}(\omega, \theta_i)$, excepto para el caso descrito en <Variación>. Obsérvese que si se usa la ecuación (9), (30), (33) o (36), la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (10b), excepto para el caso descrito con respecto a <5> el método de diseño de filtro que usa el método de máxima verosimilitud. Si la ecuación (9), (30), (33) o (36) se usa según <5> el método de diseño de filtro que usa el método de máxima verosimilitud descrito anteriormente, la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (10c). Si se usa la ecuación (29), la matriz de correlación espacial $R_{nn}(\omega)$ se puede calcular según la ecuación (27). Los $I \times |\Omega|$ filtros $W^{-}(\omega, \theta_i)$ ($1 \leq i \leq I, \omega \in \Omega$) se almacenan en el almacenamiento 290, donde $|\Omega|$ representa el número de elementos del conjunto Ω .

[Paso S2]

Los M micrófonos 200-1, ..., 200-M que componen la agrupación de micrófonos se usan para captar sonidos, donde M es un número entero mayor o igual a 2.

No hay ninguna restricción en la disposición de los M micrófonos. Sin embargo, una disposición bidimensional o tridimensional de los M micrófonos tiene la ventaja de eliminar la incertidumbre de una dirección desde la cual llegan los sonidos a ser realzados. Es decir, una disposición plana o esférica de los micrófonos puede evitar el problema con una disposición lineal horizontal de los M micrófonos de que un sonido que llega desde una dirección frontal no se puede distinguir de un sonido que llega desde arriba a la derecha, por ejemplo. Con el fin de proporcionar un intervalo amplio de direcciones que se pueden establecer como direcciones de captura de sonido, cada micrófono tiene preferiblemente una directividad capaz de captar sonidos con un cierto nivel de presión de sonido en direcciones del objetivo θ_s potenciales que son direcciones de captura de sonido. Por consiguiente, son preferibles los micrófonos que tienen una directividad relativamente débil, tales como micrófonos omnidireccionales o micrófonos unidireccionales.

[Paso S3]

El convertidor 210 AD convierte las señales analógicas (señales de captura) captadas con los M micrófonos 200-1, ..., 200-M en señales digitales $x^{-}(t) = [x_1(t), \dots, x_M(t)]^T$, donde t representa el índice de un tiempo discreto.

[Paso S4]

El generador 220 de tramas toma entradas de las señales digitales $x^{-}(t) = [x_1(t), \dots, x_M(t)]^T$ emitidas desde el convertidor 210 AD, almacena N muestras en un almacenador temporal sobre una base canal por canal, y emite señales digitales $x^{-}(k) = [x_1^{-}(k), \dots, x_M^{-}(k)]^T$ en tramas, donde k es un índice de un número de tiempo de trama y $x_m^{-}(k) = [x_m((k-1)N+1), \dots, x_m(kN)]$ ($1 \leq m \leq M$). N depende de la frecuencia de muestreo y 512 es adecuado para muestrear a 16 kHz.

[Paso S5]

La sección 230 de transformación en el dominio de frecuencia transforma las señales digitales $x^{-}(k)$ en tramas en señales en el dominio de frecuencia $X^{-}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y emite las señales en el dominio de frecuencia, donde ω es un índice de una frecuencia discreta. Una forma de transformar una señal en el dominio de

tiempo en una señal de dominio de frecuencia es una transformada de Fourier discreta rápida. Sin embargo, la forma de transformar la señal no está limitada a esto. Se puede usar otro método para transformar a una señal en el dominio de frecuencia. La señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ se emite para cada frecuencia ω y trama k a la vez.

5 [Paso S6]

La sección 240 de aplicación de filtro aplica el filtro $W^{\rightarrow}(\omega, \theta_s)$ correspondiente a una dirección del objetivo θ_s a ser realizada a la señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ en cada trama k para cada frecuencia $\omega \in \Omega$ y emite una señal de salida $Y(\omega, k, \theta_s)$ (véase la ecuación (43)). El índice s de la dirección del objetivo θ_s es $s \in \{1, \dots, 1\}$ y los filtros $W^{\rightarrow}(\omega, \theta_s)$ se almacenan en el almacenamiento 290. Por lo tanto, la sección 240 de aplicación de filtro solamente tiene que recuperar el filtro $W^{\rightarrow}(\omega, \theta_s)$ que corresponde a la dirección del objetivo θ_s a ser realizada desde el almacenamiento 290. Si el índice s de la dirección del objetivo θ_s no pertenece al conjunto $\{1, \dots, 1\}$, es decir, el filtro $W^{\rightarrow}(\omega, \theta_s)$ que corresponde a la dirección del objetivo θ_s no se ha calculado en el proceso en el paso S1, la sección 260 de diseño de filtro puede calcular en este momento el filtro $W^{\rightarrow}(\omega, \theta_s)$ que corresponde a la dirección del objetivo θ_s o se puede usar un filtro $W^{\rightarrow}(\omega, \theta_s)$ que corresponde a una dirección θ_s cercana a la dirección del objetivo θ_s .

$$Y(\omega, k, \theta_s) = \vec{W}^H(\omega, \theta_s) \vec{X}(\omega, k) \quad \forall \omega \in \Omega \quad (43)$$

[Paso S7]

La sección 250 de transformación en el dominio de tiempo transforma la señal de salida $Y(\omega, k, \theta_s)$ de cada frecuencia $\omega \in \Omega$ en una trama de orden k a un dominio de tiempo para obtener una señal $y(k)$ de trama en el dominio de tiempo en la trama de orden k , entonces combina las señales $y(k)$ en el dominio de tiempo de tramas obtenidas del orden del índice de número de tiempo de trama, y emite una señal $y(t)$ en el dominio de tiempo en la cual se realiza el sonido de la dirección del objetivo θ_s . El método para transformar una señal en el dominio de frecuencia a una señal en el dominio de tiempo es una transformación inversa de la transformada usada en el proceso en el paso S5 y puede ser una transformada de Fourier inversa discreta rápida, por ejemplo.

25 Mientras que la primera realización se ha descrito aquí en la que los filtros $W^{\rightarrow}(\omega, \theta_i)$ se calculan de antemano en el proceso en el paso S1, la sección 260 de diseño de filtro puede calcular el filtro $W^{\rightarrow}(\omega, \theta_i)$ para cada frecuencia después de que se determina la dirección del objetivo θ_s , dependiendo de la capacidad de cálculo del aparato 1 de realce de sonido directivo agudo.

<<Segunda realización de la técnica de realce de sonido directivo agudo>>

30 Las Fig. 10 y 11 ilustran una configuración funcional y un flujo de proceso de una segunda realización de una técnica de realce de sonido directivo agudo de la presente invención. Un aparato 2 de realce de sonido directivo agudo de la segunda realización incluye un convertidor 210 AD, una generador 220 de tramas, una sección 230 de transformación en el dominio de frecuencia, una sección 240 de aplicación de filtro, una sección 250 de transformación en el dominio de tiempo, una sección 261 de cálculo de filtro y un almacenamiento 290.

35 [Paso S11]

M micrófonos 200-1, ..., 200-M que componen una agrupación de micrófonos se usa para captar sonidos, donde M es un número entero mayor o igual a 2. La disposición de los M micrófonos es como se describe en la primera realización.

[Paso S12]

40 El convertidor 210 AD convierte señales analógicas (señales de captura) captadas con los M micrófonos 200-1, ..., 200-M en señales digitales $x^{\rightarrow}(t) = [x_1(t), \dots, x_M(t)]^T$, donde t representa el índice de un tiempo discreto.

[Paso S13]

45 El generador 220 de tramas toma entradas de las señales digitales $x^{\rightarrow}(t) = [x_1(t), \dots, x_M(t)]^T$ emitidas desde el convertidor 210 AD, almacena N muestras en un almacenador temporal sobre una base canal por canal, y emite señales digitales $x^{\rightarrow}(k) = [x^{\rightarrow}_1(k), \dots, x^{\rightarrow}_M(k)]^T$ en tramas, donde k es un índice de un número de tiempo de trama y $x^{\rightarrow}_m(k) [x_m((k-1)N+1), \dots, x_m(kN)]$ ($1 \leq m \leq M$). N depende de la frecuencia de muestreo y 512 es adecuado para muestrear a 16 kHz.

[Paso S14]

50 La sección 230 de transformación en el dominio de frecuencia transforma las señales digitales $x^{\rightarrow}(k)$ en tramas a señales en el dominio de frecuencia $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y emite las señales en el dominio de frecuencia, donde ω es un índice de una frecuencia discreta. Una forma de transformar una señal en el dominio de

tiempo en una señal en el dominio de frecuencia es la transformada de Fourier discreta rápida. Sin embargo, la forma de transformar la señal no está limitada a esto. Se puede usar otro método para transformar una señal en el dominio de frecuencia. La señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ se emite para cada frecuencia ω y trama k a la vez.

5 [Paso S15]

La sección 261 de cálculo de filtro calcula el filtro $W^{\rightarrow}(\omega, \theta_s, k)$ ($\omega \in \Omega$; Ω es un conjunto de frecuencias ω) que corresponde a la dirección del objetivo θ_s a ser usada en un trama actual de orden k .

Para hacerlo así, necesitan ser proporcionadas funciones de transferencia $a^{\rightarrow}(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ ($\omega \in \Omega$). Las funciones de transferencia $a^{\rightarrow}(\omega, \theta_s) = [a_1(\omega, \theta_s), \dots, a_M(\omega, \theta_s)]^T$ se puede calcular prácticamente según la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_s) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos e información de entorno tal como la relación de posición de objetos reflectores tales como un reflector, suelo, paredes o techo a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$), y la reflectancia acústica del objeto reflector. Obsérvese que si se usa $\langle 3 \rangle$ el método de diseño de filtro que usa MVDR con una o más direcciones nulas como condición de restricción, también necesitan ser obtenidas las funciones de transferencia $a^{\rightarrow}(\omega, \theta_{Nj})$ ($1 \leq j \leq B$, $\omega \in \Omega$). Las funciones de transferencia se pueden calcular prácticamente según la ecuación (17a) (para ser precisos, por la ecuación (17a) donde θ se sustituye por θ_{Nj}) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos y la información de entorno, tal como la relación de posición de objetos reflectores, tales como un reflector, un piso, una pared o el techo a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$), y la reflectancia acústica del objeto reflector.

El número Ξ de sonidos reflejados se establece en un número entero que satisface $1 \leq \Xi$. El número Ξ no está limitado y se puede establecer en un valor adecuado según la capacidad de cálculo y otros factores. Si un reflector se coloca cerca de la agrupación de micrófonos, las funciones de transferencia $a^{\rightarrow}(\omega, \theta_s)$ se pueden calcular prácticamente según la ecuación (17b) (para ser precisos, por la ecuación (17b) donde θ se sustituye por θ_s). En este caso, las funciones de transferencia $a^{\rightarrow}(\omega, \theta_{Nj})$ ($1 \leq j \leq B$, $\omega \in \Omega$) se pueden calcular prácticamente según la ecuación (17b) (para ser precisos, por la ecuación (17b) donde θ se sustituye por θ_{Nj}).

Para calcular los vectores de dirección, se pueden usar las ecuaciones (14a), (14b), (18a), (18b), (18c) o (18d), por ejemplo. Obsérvese que las funciones de transferencia obtenidas mediante mediciones reales en un entorno real, por ejemplo, se pueden usar para diseñar los filtros en lugar de usar las ecuaciones (17a) y (17b).

Entonces, la sección 261 de cálculo de filtro calcula los filtros $W^{\rightarrow}(\omega, \theta_s, k)$ ($\omega \in \Omega$) según cualquiera de las ecuaciones (9m), (29m) m (30m), (33m), (36m) y (38m) usando las funciones de transferencia $a^{\rightarrow}(\omega, \theta_s)$ ($\omega \in \Omega$) y, si se necesita, las funciones de transferencia $a^{\rightarrow}(\omega, \theta_{Nj})$ ($1 \leq j \leq B$, $\omega \in \Omega$). Obsérvese que la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (41a) o (42a). En el cálculo de la matriz de correlación espacial $Q(\omega)$, se usan las señales en el dominio de frecuencia $X^{\rightarrow}(\omega, k - i)$ ($i = 0, 1, \dots, \zeta - 1$) de un total de ζ tramas actuales y pasadas almacenadas en el almacenamiento 290, por ejemplo.

[Paso S16]

La sección 240 de aplicación de filtro aplica el filtro $W^{\rightarrow}(\omega, \theta_s, k)$ correspondiente a una dirección del objetivo θ_s a ser realizada a la señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ en cada trama k para cada frecuencia $\omega \in \Omega$ y emite una señal de salida $Y(\omega, k, \theta_s)$ (véase la ecuación (44)).

$$Y(\omega, k, \theta_s) = \vec{W}^H(\omega, \theta_s, k) \vec{X}(\omega, k) \quad \forall \omega \in \Omega \quad (44)$$

[Paso S17]

La sección 250 de transformada en el dominio de tiempo transforma la señal de salida $Y(\omega, k, \theta_s)$ de cada frecuencia $\omega \in \Omega$ de un trama de orden k a un dominio de tiempo para obtener una señal $y(k)$ de trama en el dominio de tiempo en la trama de orden k , entonces combina las señales $y(k)$ en el dominio de tiempo de tramas obtenidas del orden del índice de número de tiempo de trama, y emite una señal $y(t)$ en el dominio de tiempo en el que se realiza el sonido de la dirección del objetivo θ_s . El método para transformar una señal en el dominio de frecuencia en una señal en el dominio de tiempo es una transformación inversa del método de transformación usado en el proceso en el paso S14 y puede ser una transformada de Fourier inversa discreta rápida, por ejemplo.

50

[Ejemplo experimental de técnica de realce de sonido directivo agudo]

Se describirán los resultados de un experimento sobre la primera realización de la técnica de realce de sonido directivo agudo de la presente invención (el método de respuesta sin distorsión de varianza mínima (MVDR) bajo una única condición de restricción). Como se ilustra en la Fig. 9, se disponen linealmente 24 micrófonos y se coloca un reflector 300 de modo que la dirección a lo largo de la cual los micrófonos en la agrupación de micrófonos lineales es normal al reflector 300. Aunque no hay restricción en la forma del reflector 300, se usó un reflector plano rígido semigruoso que tenía un tamaño de 1,0 m x 1,0 m. La distancia entre micrófonos adyacentes era 4 cm y la reflectancia α del reflector 300 era 0,8. Una dirección del objetivo θ_s se estableció en 45 grados. Bajo la suposición de que los sonidos llegarían a la agrupación de micrófonos lineales como ondas planas, se calcularon las funciones de transferencia según la ecuación (17b) (véanse la ecuaciones (14a) y (18a)) y se investigaron las directividades de los filtros generados. Se usaron dos métodos convencionales (el método MVDR sin reflector y el método de conformación de haces de retardo y suma con reflector) para comparación con la técnica.

Las Fig. 12 y 13 muestran los resultados del experimento. Se puede ver que la primera realización de la técnica de realce de sonido directivo agudo de la presente invención puede lograr una directividad aguda en la dirección del objetivo en todas las bandas de frecuencia en comparación con los dos métodos convencionales. Se entenderá que la técnica de realce de sonido directivo agudo es eficaz especialmente en bandas de frecuencias más bajas. La Fig. 14 muestra la directividad de los filtros $W^-(\omega, \theta)$ generados según la primera realización de la técnica de realce de sonido directivo agudo de la presente invención. Se puede ver a partir de la Fig. 14 que la técnica realiza no solamente los sonidos directos sino también los sonidos reflejados.

El mismo experimento se dirigió con el reflector 300 colocado de modo que la superficie plana del reflector 300 formase un ángulo de 45 grados con la dirección en la que estaban dispuestos los micrófonos de la agrupación de micrófonos lineales, como se muestra en la Fig. 15. Una dirección del objetivo θ_s se estableció a 22,5 grados. Las otras condiciones experimentales fueron las mismas que las del experimento en el que el reflector 300 se colocó de modo que la dirección en la que se dispusieron los micrófonos de la agrupación de micrófonos lineales era normal al reflector 300.

Las Fig. 16 y 17 muestran los resultados del experimento. Se puede ver que la primera realización de la técnica de realce de sonido directivo agudo de la presente invención puede lograr una directividad aguda en la dirección del objetivo en todas las bandas de frecuencia en comparación con los dos métodos convencionales. Se entenderá que la técnica de realce de sonido directivo agudo es eficaz especialmente en bandas de frecuencias más bajas.

<Ejemplos de aplicaciones>

Hablando en sentido figurado, la técnica de realce de sonido directivo agudo es equivalente a la generación de una imagen clara a partir de una imagen borrosa y desenfocada y es útil para obtener información detallada acerca de un campo acústico. La siguiente es una descripción de ejemplos de servicios donde es útil la técnica de realce de sonido directivo agudo de la presente invención.

Un primer ejemplo es la creación de contenidos que son combinación de audio y video. El uso de una realización de la técnica de realce de sonido directivo agudo de la presente invención permite que el sonido del objetivo desde una gran distancia sea realizado claramente incluso en un entorno ruidoso con sonidos de ruido (sonidos distintos de los sonidos del objetivo). Por lo tanto, por ejemplo se pueden añadir a la imagen en movimiento sonidos en un área particular correspondiente a una imagen en movimiento acercada de un jugador de fútbol que dribla que fue filmada desde el exterior del campo.

Un segundo ejemplo es una aplicación a una videoconferencia (o una teleconferencia de audio). Cuando se mantiene una conferencia en una sala pequeña, la voz de un orador humano se puede realzar hasta cierto grado con varios micrófonos según una técnica convencional. Sin embargo, en una sala de conferencias grande (por ejemplo, un gran espacio donde hay oradores humanos a una distancia de 5 metros o más de los micrófonos), es difícil realzar claramente la voz de un orador humano a una distancia con las técnicas convencionales mediante el método convencional y necesita ser colocado un micrófono enfrente de cada orador humano. Por el contrario, el uso de una realización de la técnica de realce de sonido directivo agudo de la presente invención es capaz de realzar claramente sonidos desde una gran distancia y, por lo tanto, permite la construcción de un sistema de videoconferencia que es utilizable en una sala de conferencias grande sin tener que colocar un micrófono enfrente de cada orador humano.

<<Principio de la técnica de realce de punto de sonido>>

Se describirá a continuación un principio de una técnica de realce de punto de sonido de la presente invención. La técnica de realce de punto de sonido de la presente invención se basa en la naturaleza de una técnica de agrupación de micrófonos que es capaz de seguir sonidos desde cualquier dirección sobre la base de procesamiento de señal y usa positivamente los sonidos reflejados para captar sonidos con una relación SN alta. Una característica de la presente invención es un uso combinado de los sonidos reflejados y una técnica de procesamiento de señal que permite una directividad aguda. En particular, una de las características notables de la técnica de realce de punto de sonido de la presente invención es el uso de un objeto reflector para aumentar la

diferencia entre las funciones de transferencia de diferentes fuentes de sonido para una agrupación de micrófonos, a la luz del hecho de que las funciones de transferencia de fuentes de sonido situadas casi en las mismas direcciones de la agrupación de micrófonos, pero a diferentes distancias de la agrupación de micrófono a la agrupación de micrófonos son muy similares entre sí. Extrayendo las diferencias en la función de transferencia a través de procesamiento de señal, se puede lograr una técnica de realce de punto de sonido capaz de realzar sonidos según las distancias desde la agrupación de micrófonos.

Antes de la descripción, se definirán de nuevo los símbolos. El índice de una frecuencia discreta se denota por ω (El índice ω de una frecuencia discreta se puede considerar que es una frecuencia angular ω debido a que una frecuencia f y una frecuencia angular ω satisfacen la relación $\omega = 2\pi f$. Con respecto a ω , el "índice de una frecuencia discreta" también se puede conocer algunas veces simplemente como "frecuencia") y el índice del número de tiempo de trama se denota por k . La representación en el dominio de frecuencia de una trama de orden k de una señal analógica recibida en M micrófonos se denota por $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y un filtro que realiza una señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ de un sonido desde una fuente de sonido se supone que está situado en una dirección θ_s como se ve desde el centro de la agrupación de micrófonos a una distancia D_h desde el centro de la agrupación de micrófonos con una frecuencia ω se denota por $W^{\rightarrow}(\omega, \theta_s, D_h)$, donde M es un número entero mayor o igual a 2 y T representa la transposición. Se supone aquí que la distancia D_h es fija.

Mientras que el "centro de una agrupación de micrófonos" se puede determinar arbitrariamente, típicamente el centro geométrico de la agrupación de los M micrófonos se trata como el "centro de una agrupación de micrófonos". En el caso de una agrupación de micrófonos lineales, por ejemplo, el punto equidistante de los micrófonos en ambos extremos de la agrupación se trata como el "centro de la agrupación de micrófonos". En el caso de una agrupación plana de micrófonos en la que se disponen micrófonos en una matriz cuadrada de $m \times m$ ($m^2 = M$), por ejemplo, la posición en la que se cruzan las diagonales que unen los micrófonos en las esquinas se trata como el "centro de la agrupación de micrófonos".

La expresión "la fuente de sonido se supone que está situada en..." se ha usado debido a que la presencia real de una fuente de sonido en la ubicación no es esencial para la técnica de realce de punto de sonido de la presente invención. Es decir, como será evidente a partir de la descripción más tarde, la técnica de realce de punto de sonido de la presente invención realiza en esencia procesamiento de señal de aplicación de filtros a señales representadas por frecuencias y permite realizaciones en las que se crea un filtro de antemano para cada distancia discreta D_h . Por consiguiente, la presencia real de una fuente de sonido en la ubicación no se requiere incluso en la etapa donde se realiza realmente el procesamiento de realce de punto de sonido. Por ejemplo, si una fuente de sonido existe realmente en una ubicación en una dirección θ_s como se ve desde la agrupación de micrófonos y a una distancia de D_h de la agrupación de micrófonos en la etapa donde se realiza realmente el procesamiento de realce de punto de sonido, un sonido de la fuente de sonido se puede realzar eligiendo un filtro adecuado para la ubicación. Si la fuente de sonido no existe realmente en la ubicación y si se supone que no hay sonidos e incluso ningún ruido en absoluto, un sonido realzado por el filtro será idealmente un silencio total. Sin embargo, esto no es diferente de realzar un "sonido que llega de la ubicación".

Bajo estas condiciones, una señal en el dominio de frecuencia $Y(\omega, k, \theta_s, D_h)$ que resulta del realce de una señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ de un sonido de una fuente de sonido que se supone que está en una ubicación en una dirección θ_s a una distancia de D_h como se ve desde el centro de la agrupación de micrófonos (en lo sucesivo conocida como "ubicación (θ_s, D_h)" a menos que se indique de otro modo) con frecuencia ω se puede dar por la ecuación (106) (en lo sucesivo la señal resultante se conoce como señal de salida):

$$Y(\omega, k, \theta_s, D_h) = \vec{W}^H(\omega, \theta_s, D_h) \vec{X}(\omega, k) \quad (106)$$

donde H representa la transposición hermitiana.

El filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ se puede diseñar de varias formas. Se describirá aquí un diseño que usa el método de respuesta sin distorsión de varianza mínima (MVDR). En el método MVDR, un filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ está diseñado de modo que la potencia de sonidos de direcciones distintas de la dirección θ_s (en lo sucesivo, los sonidos de direcciones distintas de la dirección θ_s también se conocerán como "ruido") se minimiza a una frecuencia ω usando una matriz de correlación espacial $Q(\omega)$ bajo la condición de restricción de la ecuación (108) (véase la ecuación (107)). Se debería señalar que la matriz de correlación espacial $Q(\omega)$ se especifica como $Q(\omega, D_h)$ debido a que se supone aquí que la dirección D_h es fija. Suponiendo que una fuente de sonido se sitúa en una posición (θ_s, D_h), entonces $a^{\rightarrow}(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ representa las funciones de transferencia a una frecuencia ω entre la fuente de sonido y los M micrófonos. En otras palabras, $a^{\rightarrow}(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ representa las funciones de transferencia de un sonido desde la posición (θ_s, D_h) hasta los micrófonos incluidos en la agrupación de micrófonos a la frecuencia ω . La matriz de correlación espacial $Q(\omega)$ representa la correlación entre los componentes $X_i(\omega, k), \dots, X_M(\omega, k)$ de una señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ a la frecuencia ω y tiene $E[X_i(\omega, k)X_j^*(\omega, k)]$ ($1 \leq i \leq M, 1 \leq j \leq M$) como sus elementos (i, j). El operador $E[\cdot]$ representa una operación de promediado estadístico y el símbolo $*$ es un operador conjugado complejo. La matriz de correlación espacial $Q(\omega)$ se puede expresar usando los valores estadísticos de $X_1(\omega, k), \dots, X_M(\omega, k)$ obtenidos a partir de la observación o se

pueden expresar usando funciones de transferencia. Este último caso, donde la matriz de correlación espacial $Q(\omega)$ se expresa usando funciones de transferencia, se describirá momentáneamente en lo sucesivo.

$$\min_{\vec{W}(\omega, \theta_s, D_h)} \left(\vec{W}^H(\omega, \theta_s, D_h) Q(\omega, D_h) \vec{W}(\omega, \theta_s, D_h) \right) \quad (107)$$

$$\vec{W}^H(\omega, \theta_s, D_h) \vec{a}(\omega, \theta_s, D_h) = 1, 0 \quad (108)$$

5 Es conocido que el filtro $W^-(\omega, \theta_s, D_h)$ que es una solución óptima de la ecuación (107) se puede dar por la ecuación (109) (véase la Referencia 1 enumerada más tarde).

$$\vec{W}(\omega, \theta_s, D_h) = \frac{Q^{-1}(\omega, D_h) \vec{a}(\omega, \theta_s, D_h)}{\vec{a}^H(\omega, \theta_s, D_h) Q^{-1}(\omega, D_h) \vec{a}(\omega, \theta_s, D_h)} \quad (109)$$

10 Como se apreciará a partir del hecho de que la matriz inversa de la matriz de correlación espacial $Q(\omega, D_h)$ se incluye en la ecuación (109), la estructura de la matriz de correlación espacial $Q(\omega, D_h)$ es importante para lograr una directividad aguda. Se apreciará a partir de la ecuación (107) que la potencia de ruido depende de la estructura de la matriz de correlación espacial $Q(\omega, D_h)$.

15 Un conjunto de índices p de direcciones desde las cuales llega ruido se denota por $\{1, 2, \dots, P - 1\}$. Se supone que el índice s de la dirección del objetivo θ_s no pertenece al conjunto $\{1, 2, \dots, P - 1\}$. Suponiendo que $P - 1$ ruidos vienen de direcciones arbitrarias, la matriz de correlación espacial $Q(\omega, D_h)$ se puede dar por la ecuación (110a). Con el fin de diseñar un filtro que funcione suficientemente en presencia de muchos ruidos, es preferible que P sea un valor relativamente grande. Se supone aquí que P es un número entero del orden de M . Mientras que la descripción se da como si la dirección θ_s es una dirección constante (y por lo tanto direcciones distintas de la dirección θ_s se describen como direcciones desde las cuales llega ruido) por claridad de la explicación del principio de la técnica de realce de punto de sonido de la presente invención, la dirección θ_s en realidad pueden ser cualquier dirección que pueda ser un objetivo de realce de sonido. Usualmente, una pluralidad de direcciones pueden ser las direcciones θ_s . A la luz de esto, la diferenciación entre la dirección θ_s y las direcciones de ruido es subjetiva. Es más correcto considerar que una dirección seleccionada a partir de P direcciones diferentes que están predeterminadas como una pluralidad de direcciones posibles a partir de la cuales puede llegar cualquier sonido, incluyendo un sonido o ruido del objetivo, es la dirección que puede ser un objetivo de realce de sonido y las otras direcciones son direcciones de ruido. Por lo tanto, la matriz de correlación espacial $Q(\omega, D_h)$ se puede representar por las funciones de transferencia $a^-(\omega, \theta_\phi, D_h) = [a_1(\omega, \theta_\phi, D_h), \dots, a_M(\omega, \theta_\phi, D_h)]^T$ ($\phi \in \Phi$) de sonidos que vienen de direcciones θ_ϕ incluidas en una pluralidad de direcciones posibles que están a una distancia D_h del centro de la agrupación de micrófonos y desde las cuales pueden llegar sonidos a los micrófonos y se pueden escribir como la ecuación (110b), donde Φ es la unión del conjunto $\{1, 2, \dots, P - 1\}$ y un conjunto $\{s\}$. Obsérvese que $|\Phi| = P$ y $|\Phi|$ representa el número de elementos del conjunto Φ .

$$Q(\omega, D_h) = \vec{a}(\omega, \theta_s, D_h) \vec{a}^H(\omega, \theta_s, D_h) + \sum_{p \in \{1, \dots, P-1\}} \vec{a}(\omega, \theta_p, D_h) \vec{a}^H(\omega, \theta_p, D_h) \quad (110a)$$

$$Q(\omega, D_h) = \sum_{\phi \in \Phi} \vec{a}(\omega, \theta_\phi, D_h) \vec{a}^H(\omega, \theta_\phi, D_h) \quad (110b)$$

30 Aquí, se supone que la función de transferencia $a^-(\omega, \theta_s, D_h)$ de un sonido de la dirección θ_s y las funciones de transferencia $a^-(\omega, \theta_p, D_h) = [a_1(\omega, \theta_p, D_h), \dots, a_M(\omega, \theta_p, D_h)]^T$ de los sonidos de las direcciones $p \in \{1, 2, \dots, P - 1\}$ son ortogonales entre sí. Es decir, se supone que hay P sistemas base ortogonales que satisfacen la condición dada por la ecuación (111). El símbolo \perp representa ortogonalidad. Si $A^- \perp B^-$, el producto interno de los vectores A^- y B^- es cero. Se supone aquí que $P \leq M$. Obsérvese que si la condición dada por la ecuación (111) se puede relajar para suponer que hay P sistemas base que se pueden considerar aproximadamente como sistemas base ortogonales, P es preferiblemente un valor del orden de M o un valor relativamente grande mayor o igual a M .

$$\vec{a}(\omega, \theta_s, D_h) \perp \vec{a}(\omega, \theta_1, D_h) \perp \dots \perp \vec{a}(\omega, \theta_{P-1}, D_h) \quad (111)$$

40 Entonces, la matriz de correlación espacial $Q(\omega, D_h)$ se puede expandir como la ecuación (112). La ecuación (112) supone que la matriz de correlación espacial $Q(\omega, D_h)$ se puede descomponer en una matriz $V(\omega, D_h) = [a^-(\omega, \theta_s, D_h), a^-(\omega, \theta_1, D_h), \dots, a^-(\omega, \theta_{P-1}, D_h)]^T$ compuesta por P funciones de transferencia que satisfacen la ortogonalidad y una matriz unidad $\Lambda(\omega, D_h)$. Aquí, ρ es un valor propio de una función de transferencia $a^-(\omega, \theta_\phi, D_h)$ que satisface la ecuación (111) para la matriz de correlación espacial $Q(\omega, D_h)$ y es un valor real.

$$Q(\omega, D_h) = \rho \vec{V}(\omega, D_h) \bar{\Lambda}(\omega, D_h) \vec{V}^H(\omega, D_h) \quad (112)$$

Entonces, la matriz inversa de la matriz de correlación espacial $Q(\omega)$ se puede dar por la ecuación (113).

$$Q^{-1}(\omega, D_h) = \frac{1}{\rho} \vec{V}^H(\omega, D_h) \bar{\Lambda}^{-1}(\omega, D_h) \vec{V}(\omega, D_h) \quad (113)$$

5 La sustitución de la ecuación (113) en la ecuación (107) muestra que se minimiza la potencia de ruido. Si se minimiza la potencia de ruido, significa que se logra la directividad en la dirección θ_s . Por lo tanto, la ortogonalidad entre las funciones de transferencia de sonidos de diferentes direcciones es una condición importante para lograr la directividad en la dirección θ_s .

La razón de por qué es difícil para las técnicas convencionales lograr una directividad aguda en una dirección θ_s se tratará a continuación.

10 Las técnicas convencionales asumieron en el diseño de filtros que las funciones de transferencia se componen de los de sonidos directos. En realidad, hay sonidos reflejados que se producen por la reflexión de sonidos de la misma fuente de sonido fuera de superficies tales como las paredes y el techo y llegan a los micrófonos. Sin embargo, las técnicas convencionales consideraban los sonidos reflejados como un factor que degradaba la directividad e ignoraban la presencia de los sonidos reflejados. Suponiendo que los sonidos llegan a una agrupación de
15 micrófonos lineales como ondas planas, la técnica convencional trataba las funciones de transferencia $\vec{a}_{\text{conv}}(\omega, \theta) = [a_1(\omega, \theta), \dots, a_M(\omega, \theta)]^T$ como $\vec{a}_{\text{conv}}(\omega, \theta) = h_{\vec{d}}^{-1}(\omega, \theta)$ donde $h_{\vec{d}}^{-1}(\omega, \theta) = [h_{d1}(\omega, \theta), \dots, h_{dM}(\omega, \theta)]^T$ representa vectores de dirección de solamente un sonido directo que llega de una dirección θ (dado que las ondas de sonido se considera que son ondas planas, los vectores de dirección no dependen de la distancia D). Obsérvese que un vector de dirección es un vector complejo donde se disponen las características de respuesta de fase de los micrófonos a
20 una frecuencia ω con respecto a un punto de referencia para una onda de sonido de una dirección θ vista desde el centro de la agrupación de micrófonos.

Se supone en lo sucesivo momentáneamente que el sonido llega al micrófono lineal como ondas planas. Supongamos que un elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h_{\vec{d}}^{-1}(\omega, \theta)$ de un sonido directo se da, por ejemplo, por la ecuación (114c), donde u representa la distancia entre micrófonos adyacentes, j es una unidad
25 imaginaria. En este caso, el punto de referencia es el punto medio de la longitud total de la agrupación de micrófonos lineales (el centro de la agrupación de micrófonos lineales). La dirección θ se define como el ángulo formado por la dirección desde la cual llega un sonido directo y la dirección en la que están dispuestos los micrófonos incluidos en la agrupación de micrófonos lineales, como se ve desde el centro de la agrupación de micrófonos lineales (véase la Fig. 9). Obsérvese que un vector de dirección se puede expresar de varias formas. Por
30 ejemplo, suponiendo que el punto de referencia es la posición del micrófono en un extremo de la agrupación de micrófonos lineales, un elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h_{\vec{d}}^{-1}(\omega, \theta)$ de un sonido directo se puede dar por la ecuación (114d). En la siguiente descripción, la suposición es que el elemento de orden m $h_{dm}(\omega, \theta)$ del vector de dirección $h_{\vec{d}}^{-1}(\omega, \theta)$ de un sonido directo se puede escribir como la ecuación (114c).

$$h_{dm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos\theta\right] \quad (114c)$$

$$h_{dm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos\theta\right] \quad (114d)$$

35 El producto interno $\gamma_{\text{conv}}(\omega, \theta)$ de una función de transferencia de una dirección θ y una función de transferencia de una dirección del objetivo θ_s se puede dar por la ecuación (115), donde $\theta \neq \theta_s$.

$$\begin{aligned} \gamma_{\text{conv}}(\omega, \theta) &= \vec{a}_{\text{conv}}^H(\omega, \theta_s) \vec{a}_{\text{conv}}(\omega, \theta) \\ &= \vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\ &= \sum_{m=1}^M \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)(\cos\theta - \cos\theta_s)\right] \end{aligned} \quad (115)$$

En lo sucesivo, $\gamma_{\text{conv}}(\omega, \theta)$ se conoce como coherencia. La dirección θ en la cual la coherencia $\gamma_{\text{conv}}(\omega, \theta)$ es 0 se puede dar por la ecuación (116), donde q es un número entero arbitrario, excepto 0. Dado que $0 < \theta < \pi/2$, el intervalo de q está limitado para cada banda de frecuencia.

$$\theta = \arccos\left(\frac{2q\pi c}{M\omega u} + \cos\theta_s\right) \quad (116)$$

5 Dado que solamente los parámetros relacionados con el tamaño de la agrupación de micrófonos (M y u) se pueden cambiar en la ecuación (116), es difícil reducir la coherencia $\gamma_{\text{conv}}(\omega, \theta)$ sin cambiar ninguno de los parámetros relacionados con el tamaño de la agrupación de micrófonos si la diferencia (diferencia angular) $|\theta - \theta_s|$ entre direcciones es pequeña. Si este es el caso, la potencia de ruido no se reduce a un valor suficientemente pequeño y resultará una directividad que tiene una anchura de haz amplia en la dirección del objetivo θ_s como se ilustra esquemáticamente en la Fig. 5A.

10 La técnica de realce de punto de sonido de la presente invención se basa en la consideración descrita anteriormente y se caracteriza por tener en cuenta positivamente los sonidos reflejados, a diferencia de la técnica convencional, sobre la base de un entendimiento de que con el fin de diseñar un filtro que proporcione una directividad aguda en la dirección θ_s , es importante permitir que la coherencia sea reducida a un valor suficientemente pequeño incluso cuando la diferencia (diferencia angular) $|\theta - \theta_s|$ entre direcciones es pequeña.

15 Dos tipos de ondas planas, esto es, sonidos directos desde una fuente de sonido y sonidos reflejados producidos por reflexión de ese sonido fuera de un objeto 300 reflector, juntos entran en los micrófonos de una agrupación de micrófonos. Permitamos que el número de sonidos reflejados sea denotado por Ξ . Aquí, Ξ es un número entero predeterminado mayor o igual a 1. Entonces, una función de transferencia $\vec{a}^-(\omega, \theta) = [a_1(\omega, \theta), \dots, a_M(\omega, \theta)]^T$ se puede expresar mediante la suma de una función de transferencia de un sonido directo que viene de una dirección que puede ser un objetivo de realce de sonido y llega directamente a la agrupación de micrófonos y la función o funciones de transferencia de uno o más sonidos reflejados que se producen por la reflexión de ese sonido fuera de un objeto reflector y llegan a la agrupación de micrófonos. Específicamente, la función de transferencia se puede representar como la suma del vector de dirección del sonido directo y el vector de dirección de Ξ sonidos reflejados cuyas descomposiciones debidas a la reflexión y las diferencias de tiempo de llegada del sonido directo se corrigen, como se muestra en la ecuación (117a), donde $\tau_\xi(\theta)$ es la diferencia de tiempo de llegada entre el sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y α_ξ ($1 \leq \xi \leq \Xi$) es un coeficiente para tener en cuenta las descomposiciones de los sonidos debidas a la reflexión. Aquí, $\vec{h}_{r\xi}^-(\omega, \theta) = [h_{r1\xi}^-(\omega, \theta), \dots, h_{rM\xi}^-(\omega, \theta)]^T$ representa los vectores de dirección de los sonidos reflejados correspondientes al sonido directo desde la dirección θ . Típicamente, α_ξ ($1 \leq \xi \leq \Xi$) es menor o igual a 1 ($1 \leq \xi \leq \Xi$). Para cada sonido reflejado, si el número de reflexiones en el camino desde la fuente de sonido a los micrófonos es 1, α_ξ ($1 \leq \xi \leq \Xi$) se puede considerar que representa la reflectancia acústica del objeto desde el cual se reflejó el sonido reflejado de orden ξ .

$$\vec{a}(\omega, \theta) = \vec{h}_d(\omega, \theta) + \sum_{\xi=1}^{\Xi} \alpha_\xi \exp[-j\omega\tau_\xi(\theta)] \cdot \vec{h}_{r\xi}^-(\omega, \theta) \quad (117a)$$

35 Dado que se proporcionan uno o más sonidos reflejados a la agrupación de micrófonos compuesta por M micrófonos, son necesarios uno o más objetos reflectores. Desde este punto de vista, una fuente de sonido, la agrupación de micrófonos, y uno o más objetos reflectores están preferiblemente en tal relación de posición que un sonido de la fuente de sonido se refleja fuera al menos un objeto reflector antes de llegar a la agrupación de micrófonos, suponiendo que la fuente de sonido está situada en la dirección del objetivo para la mejora de sonido. Cada uno de los objetos reflectores tiene una forma bidimensional (por ejemplo, una placa plana) o una forma tridimensional (por ejemplo, una forma parabólica). Cada objeto reflector es preferiblemente de aproximadamente el tamaño de la agrupación de micrófonos o mayor (mayor en un factor de 1 a 2). Con el fin de usar eficazmente los sonidos reflejados, la reflectancia α_ξ ($1 \leq \xi \leq \Xi$) de cada objeto reflector es preferiblemente al menos mayor que 0, y más preferiblemente, la amplitud de un sonido reflejado que llega a la agrupación de micrófonos es mayor que la amplitud del sonido directo en un factor de 0,2 o mayor. Por ejemplo, cada objeto reflector es un sólido rígido. Cada objeto reflector puede ser un objeto móvil (por ejemplo, un reflector) o un objeto inamovible (tal como un suelo, una pared o un techo). Obsérvese que si un objeto inamovible se establece como objeto reflector, el vector de dirección para el objeto reflector necesita ser cambiado a medida que se vuelve a situar la agrupación de micrófonos (véanse las funciones $\psi(\theta)$ y $\psi_\xi(\theta)$ descritas más tarde) y, en consecuencia, el filtro necesita ser recalculado (restablecido).

Por lo tanto, los objetos reflectores son preferiblemente accesorios de la agrupación de micrófonos en aras de la robustez frente a cambios de entorno (en este caso, Ξ sonidos reflejados supuestos se considera que son sonidos reflejados fuera de los objetos reflectores). Aquí los “accesorios de la agrupación de micrófonos” son “objetos tangibles capaces de seguir los cambios de la posición y la orientación de la agrupación de micrófonos mientras que se mantiene la relación de posición (relación geométrica) con la agrupación de micrófonos). Un ejemplo simple puede ser una configuración donde los objetos reflectores se fijan a la agrupación de micrófonos.

Con el fin de describir concretamente ventajas de la técnica de realce de punto de sonido de la presente invención, se supone en lo siguiente que $\Xi = 1$, los sonidos se reflejan una vez y un objeto reflector existe a una distancia de L metros del centro de la agrupación de micrófonos. El objeto reflector es un objeto grueso y rígido. Dado que $\Xi = 1$ en este caso, el símbolo que representa esto se omite y, por lo tanto, la ecuación (117a) se puede reescribir como la ecuación (117b):

$$\vec{a}(\omega, \theta) = \vec{h}_d(\omega, \theta) + \alpha \exp[-j\omega\tau(\theta)] \cdot \vec{h}_r(\omega, \theta) \quad (117b)$$

Se representa un elemento de orden m del vector de dirección $\vec{h}_r(\omega, \theta) = [h_{r1}(\omega, \theta), \dots, h_{rM}(\omega, \theta)]^T$ de un sonido reflejado se puede dar por la ecuación (118a) de la misma forma que el vector de dirección de un sonido directo (véase la ecuación (114c)). La función $\psi(\theta)$ emite la dirección desde la cual llega un sonido reflejado. Obsérvese que si el vector de dirección de un sonido directo se escribe como la ecuación (114d), un elemento de orden m del vector de dirección $\vec{h}_r(\omega, \theta) = [h_{r1}(\omega, \theta), \dots, h_{rM}(\omega, \theta)]^T$ de un sonido reflejado se da por la ecuación (118b). Si $\Xi \leq 2$, un elemento de orden m de un vector de dirección de orden ξ ($1 \leq \xi \leq \Xi$) $\vec{h}_{r\xi}(\omega, \theta) = [h_{r1\xi}(\omega, \theta), \dots, h_{rM\xi}(\omega, \theta)]^T$ se da por la ecuación (118c) o la ecuación (118d). La función $\psi_\xi(\theta)$ emite la dirección desde la cual llega el sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$).

$$h_{rm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos(\Psi(\theta))\right] \quad (118a)$$

$$h_{rm}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos(\Psi(\theta))\right] \quad (118b)$$

$$h_{rm\xi}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}\left(m - \frac{M+1}{2}\right)\cos(\Psi_\xi(\theta))\right] \quad (118c)$$

$$h_{rm\xi}(\omega, \theta) = \exp\left[-\frac{j\omega u}{c}(m-1)\cos(\Psi_\xi(\theta))\right] \quad (118d)$$

Dado que la ubicación de un objeto reflector se puede establecer según sea adecuado, la dirección desde la cual llega un sonido reflejado se puede tratar como un parámetro variable.

Suponiendo que un objeto reflector de placa plana está cerca de la agrupación de micrófonos (la distancia L no es extremadamente grande en comparación con el tamaño de la agrupación de micrófonos), la coherencia $\gamma(\omega, \theta)$ se da por la ecuación (119), donde $\theta \neq \theta_s$.

$$\begin{aligned} \gamma(\omega, \theta) &= \vec{a}^H(\omega, \theta_s) \vec{a}(\omega, \theta) \\ &= \vec{h}_d^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\ &\quad + \alpha \exp[-j\omega\tau(\theta)] \cdot \vec{h}_d^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) \\ &\quad + \alpha \exp[j\omega\tau(\theta_s)] \cdot \vec{h}_r^H(\omega, \theta_s) \vec{h}_d(\omega, \theta) \\ &\quad + \alpha^2 \exp[-j\omega(\tau(\theta) - \tau(\theta_s))] \cdot \vec{h}_r^H(\omega, \theta_s) \vec{h}_r(\omega, \theta) \end{aligned} \quad (119)$$

Será evidente a partir de la ecuación (119) que la coherencia $\gamma(\omega, \theta)$ de la ecuación (119) puede ser menor que la coherencia $\gamma_{\text{conv}}(\omega, \theta)$ de la técnica convencional de la ecuación (115). Dado que los parámetros ($\psi(\theta)$ y L) que se pueden cambiar volviendo a situar o reorientando el objeto reflector se incluyen en el segundo a cuarto términos de la ecuación (119), hay una posibilidad de que se pueda eliminar el primer término, $h_d^{-H}(\omega, \theta) h_d^{-}(\omega, \theta)$.

- 5 Por ejemplo, si se coloca un reflector plano en tal posición que la dirección a lo largo de la cual se disponen los micrófonos en una agrupación de micrófonos lineales es normal al reflector, $\psi(\theta) = \pi - \theta$ se mantiene para la función $\psi(\theta)$ y la ecuación (120) recoge la diferencia $\tau(\theta)$ en el tiempo de llegada entre un sonido directo y un sonido reflejado. Por lo tanto, las condiciones de las ecuaciones (121) y (122) se generan para los elementos de la ecuación (119). Aquí el símbolo * es un operador conjugado complejo.

$$\tau(\theta) = \begin{cases} (2L \cos \theta)/c & \left(0 < \theta \leq \frac{\pi}{4} \right) \\ (2L \sin \theta \tan \theta)/c & \left(\frac{\pi}{4} < \theta < \frac{\pi}{2} \right) \end{cases} \quad (120)$$

$$\bar{h}_d^H(\omega, \theta_s) \bar{h}_d(\omega, \theta) = \bar{h}_r^H(\omega, \theta_s) \bar{h}_r(\omega, \theta) \quad (121)$$

$$10 \quad \bar{h}_d^H(\omega, \theta_s) \bar{h}_r(\omega, \theta) = \left[\bar{h}_r^H(\omega, \theta_s) \bar{h}_d(\omega, \theta) \right]^* \quad (122)$$

Dado que el valor absoluto de $h_d^{-H}(\omega, \theta) h_r^{-}(\omega, \theta)$ es suficientemente menor que $h_d^{-H}(\omega, \theta) h_d^{-}(\omega, \theta)$, el segundo y tercer términos de la ecuación (119) son despreciables. Entonces la coherencia $\gamma(\omega, \theta)$ se puede aproximar como la ecuación (123):

$$\tilde{\gamma}(\omega, \theta) \approx \left\{ 1 + \alpha^2 \exp[-j\omega(\tau(\theta) - \tau(\theta_s))] \right\} \bar{h}_d^H(\omega, \theta_s) \bar{h}_d(\omega, \theta) \quad (123)$$

- 15 Incluso si $h_d^{-H}(\omega, \theta) h_d^{-}(\omega, \theta) \neq 0$, una coherencia aproximada $\tilde{\gamma}(\omega, \theta)$ tiene una solución mínima θ de la ecuación (124), donde q es un número entero positivo arbitrario. El intervalo de q está restringido para cada frecuencia.

$$\theta = \begin{cases} \arccos\left(\frac{(2q+1)\pi c}{2\omega L} + \cos \theta_s\right) & \left(0 < \theta \leq \frac{\pi}{4} \right) \\ \frac{(2q+1)\pi c}{4\omega L} + \frac{1}{2} \sqrt{\left(\frac{(2q+1)\pi c}{4\omega L}\right)^2 + 4} & \left(\frac{\pi}{4} < \theta < \frac{\pi}{2} \right) \end{cases} \quad (124)$$

- 20 Es decir, no solamente se puede suprimir la coherencia en una dirección dada por la ecuación (116) sino también la coherencia en una dirección dada por la ecuación (124). Dado que la supresión de coherencia puede reducir la potencia de ruido, una directividad aguda se puede lograr como se muestra esquemáticamente en la Fig. 5B.

- Mientras que las Fig. 5A y 5B muestran esquemáticamente la diferencia entre la directividad lograda por la técnica de realce de sonido directivo agudo de la presente invención y la directividad lograda mediante una técnica convencional, la Fig. 6 muestra específicamente la diferencia entre θ dada por la ecuación (116) y θ dada por la ecuación (124). Aquí, $\omega = 2\pi \times 1000$ [rad/s], $L = 0,70$ [m], y $\theta_s = \pi/4$ [rad]. La dependencia de la dirección de la coherencia normalizada se muestra en la Fig. 6 para comparación entre las técnicas. La dirección indicada por un círculo es θ dada por la ecuación (116) y las direcciones indicadas por el símbolo + son θ dadas por la ecuación (124). Como se puede ver a partir de la Fig. 6, según la técnica convencional, θ que produce una coherencia de 0 para $\theta_s = \pi/4$ [rad] existe solamente en la dirección indicada por el círculo, mientras que según el principio de la técnica de realce de sonido directivo agudo de la presente invención, θ que produce una coherencia de 0 para $\theta_s = \pi/4$ [rad] existe en muchas direcciones indicadas por el símbolo +. Especialmente, las direcciones indicadas por el símbolo + existen mucho más cerca de $\theta_s = \pi/4$ [rad] que la dirección indicada por el círculo. Por lo tanto, se entenderá que la técnica de la presente invención logra una directividad más aguda que la técnica convencional.

Mientras que por claridad de la explicación del principio de la técnica de realce de punto de sonido de la presente invención, se ha supuesto en lo precedente que ondas de sonidos llegan como ondas planas, la esencia de la

técnica de realce de sonido de punto de la presente invención es que la función de transferencia $a^{\rightarrow}(\omega, \theta, D) = [a_1(\omega, \theta, D), \dots, a_M(\omega, \theta, D)]^T$ se representa mediante la suma del vector de dirección de un sonido directo y los vectores de dirección de Ξ sonidos reflejados, como se muestra en la Ecuación (117a), por ejemplo, como es evidente a partir de la descripción precedente. Por consiguiente, se entenderá que la técnica no está limitada a las ondas de sonido que llegan como ondas planas, sino que es capaz de lograr el realce del sonido de los sonidos que llegan como ondas esféricas con una directividad más alta que la técnica convencional.

Se describirán las funciones de transferencia $a^{\rightarrow}(\omega, \theta, D)$ de ondas de sonido que llegan como ondas esféricas. Dos tipos de ondas esféricas, esto es, sonidos directos de una fuente de sonido y sonidos reflejados producidos por reflexión de ese sonido fuera de un objeto 300 reflector, juntos entran en los micrófonos de una agrupación de micrófonos. Permitamos que el número de sonidos reflejados sea denotado por Ξ . Aquí, Ξ es un número entero predeterminado mayor o igual a 1. Entonces, una función de transferencia $a^{\rightarrow}(\omega, \theta, D) = [a_1(\omega, \theta, D), \dots, a_M(\omega, \theta, D)]^T$ se puede expresar mediante la suma de una función de transferencia de un sonido directo que viene de una posición (θ_s, D) que puede ser un objetivo de realce de sonido y llega directamente a la agrupación de micrófonos y la función o funciones de transferencia de uno o más sonidos reflejados que se producen por la reflexión de ese sonido fuera de un objeto reflector y llegan a la agrupación de micrófonos. Específicamente, la función de transferencia se puede representar como la suma del vector de dirección del sonido directo y del vector de dirección de Ξ sonidos reflejados cuyas descomposiciones debidas a la reflexión y las diferencias de tiempo de llegada del sonido directo se corrigen, como se muestra en la ecuación (125), donde $\tau_{\xi}(\theta, D)$ es la diferencia de tiempo de llegada entre el sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y α_{ξ} ($1 \leq \xi \leq \Xi$) es un coeficiente para tener en cuenta las descomposiciones de los sonidos debidas a la reflexión. Aquí $h^{\rightarrow}_d(\omega, \theta, D_h) = [h_{d1}(\omega, \theta, D_h), \dots, h_{dM}(\omega, \theta, D_h)]^T$ representa el vector de dirección de un sonido directo de la posición (θ_s, D) y $h^{\rightarrow}_{r\xi}(\omega, \theta, D) = [h_{r1\xi}(\omega, \theta, D), \dots, h_{rM\xi}(\omega, \theta, D)]^T$ representa el vector de dirección de un sonido reflejado correspondiente al sonido directo de la posición (θ_s, D) . Se añadirá aquí una nota acerca del término “vector de dirección”. Un “vector de dirección” también se denomina “vector de dirección” y, como su nombre sugiere, representa típicamente un vector complejo que es dependiente de la “dirección”. Desde este punto de vista, es más preciso hacer referencia a un vector complejo que es dependiente de una posición (θ_s, D) como “vector de dirección extendida”, por ejemplo. Sin embargo, en aras de la simplicidad, el vector complejo que es dependiente de una posición (θ_s, D) también se conocerá simplemente como el “vector de dirección” en la presente memoria. Típicamente, α_{ξ} ($1 \leq \xi \leq \Xi$) es menor o igual a 1 ($1 \leq \xi \leq \Xi$). Para cada sonido reflejado, si el número de reflexiones en el camino desde la fuente de sonido a los micrófonos es 1, α_{ξ} ($1 \leq \xi \leq \Xi$) se puede considerar que representa la reflectancia acústica del objeto desde el cual se reflejó el sonido reflejado de orden ξ .

$$\bar{a}(\omega, \theta, D) = \bar{h}_d(\omega, \theta, D) + \sum_{\xi=1}^{\Xi} \alpha_{\xi} \exp[-j\omega\tau_{\xi}(\theta, D)] \cdot \bar{h}_{r\xi}(\omega, \theta, D) \quad (125)$$

En la ecuación (125), un elemento de orden m $h_{dm}(\omega, \theta, D_h)$ del vector de dirección $h^{\rightarrow}_d(\omega, \theta, D_h)$ del sonido directo se puede dar por la ecuación (125a), por ejemplo. Aquí m es un número entero que satisface $1 \leq m \leq M$, c representa la velocidad del sonido y j es una unidad imaginaria. En un sistema de coordenadas espaciales adecuadamente establecido, $\vec{v}_{\theta, D}^{(d)}$ representa un vector de posición de una posición (θ, D) , \vec{u}_m representa un vector de posición de un micrófono de orden m , el símbolo $\|\cdot\|$ representa una norma, y $f(\|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|)$ es una función que representa una distancia de descomposición de una onda de sonido. Por ejemplo, $f(\|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|) = 1/\|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|$ y en este caso la ecuación (125a) se puede escribir como la ecuación (125b).

$$h_{dm}(\omega, \theta, D) = f(\|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|) \cdot \exp\left[-\frac{j\omega}{c} \|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|\right] \quad (125a)$$

$$h_{dm}(\omega, \theta, D) = \frac{1}{\|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|} \exp\left[-\frac{j\omega}{c} \|\vec{v}_{\theta, D}^{(d)} - \vec{u}_m\|\right] \quad (125b)$$

En la ecuación (125), un elemento de orden m $h_{rm\xi}(\omega, \theta, D)$ del vector de dirección $h^{\rightarrow}_{r\xi}(\omega, \theta, D) = [h_{r1\xi}(\omega, \theta, D), \dots, h_{rM\xi}(\omega, \theta, D)]^T$ se puede dar por la ecuación (126a), como el vector de dirección del sonido directo (véase la

ecuación (125a)). Aquí, m es un número entero que satisface $1 \leq m \leq M$, c representa la velocidad del sonido y j es una unidad imaginaria. En el sistema de coordenadas espaciales, $\vec{v}_{\theta,D}^{(\xi)}$ representa un vector de posición de una posición que es una imagen espejo de una posición (θ, D) con respecto a la superficie reflectora de un reflector de orden ξ , \vec{u}_m representa el vector de posición del micrófono de orden m , el símbolo $\|\cdot\|$ representa una

5 norma, y $f(\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|)$ es una función que representa una distancia de descomposición de una onda de sonido. Por ejemplo, $f(\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|)$ es una función que representa una descomposición de distancia de una onda de sonido. Por ejemplo, $f(\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|) = 1/\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|$ y en este caso la ecuación (126a) se puede escribir como la ecuación (126b).

$$h_{rm\xi}(\omega, \theta, D) = f\left(\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|\right) \cdot \exp\left[-\frac{j\omega}{c}\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|\right] \quad (126a)$$

$$h_{rm\xi}(\omega, \theta, D) = \frac{1}{\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|} \exp\left[-\frac{j\omega}{c}\|\vec{v}_{\theta,D}^{(\xi)} - \vec{u}_m\|\right] \quad (126b)$$

10 Obsérvese que una diferencia de tiempo de llegada de orden ξ $\tau_\xi(\theta, D)$ y vector de posición $\vec{v}_{\theta,D}^{(\xi)}$ se pueden calcular teóricamente sobre la base de la relación de posición entre la posición (θ, D) , la agrupación de micrófonos y el objeto reflector de orden ξ cuando se determina la relación de posición.

A diferencia de las técnicas convencionales, la técnica de realce de punto de sonido de la presente invención tiene en cuenta positivamente los sonidos reflejados y, por lo tanto, es capaz de un realce de punto de sonido directivo
 15 agudo. Esto se describirá tomando dos fuentes de sonido a modo de ejemplo. Es difícil el realce de puntos de los sonidos que emanan de dos fuentes de sonido A y B a diferentes distancias de una agrupación de micrófonos pero en aproximadamente las mismas direcciones vistas desde la agrupación de micrófonos como se ilustra en la Fig. 18A solamente de sonidos directos de las dos fuentes de sonido por la siguiente razón. Dado el hecho de que $\theta_{[A]} \approx \theta_{[B]}$ y $D_{[A]} \neq D_{[B]}$, hay una diferencia entre un valor de función de descomposición $f(\|\vec{v}_{\theta_{[A]}, D_{[A]}}^{(d)} - \vec{u}_m\|)$ que aparece en el vector de dirección $\vec{h}_d(\omega, \theta_{[A]}, D_{[A]})$ de un sonido directo correspondiente a la posición $(\theta_{[A]}, D_{[A]})$ de la fuente de sonido A y un valor de función de descomposición $f(\|\vec{v}_{\theta_{[B]}, D_{[B]}}^{(d)} - \vec{u}_m\|)$ que aparece en el vector de dirección $\vec{h}_d(\omega, \theta_{[B]}, D_{[B]})$ de un sonido directo correspondiente a la posición $(\theta_{[B]}, D_{[B]})$ de la fuente de sonido B como función de la distancia desde la agrupación de micrófonos. Sin embargo, en realidad, la distinción entre la intensidad de una señal de fuente (volumen de sonido) y su valor de función de descomposición no se puede hacer a partir de la
 20 intensidad de un sonido (volumen de sonido) captado con la agrupación de micrófonos. Es decir, si $a_{conv}(\omega, \theta, D) = \vec{h}_d(\omega, \theta, D)$ como en la técnica convencional, las funciones de transferencia de sonidos directos no son suficientes como indicación para diferenciar entre las distancias de fuentes de sonido en aproximadamente las mismas direcciones y, por lo tanto, es difícil diseñar filtros capaces de realce de punto, como es evidente a partir de las ecuaciones (109), (110a) y (110b).

30 Por el contrario, la técnica de realce de punto de sonido de la presente invención tiene en cuenta positivamente los sonidos reflejados, por lo tanto existen fuentes de sonido virtuales $A(\xi)$ y $B(\xi)$ de sonidos reflejados de orden ξ en posiciones de imágenes espejo de las fuentes de sonido A y B con respecto a la superficie reflectora del reflector 300 de orden ξ desde el punto de vista de la agrupación de micrófonos como se ilustra en la Fig. 18B. Esto es equivalente a que los sonidos que emanan de las fuentes de sonido A y B y que se reflejan en el reflector 300 de orden ξ vienen desde las fuentes de sonido virtuales $A(\xi)$ y $B(\xi)$. Hay una diferencia significativa entre el sonido reflejado de orden ξ desde la fuente de sonido virtual $A(\xi)$ y el sonido reflejado de orden ξ desde la fuente de
 35 sonido virtual $B(\xi)$ en el vector de posición $\vec{V}_{\theta_{[A(\xi)], D_{[A(\xi)]}}^{(\xi)}}$ y $\vec{V}_{\theta_{[B(\xi)], D_{[B(\xi)]}}^{(\xi)}}$ y en la diferencia de tiempo de llegada $\tau_\xi(\theta_{[A]}, D_{[A]})$ y $\tau_\xi(\theta_{[B]}, D_{[B]})$. Las funciones de transferencia $\vec{a}(\omega_{[A]}, \theta_{[A]}, D_{[A]})$ y $\vec{a}(\omega_{[B]}, \theta_{[B]}, D_{[B]})$ que corresponden a las posiciones $(\theta_{[A]}, D_{[A]})$ y $(\theta_{[B]}, D_{[B]})$, respectivamente, se pueden dar por las ecuaciones (127a) y (127b), respectivamente. La presencia del segundo término de las ecuaciones (127a) y (127b) proporciona una diferencia significativa entre funciones de transferencia correspondientes a diferentes posiciones a pesar de que $\theta_{[A]} \approx \theta_{[B]}$. Extrayendo la diferencia entre las funciones de transferencia mediante el método de conformación de haces, se puede realizar el realce de punto de los sonidos según las posiciones de las fuentes de sonido asumidas.

40

$$\begin{aligned}\bar{a}(\omega, \theta_{[A]}, D_{[A]}) &= \bar{h}_d(\omega, \theta_{[A]}, D_{[A]}) \\ &+ \sum_{\xi=1}^{\Xi} \alpha_{\xi} \exp[-j\omega\tau_{\xi}(\theta_{[A]}, D_{[A]})] \cdot \bar{h}_{r\xi}(\omega, \theta_{[A]}, D_{[A]})\end{aligned}\quad (127a)$$

$$\begin{aligned}\bar{a}(\omega, \theta_{[B]}, D_{[B]}) &= \bar{h}_d(\omega, \theta_{[B]}, D_{[B]}) \\ &+ \sum_{\xi=1}^{\Xi} \alpha_{\xi} \exp[-j\omega\tau_{\xi}(\theta_{[B]}, D_{[B]})] \cdot \bar{h}_{r\xi}(\omega, \theta_{[B]}, D_{[B]})\end{aligned}\quad (127b)$$

$$\bar{h}_d(\omega, \theta_{[A]}, D_{[A]}) = [h_{d1}(\omega, \theta_{[A]}, D_{[A]}), \dots, h_{dM}(\omega, \theta_{[A]}, D_{[A]})]^T$$

$$\bar{h}_d(\omega, \theta_{[B]}, D_{[B]}) = [h_{d1}(\omega, \theta_{[B]}, D_{[B]}), \dots, h_{dM}(\omega, \theta_{[B]}, D_{[B]})]^T$$

$$\bar{h}_{r\xi}(\omega, \theta_{[A]}, D_{[A]}) = [h_{r1\xi}(\omega, \theta_{[A]}, D_{[A]}), \dots, h_{rM\xi}(\omega, \theta_{[A]}, D_{[A]})]^T$$

$$\bar{h}_{r\xi}(\omega, \theta_{[B]}, D_{[B]}) = [h_{r1\xi}(\omega, \theta_{[B]}, D_{[B]}), \dots, h_{rM\xi}(\omega, \theta_{[B]}, D_{[B]})]^T$$

$$h_{dm}(\omega, \theta_{[A]}, D_{[A]}) = f\left(\left\| \bar{v}_{\theta_{[A]}, D_{[A]}}^{(d)} - \bar{u}_m \right\|\right) \cdot \exp\left[-\frac{j\omega}{c} \left\| \bar{v}_{\theta_{[A]}, D_{[A]}}^{(d)} - \bar{u}_m \right\|\right]$$

$$h_{dm}(\omega, \theta_{[B]}, D_{[B]}) = f\left(\left\| \bar{v}_{\theta_{[B]}, D_{[B]}}^{(d)} - \bar{u}_m \right\|\right) \cdot \exp\left[-\frac{j\omega}{c} \left\| \bar{v}_{\theta_{[B]}, D_{[B]}}^{(d)} - \bar{u}_m \right\|\right]$$

$$h_{rm\xi}(\omega, \theta_{[A]}, D_{[A]}) = f\left(\left\| \bar{v}_{\theta_{[A(\xi)], D_{[A(\xi)]}}^{(\xi)} - \bar{u}_m \right\|\right) \cdot \exp\left[-\frac{j\omega}{c} \left\| \bar{v}_{\theta_{[A(\xi)], D_{[A(\xi)]}}^{(\xi)} - \bar{u}_m \right\|\right]$$

$$h_{rm\xi}(\omega, \theta_{[B]}, D_{[B]}) = f\left(\left\| \bar{v}_{\theta_{[B(\xi)], D_{[B(\xi)]}}^{(\xi)} - \bar{u}_m \right\|\right) \cdot \exp\left[-\frac{j\omega}{c} \left\| \bar{v}_{\theta_{[B(\xi)], D_{[B(\xi)]}}^{(\xi)} - \bar{u}_m \right\|\right]$$

Hasta ahora, la distancia D_h se ha fijado con el fin de explicar qué tan alta directividad se puede lograr. Por consiguiente, las matrices de correlación espacial $Q(\omega)$ se han escrito como (110a) y (110b). Sin embargo, teniendo en cuenta la correlación entre las funciones de transferencia de M canales para diferentes distancias D_{δ} ($\delta = 1, 2, \dots, G$), la cantidad de información que se refiere a un campo de sonido se puede aumentar para construir una matriz de correlación espacial que proporciona filtros más precisos. La matriz de correlación espacial $Q(\omega)$ se puede dar por la

ecuación (110c). Un conjunto al que pertenecen los índices Φ de direcciones θ_ϕ se denota por Φ ($|\Phi| = P$) y un conjunto al que pertenecen los índices δ de distancias D_δ se denota por Δ ($|\Delta| = G$).

$$Q(\omega) = \sum_{\phi \in \Phi} \sum_{\delta \in \Delta} \bar{a}(\omega, \theta_\phi, D_\delta) \bar{a}^H(\omega, \theta_\phi, D_\delta) \quad (110c)$$

Entonces, usando la matriz de correlación espacial $Q(\omega)$ dada por la ecuación (110c), un filtro $W^-(\omega, \theta_s, D_h)$ diseñado mediante el método de respuesta sin distorsión de varianza mínima (MVDR) se puede escribir como la ecuación (109a) en lugar de la ecuación (109).

$$\bar{W}(\omega, \theta_s, D_h) = \frac{Q^{-1}(\omega) \bar{a}(\omega, \theta_s, D_h)}{\bar{a}^H(\omega, \theta_s, D_h) Q^{-1}(\omega) \bar{a}(\omega, \theta_s, D_h)} \quad (109a)$$

Como se ha descrito, la esencia de la técnica de realce de punto de sonido de la presente invención es que la función de transferencia $\bar{a}^-(\omega, \theta, D) = [a_1(\omega, \theta, D), \dots, a_M(\omega, \theta, D)]^T$ se representa mediante la suma del vector de dirección de un sonido directo y los vectores de dirección de M sonidos reflejados. Dado que esto no afecta al concepto de diseño de filtro, los filtros $W^-(\omega, \theta_s, D_h)$ se pueden diseñar mediante un método distinto del método de respuesta sin distorsión de varianza mínima (MVDR).

Se describirán métodos distintos del método MVDR descrito anteriormente. Son: <1> un método de diseño de filtro basado en el criterio de maximización de SNR, <2> un método de diseño de filtro basado en inversión de potencia, <3> un método de diseño de filtro que usa MVDR con uno o más puntos de supresión (direcciones en las que se suprime la ganancia de ruido) como una condición de restricción, <4> un método de diseño de filtro que usa la conformación de haces de retardo y suma, <5> un método de diseño de filtro que usa el método de máxima verosimilitud y <6> un método de diseño de filtro que usa el método de agrupación adaptativa de micrófonos para reducción de ruido (AMNOR). Para <1> el método de diseño de filtro basado en el criterio de maximización de SNR y <2> el método de diseño de filtro basado en la inversión de potencia, se hace referencia a la Referencia 2 enumerada a continuación. Para <3> el método de diseño de filtro que usa MVDR con uno o más puntos de supresión (direcciones en las que se suprime la ganancia de ruido) como condición de restricción, se hace referencia a la Referencia 3 enumerada a continuación. Para <6> el método de diseño de filtro que usa el método de agrupación adaptativa de micrófonos para reducción de ruido (AMNOR), se hace referencia a la Referencia 4 enumerada a continuación.

<1> Método de diseño de filtro basado en el criterio de maximización de SNR

En el método de diseño de filtro basado en el criterio de maximización de SNR, se determina un filtro $W^-(\omega, \theta_s, D_h)$ sobre la base de un criterio de maximización de la relación SN (SNR) de una posición (θ_s, D_h) . La matriz de correlación espacial para un sonido desde la posición (θ_s, D_h) se denota por $R_{ss}(\omega)$ y una matriz de correlación espacial para un sonido de una posición distinta de la posición (θ_s, D_h) se denota por $R_{nn}(\omega)$. Entonces la SNR se puede dar por la ecuación (128). Aquí, $R_{ss}(\omega)$ se puede dar por la ecuación (129) y $R_{nn}(\omega)$ se puede dar por la ecuación (130). Las funciones de transferencia $\bar{a}^-(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ se pueden dar por la ecuación (125), por ejemplo (para ser precisos, la ecuación (125) donde θ se sustituye por θ_s y D se sustituye por D_h). Un conjunto al que pertenecen los índices Φ de direcciones θ_ϕ se denota por Φ ($|\Phi| = P$) y un conjunto al que pertenecen los índices δ de distancias D_δ se denota por Δ ($|\Delta| = G$).

$$SNR = \frac{\bar{W}^H(\omega, \theta_s, D_h) R_{ss}(\omega) \bar{W}(\omega, \theta_s, D_h)}{\bar{W}^H(\omega, \theta_s, D_h) R_{nn}(\omega) \bar{W}(\omega, \theta_s, D_h)} \quad (128)$$

$$R_{ss}(\omega) = \bar{a}(\omega, \theta_s, D_h) \bar{a}^H(\omega, \theta_s, D_h) \quad (129)$$

$$R_{nn}(\omega) = \left(\sum_{\phi \in \Phi} \sum_{\delta \in \Delta} \bar{a}(\omega, \theta_\phi, D_\delta) \bar{a}^H(\omega, \theta_\phi, D_\delta) \right) - R_{ss}(\omega) \quad (130)$$

El filtro $W^-(\omega, \theta_s, D_h)$ que maximiza la SNR de la ecuación (128) se puede obtener estableciendo el gradiente relacionado con el filtro $W^-(\omega, \theta_s, D_h)$ a cero, es decir, por la ecuación (131).

$$\nabla_{\vec{W}(\omega, \theta_s, D_h)} [\text{SNR}] = 0 \quad (131)$$

donde

$$\nabla_{\vec{W}(\omega, \theta_s, D_h)} [\text{SNR}] = \frac{2R_{ss}(\omega)\vec{W}(\omega, \theta_s, D_h)\left(\vec{W}^H(\omega, \theta_s, D_h)R_{nn}(\omega)\vec{W}(\omega, \theta_s, D_h)\right)}{\left(\vec{W}^H(\omega, \theta_s, D_h)R_{nn}(\omega)\vec{W}(\omega, \theta_s, D_h)\right)^2} - \frac{2R_{nn}(\omega)\vec{W}(\omega, \theta_s, D_h)\left(\vec{W}^H(\omega, \theta_s, D_h)R_{ss}(\omega)\vec{W}(\omega, \theta_s, D_h)\right)}{\left(\vec{W}^H(\omega, \theta_s, D_h)R_{nn}(\omega)\vec{W}(\omega, \theta_s, D_h)\right)^2}$$

De esta manera, el filtro $W^-(\omega, \theta_s, D_h)$ que maximiza la SNR de la ecuación (128) se puede dar por la ecuación (132):

$$\vec{W}(\omega, \theta_s, D_h) = R_{nn}^{-1}(\omega)\vec{a}(\omega, \theta_s, D_h) \quad (132)$$

La ecuación (132) incluye la matriz inversa de la matriz de correlación espacial $R_{nn}(\omega)$ de un sonido de una posición distinta de la posición (θ_s, D_h) . Es conocido que la matriz inversa de $R_{nn}(\omega)$ se puede sustituir con la matriz inversa de una matriz de correlación espacial $R_{xx}(\omega)$ de una entrada completa que incluye sonidos de (1) la posición (θ_s, D_h) y (2) sonidos de una posición distinta de la dirección (θ_s, D_h) . Aquí, $R_{xx}(\omega) = R_{ss}(\omega) + R_{nn}(\omega) = Q(\omega)$. Es decir, el filtro $W^-(\omega, \theta_s, D_h)$ que maximiza la SNR de la ecuación (128) se puede obtener por la ecuación (133):

$$\vec{W}(\omega, \theta_s, D_h) = R_{xx}^{-1}(\omega)\vec{a}(\omega, \theta_s, D_h) \quad (133)$$

<2> Método de diseño de filtro basado en inversión de potencia

En el método de diseño de filtro basado en inversión de potencia, un filtro $W^-(\omega, \theta_s, D_h)$ se determina sobre la base de un criterio de minimización de la potencia de salida promedio de un conformador de haces mientras que un coeficiente de filtro para un micrófono se fija a un valor constante. Aquí, se describirá un ejemplo donde se fija el coeficiente de filtro para el primer micrófono entre M micrófonos. En este método de diseño, se diseña un filtro $W^-(\omega, \theta_s, D_h)$ que minimiza la potencia de sonidos de todas las posiciones (todas las posiciones que se puede suponer que son posiciones de fuente de sonido) usando una matriz de correlación espacial $R_{xx}(\omega)$ (véase la ecuación (134)) bajo la condición de restricción de la ecuación (135). Las funciones de transferencia $\vec{a}^-(\omega, \theta_s, d_h) = [a_1^-(\omega, \theta_s, D_h), \dots, a_M^-(\omega, \theta_s, D_h)]^T$ se pueden dar por la ecuación (125), por ejemplo (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_s y D se sustituye por D_h).

$$\min_{\vec{W}(\omega, \theta_s, D_h)} \left(\vec{W}^H(\omega, \theta_s, D_h)R_{xx}(\omega)\vec{W}(\omega, \theta_s, D_h) \right) \quad (134)$$

$$\vec{W}^H(\omega, \theta_s, D_h)\vec{G} = \vec{G}^H R_{xx}^{-1}(\omega)\vec{G} \quad (135)$$

donde

$$\vec{G} = [1, 0, \dots, 0]^T$$

Es conocido que el filtro $W^-(\omega, \theta_s, D_h)$ que es una solución óptima de la ecuación (134) se puede dar por la ecuación (136) (véase la Referencia 2 enumerada a continuación).

$$\vec{W}(\omega, \theta_s, D_h) = R_{xx}^{-1}(\omega)\vec{G} \quad (136)$$

<3> Método de diseño de filtro que usa MVDR con uno o más puntos de supresión como condición de restricción

En el método MVDR descrito anteriormente, un filtro $W^-(\omega, \theta_s, D_h)$ se ha diseñado bajo la condición de restricción única de que se obtiene un filtro que minimiza la potencia de salida promedio de un conformador de haces dada por la ecuación (107) (es decir, la potencia de ruido que es sonidos de direcciones distintas de una posición (θ_s, D_h) bajo la condición de restricción de que el filtro pasa sonidos desde una posición (θ_s, D_h) en todas las bandas de frecuencia como se expresa por la ecuación (108). Según el método, la potencia de ruido se puede suprimir de

manera general. Sin embargo, el método no es necesariamente preferible si se conoce previamente que hay una fuente o fuentes de ruido que tienen una potencia fuerte en una o más direcciones particulares. Si este es el caso, se requiere un filtro que suprima fuertemente una o más direcciones conocidas particulares (es decir, puntos de supresión) en las que existe la fuente o fuentes de ruido. Por lo tanto, el método de diseño de filtro descrito aquí obtiene un filtro que minimiza la potencia de salida promedio del conformador de haces dada por la ecuación (107) (es decir, minimiza la potencia de salida promedio de sonidos de direcciones distintas de una posición (θ_s, D_h) y los puntos de supresión) bajo las condiciones de restricción que (1) el filtro pasa sonidos desde la posición (θ_s, D_h) en todas las bandas de frecuencia y que (2) el filtro suprime sonidos de B puntos de supresión conocidos $(\theta_{N1}, D_{G1}), (\theta_{N2}, D_{G2}), \dots, (\theta_{NB}, D_{GB})$. (B es un número entero predeterminado mayor o igual a 1) en todas las bandas de frecuencia. Permitamos que un conjunto de índices Φ de direcciones desde las cuales llega ruido sean denotadas por $\{1, 2, \dots, P\}$, entonces $N_j \in \{1, 2, \dots, P\}$ (donde $j \in \{1, 2, \dots, B\}$) y $B \leq P - 1$, como se ha descrito anteriormente. Permitamos que un conjunto de índices δ de distancias a fuentes de sonido sean denotados por $\{1, 2, \dots, G\}$, entonces $G_j \in \{1, 2, \dots, G\}$ (donde $j \in \{1, 2, \dots, B\}$) y $B \leq G-1$.

Permitamos que $a^{\rightarrow}(\omega, \theta_i, D_g) = [a_1(\omega, \theta_i, D_g), \dots, a_M(\omega, \theta_i, D_g)]^T$ sean funciones de transferencia entre una fuente de sonido que se supone que está situada en una posición (θ_i, D_g) y los M micrófonos a una frecuencia ω , en otras palabras, funciones de transferencia de un sonido desde una posición (θ_i, D_g) a una frecuencia ω que llega a los micrófonos de una agrupación de micrófonos, entonces se puede dar una condición de restricción por la ecuación (137). Aquí, para los índices i y g, $(i, g) \in \{(s, h), (N1, G1), (N2, G2), \dots, (NB, GB)\}$, las funciones de transferencia $a^{\rightarrow}(\omega, \theta_i, D_g) = [a_1(\omega, \theta_i, D_g), \dots, a_M(\omega, \theta_i, D_g)]^T$ se pueden dar por la ecuación (125) (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_i y D se sustituye por D_h), y $f_{i,g}(\omega)$ representa una característica de paso a frecuencia ω para una posición (θ_i, D_g) .

$$\bar{W}^H(\omega, \theta_s, D_h) \bar{a}(\omega, \theta_i, D_g) = f_{i,g}(\omega) \quad (i, g) \in \{(s, h), (N1, G1), (N2, G2), \dots, (NB, GB)\} \quad (137)$$

La ecuación (137) se puede representar como una matriz, por ejemplo, escrita como la ecuación (138). Aquí, $A^{\rightarrow}(\omega, \theta_s, D_h) = [[a^{\rightarrow}(\omega, \theta_s, D_h), a^{\rightarrow}(\omega, \theta_{N1}, D_{G1}), \dots, a^{\rightarrow}(\omega, \theta_{NB}, D_{GB})]$.

$$\bar{W}^H(\omega, \theta_s, D_h) \bar{A}(\omega, \theta_s, D_h) = \bar{F} \quad (138)$$

donde

$$\bar{F} = [f_{s,h}(\omega), f_{N1,G1}(\omega), \dots, f_{NB,GB}(\omega)]$$

Teniendo en consideración las condiciones de restricción que (1) el filtro pasa sonidos de la posición (θ_s, D_h) en todas las bandas de frecuencia y que (2) el filtro suprime sonidos de B puntos de supresión conocidos $(\theta_{N1}, D_{G1}), (\theta_{N2}, D_{G2}), \dots, (\theta_{NB}, D_{GB})$, en todas las bandas de frecuencia, se debería establecer idealmente $f_{s,h}(\omega) = 1,0$ y $f_{i,g}(\omega) = 0,0$ ($(i, g) \in \{(N1, G1), (N2, G2), \dots, (NB, GB)\}$). Esto significa que el filtro pasa completamente sonidos en todas las bandas de frecuencia desde la posición (θ_s, D_h) y bloquea completamente sonidos en todas las bandas de frecuencia de B puntos de supresión conocidos $(\theta_{N1}, D_{G1}), (\theta_{N2}, D_{G2}), \dots, (\theta_{NB}, D_{GB})$. En realidad, sin embargo, es difícil en algunas situaciones efectuar tal control como que pasan completamente todas las bandas de frecuencia o que se bloquean completamente todas las bandas de frecuencia. En tal caso, el valor absoluto de $f_{s,h}(\omega)$ se establece en un valor cercano a 1,0 y el valor absoluto de $f_{i,g}(\omega)$ ($(i, g) \in \{(N1, G1), (N2, G2), \dots, (NB, GB)\}$) se establece en un valor cercano a 0,0. Por supuesto, $f_{i,g_1}(\omega)$ y $f_{j,g_2}(\omega)$ ($i \neq j; i, j \in \{N1, N2, \dots, NB\}$) pueden ser iguales o diferentes.

Según el método de diseño de filtro descrito aquí, el filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ que es una solución óptima de la ecuación (107) bajo la condición de restricción dada por la ecuación (138) se puede dar por la ecuación (139) (véase la Referencia 3 enumerada a continuación). Mientras que se ha usado una matriz de correlación espacial $Q(\omega)$ que se puede dar por la ecuación (110c), se puede usar una matriz de correlación espacial dada por la ecuación (110a) o (110b).

$$\bar{W}(\omega, \theta_s, D_h) = Q^{-1}(\omega) \bar{A}(\omega, \theta_s, D_h) \left(\bar{A}^H(\omega, \theta_s, D_h) Q^{-1}(\omega) \bar{A}(\omega, \theta_s, D_h) \right)^{-1} \bar{F} \quad (139)$$

<4> Método de diseño de filtro que usa conformación de haces de retardo y suma

Suponiendo que los sonidos directos y reflejados que llegan son ondas planas, entonces un filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ se puede dar por la ecuación (140) según la conformación de haces de retardo y suma. Es decir, el filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ se puede obtener normalizando una función de transferencia $a^{\rightarrow}(\omega, \theta_s, D_h)$. La función de transferencia $a^{\rightarrow}(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ se puede dar por la ecuación (125) (para ser precisos, por la ecuación (125)

donde θ se sustituye por θ_s y D se sustituye por D_h). El método de diseño de filtro no logra necesariamente una alta precisión de filtrado, sino que solamente requiere una pequeña cantidad de cálculo.

$$\vec{W}(\omega, \theta_s, D_h) = \frac{\vec{a}(\omega, \theta_s, D_h)}{\vec{a}^H(\omega, \theta_s, D_h)\vec{a}(\omega, \theta_s, D_h)} \quad (140)$$

<5> Método de diseño de filtro que usa el método de máxima verosimilitud

- 5 Excluyendo la información espacial que se refiere a sonidos de una dirección del objetivo de una matriz de correlación espacial $Q(\omega, D_h)$ en el método MVDR descrito anteriormente, se puede mejorar la flexibilidad de supresión de ruido y se puede suprimir aún más la potencia de ruido. Por lo tanto, en el método de diseño de filtro descrito aquí, la matriz de correlación espacial $Q(\omega, D_h)$ se escribe como el segundo término del lado derecho de la ecuación (110a), es decir, la ecuación (110d). Un filtro $W^-(\omega, \theta_s, D_h)$ se puede dar por la ecuación (109) o (139).
10 Aquí, la matriz de correlación espacial incluida en las ecuaciones (109) y (139) es una matriz de correlación espacial dada por la ecuación (110d).

$$Q(\omega, D_h) = \sum_{p \in \{1, \dots, P-1\}} \vec{a}(\omega, \theta_p, D_h)\vec{a}^H(\omega, \theta_p, D_h) \quad (110d)$$

- 15 Alternativamente, la información espacial que se refiere a sonidos de la posición (θ_s, D_h) se puede excluir de la matriz de correlación espacial $Q(\omega)$. En ese caso, una matriz de correlación espacial $Q(\omega)$ se da por la ecuación (110e) en el método de diseño de filtro descrito aquí. Un filtro $W^-(\omega, \theta_s, D_h)$ se puede dar por la ecuación (109) o (139). Aquí, la matriz de correlación espacial incluida en las ecuaciones (109) y (139) se da por la ecuación (110e).

$$Q(\omega) = \left(\sum_{\phi \in \Phi} \sum_{\delta \in \Delta} \vec{a}(\omega, \theta_\phi, D_\delta)\vec{a}^H(\omega, \theta_\phi, D_\delta) \right) - \vec{a}(\omega, \theta_s, D_h)\vec{a}^H(\omega, \theta_s, D_h) \quad (110e)$$

<6> Método de diseño de filtro que usa el método AMNOR

- 20 El método AMNOR obtiene un filtro que permite alguna cantidad de descomposición D de un sonido de una dirección del objetivo mediante una solución de compromiso de la cantidad de descomposición D del sonido de la dirección del objetivo frente a la potencia del ruido restante en una señal de salida de filtro (por ejemplo, la cantidad de descomposición D se mantiene en un cierto umbral D^\wedge o menos) y, cuando se introduce una señal mezclada de [a] una señal producida aplicando funciones de transferencia entre una fuente de sonido y los micrófonos a una señal virtual (en lo sucesivo conocida como la señal virtual) de una dirección del objetivo y [b] ruido (obtenido
25 mediante observación con M micrófonos en un entorno ruidoso sin un sonido de la dirección del objetivo), emite una señal de salida de filtro que reproduce mejor la señal virtual en términos de error cuadrático mínimo (es decir, se minimiza la potencia de ruido contenido en una señal de salida de filtro).

- 30 El método de diseño de filtro descrito aquí incorpora el concepto de distancia en el método AMNOR y se puede considerar similar al método AMNOR. Específicamente, el método obtiene un filtro que permite alguna cantidad de descomposición D de un sonido de una posición (θ_s, D_h) mediante una solución de compromiso de la cantidad de descomposición D del sonido de la posición (θ_s, D_h) frente a la potencia de ruido restante en una señal de salida de filtro (por ejemplo, la cantidad de descomposición D se mantiene en cierto umbral D^\wedge o menos) y, cuando se introduce una señal mezclada de [a] una señal producida aplicando funciones de transferencia entre una fuente de sonido y los micrófonos a una señal de objetivo virtual de una posición (θ_s, D_h) (en lo sucesivo conocida como la
35 señal de objetivo virtual) y [b] ruido (obtenido mediante observación con M micrófonos en un entorno ruidoso sin un sonido de la posición (θ_s, D_h)), emite una señal de salida de filtro que reproduce la mejor señal de objetivo virtual en términos de error cuadrático mínimo (es decir, se minimiza la potencia de ruido contenido en una señal de salida de filtro).

- 40 Según el método de diseño de filtro descrito aquí, un filtro $W^-(\omega, \theta_s, D_h)$ se puede dar por la ecuación (141) como en el método AMNOR (véase la Referencia 4 enumerada a continuación). Aquí, $R_{ss}(\omega)$ se puede dar por la ecuación (126) y $R_{nn}(\omega)$ se puede dar por la ecuación (127). Las funciones de transferencia $a^-(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ se puede dar por la ecuación (125) (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_s y D se sustituye por D_h).

$$\vec{W}(\omega, \theta_s, D_h) = P_s \vec{a}(\omega, \theta_s, D_h) (R_{nn}(\omega) + P_s R_{ss}(\omega))^{-1} \quad (141)$$

5 P_s es un coeficiente que asigna una ponderación al nivel de la señal de objetivo virtual y denominado el nivel de
 señal de objetivo virtual. El nivel de señal de objetivo virtual P_s es una constante que no es dependiente de las
 frecuencias. El nivel de señal de objetivo virtual P_s se puede determinar empíricamente o se puede determinar de
 modo que la diferencia entre la cantidad de descomposición D de un sonido desde la posición (θ_s, D_h) y el umbral D^\wedge
 10 está dentro de un margen de error predeterminado arbitrariamente. Se describirá este último caso. La respuesta en
 frecuencia $F(\omega)$ del filtro $W^\rightarrow(\omega, \theta_s, D_h)$ a un sonido de una posición (θ_s, D_h) se puede dar por la ecuación (142).
 Permitamos que la cantidad de descomposición $D(P_s)$ cuando se usa el filtro $W^\rightarrow(\omega, \theta_s, D_h)$ dado por la ecuación
 (141) sea denotada por $D(P_s)$, entonces la cantidad de descomposición $D(P_s)$ se puede definir por la ecuación (143).
 Aquí, ω_0 representa el límite superior de la frecuencia ω (típicamente, una frecuencia más alta adyacente a una
 15 frecuencia discreta ω). La cantidad de descomposición $D(P_s)$ es una función monótonamente decreciente de P_s . Por
 lo tanto, un nivel de señal de objetivo virtual P_s de manera que la diferencia entre la cantidad de descomposición
 $D(P_s)$ y el umbral D^\wedge está dentro de un margen de error predeterminado arbitrariamente se puede obtener
 obteniendo repetidamente la cantidad de descomposición $D(P_s)$ mientras que se cambia P_s con la monotonía de
 $D(P_s)$.

$$F(\omega) = \vec{W}^H(\omega, \theta_s, D_h) \vec{a}(\omega, \theta_s, D_h) \quad (142)$$

$$D(P_s) = \frac{1}{2\omega_0} \int_{-\omega_0}^{\omega_0} |1 - F(\omega)|^2 d\omega \quad (143)$$

15 <Variación>

En la descripción precedente, las matrices de correlación espacial $Q(\omega)$, $R_{ss}(\omega)$ y $R_{nn}(\omega)$ se expresan usando
 funciones de transferencia. Sin embargo, las matrices de correlación espacial $Q(\omega)$, $R_{ss}(\omega)$ y $R_{nn}(\omega)$ también se
 pueden expresar usando las señales en el dominio de frecuencia $X^\rightarrow(\omega, k)$ descritas anteriormente. Mientras que la
 20 matriz de correlación espacial $Q(\omega)$ se describirá a continuación, la siguiente descripción también se aplica a $R_{ss}(\omega)$
 y $R_{nn}(\omega)$. ($Q(\omega)$ se puede sustituir por $R_{ss}(\omega)$ o $R_{nn}(\omega)$). La matriz de correlación espacial $R_{ss}(\omega)$ se puede obtener
 usando representaciones en el dominio de frecuencia de señales analógicas obtenidas mediante observación con
 una agrupación de micrófonos (que incluye M micrófonos) en un entorno donde solamente existen sonidos de
 25 una posición (θ_s, D_h) . La matriz de correlación espacial $R_{nn}(\omega)$ se puede obtener usando representaciones en el dominio
 de frecuencia de una señal analógica obtenida mediante observación con una agrupación de micrófonos (que
 incluye M micrófonos) en un entorno donde no existen sonidos de una posición (θ_s, D_h) (es decir, un entorno
 ruidoso).

La matriz de correlación espacial $Q(\omega)$ que usa las señales en el dominio de frecuencia $X^\rightarrow(\omega, k) = [X_1(\omega, k), \dots,$
 $X_M(\omega, k)]^T$ se puede dar por la ecuación (144). Aquí, el operador $E[\cdot]$ representa una operación de promediado
 30 estadístico. Cuando se ve una serie de tiempo discreta de una señal analógica recibida con una agrupación de
 micrófonos (que incluye M micrófonos) como un proceso estocástico, el operador $E[\cdot]$ representa una operación de
 valor de media aritmética (valor esperado) si el proceso estocástico es un denominado proceso estacionario en
 sentido amplio o un proceso estacionario de segundo orden. En este caso, la matriz de correlación espacial $Q(\omega)$ se
 puede dar por la ecuación (145) usando las señales en el dominio de frecuencia $X^\rightarrow(\omega, k - i)$ ($i = 0, 1, \dots, \zeta - 1$) de
 35 un total de ζ tramas actuales y pasadas almacenadas en una memoria, por ejemplo. Cuando $i = 0$, un trama de
 orden k es la trama actual. Obsérvese que la matriz de correlación espacial $Q(\omega)$ dada por la ecuación (144) o (145)
 se puede recalculer para cada trama o se puede calcular en un intervalo regular o irregular, o se puede calcular
 antes de la implementación de una realización, que se describirá más tarde (especialmente cuando se usa $R_{ss}(\omega)$ o
 40 $R_{nn}(\omega)$, la matriz de correlación espacial $Q(\omega)$ se calcula preferiblemente de antemano usando señales en el dominio
 de frecuencia obtenidas antes de la implementación de la realización). Si la matriz de correlación espacial $Q(\omega)$ se
 recalcula para cada trama, la matriz de correlación espacial $Q(\omega)$ depende de las tramas actuales y pasadas y, por
 lo tanto, la matriz de correlación espacial se representará explícitamente como $Q(\omega, k)$ como en las ecuaciones
 (144a) y (145a).

$$Q(\omega) = E[\bar{X}(\omega, k)\bar{X}^H(\omega, k)] \quad (144)$$

$$Q(\omega) = \sum_{i=0}^{\zeta-1} \bar{X}(\omega, k-i)\bar{X}^H(\omega, k-i) \quad (145)$$

$$Q(\omega, k) = E[\bar{X}(\omega, k)\bar{X}^H(\omega, k)] \quad (144a)$$

$$Q(\omega, k) = \sum_{i=0}^{\zeta-1} \bar{X}(\omega, k-i)\bar{X}^H(\omega, k-i) \quad (145a)$$

Si se usa la matriz de correlación espacial $Q(\omega, k)$ representada por la ecuación (144a) o (145a), el filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ también depende de las tramas actuales y pasadas y, por lo tanto, se representa explícitamente como $W^{\rightarrow}(\omega, \theta_s, D_h, k)$. Entonces, un filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ representado por cualquiera de las ecuaciones (109), (132), (133), (136), (139) y (141) descritas con los métodos de diseño de filtro descritos anteriormente se reescribe como las ecuaciones (109m), (132m), (133m), (136m), (139m) o (141m).

$$\bar{W}(\omega, \theta_s, D_h, k) = \frac{Q^{-1}(\omega, k)\bar{a}(\omega, \theta_s, D_h)}{\bar{a}^H(\omega, \theta_s, D_h)Q^{-1}(\omega, k)\bar{a}(\omega, \theta_s, D_h)} \quad (109m)$$

$$\bar{W}(\omega, \theta_s, D_h, k) = R_{nn}^{-1}(\omega, k)\bar{a}(\omega, \theta_s, D_h) \quad (132m)$$

$$\bar{W}(\omega, \theta_s, D_h, k) = R_{xx}^{-1}(\omega, k)\bar{a}(\omega, \theta_s, D_h) \quad (133m)$$

$$\bar{W}(\omega, \theta_s, D_h, k) = R_{xx}^{-1}(\omega, k)\bar{G} \quad (136m)$$

$$\bar{W}(\omega, \theta_s, D_h, k) = Q^{-1}(\omega, k)\bar{A}(\omega, \theta_s, D_h)\left(\bar{A}^H(\omega, \theta_s, D_h)Q^{-1}(\omega, k)\bar{A}(\omega, \theta_s, D_h)\right)^{-1}\bar{F} \quad (139m)$$

$$\bar{W}(\omega, \theta_s, D_h, k) = P_s\bar{a}(\omega, \theta_s, D_h)(R_{nn}(\omega, k) + P_sR_{ss}(\omega, k))^{-1} \quad (141m)$$

<<Primera realización de la técnica de realce de punto de sonido>>

Las Fig. 19 y 20 ilustran una configuración funcional y un flujo de proceso de una primera realización de una técnica de realce de punto de sonido de la presente invención. Un aparato 3 de realce de punto de sonido de la primera realización incluye un convertidor 610 AD, un generador 620 de tramas, una sección 630 de transformada en el dominio de frecuencia, una sección 640 de aplicación de filtro, una sección 650 de transformada en el dominio de tiempo, una sección 660 de diseño de filtro y almacenamiento 690.

[Paso S21]

La sección 660 de diseño de filtro calcula de antemano un filtro $W^{\rightarrow}(\omega, \theta_i, D_g)$ para cada frecuencia para cada una de las posibles posiciones discretas (θ_i, D_g) a partir de las cuales pueden llegar los sonidos a ser realizados. La sección 660 de diseño de filtro calcula los filtros $W^{\rightarrow}(\omega, \theta_1, D_1), \dots, W^{\rightarrow}(\omega, \theta_i, D_1), \dots, W^{\rightarrow}(\omega, \theta_i, D_1), \dots, W^{\rightarrow}(\omega, \theta_1, D_2), \dots, W^{\rightarrow}(\omega, \theta_i, D_2), \dots, W^{\rightarrow}(\omega, \theta_i, D_2), \dots, W^{\rightarrow}(\omega, \theta_1, D_g), \dots, W^{\rightarrow}(\omega, \theta_i, D_g), \dots, W^{\rightarrow}(\omega, \theta_i, D_g), \dots, W^{\rightarrow}(\omega, \theta_1, D_G), \dots, W^{\rightarrow}(\omega, \theta_i, D_G), \dots, W^{\rightarrow}(\omega, \theta_i, D_G)$ ($1 \leq i \leq I, 1 \leq g \leq G, \omega \in \Omega$; i y g son números enteros y Ω es un conjunto de frecuencias ω), donde I es el número total de direcciones discretas a partir de las cuales pueden llegar los sonidos a ser realizados (I es un número entero predeterminado mayor o igual a 1 y satisface $I \leq P$) y G es el número de las distancias discretas (G un número entero predeterminado mayor o igual a 1).

Para hacerlo así, las funciones de transferencia $a^{\rightarrow}(\omega, \theta_i, D_g) = [a_1(\omega, \theta_i, D_g), \dots, a_M(\omega, \theta_i, D_g)]^T$ ($1 \leq i \leq I, 1 \leq g \leq G, \omega \in \Omega$) necesitan ser obtenidas excepto para el caso de <Variación> descrito anteriormente. Las funciones de transferencia $a^{\rightarrow}(\omega, \theta_i, D_g) = [a_1(\omega, \theta_i, D_g), \dots, a_M(\omega, \theta_i, D_g)]^T$ se pueden calcular prácticamente según la ecuación (125) (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_i y D se sustituye por D_g) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos y la información de entorno, tal como la relación de posición de un objeto reflector, tal como un reflector, suelo, paredes y techo a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y la reflectancia acústica del objeto reflector. Obsérvese que si se usa <3> el método de diseño de filtro que usa MVDR

con uno o más puntos de supresión como condición de restricción, los índices (i, g) de las direcciones usadas para calcular las funciones de transferencia $a^{\rightarrow}(\omega, \theta_i, D_g)$ ($1 \leq i \leq I, 1 \leq g \leq G, \omega \in \Omega$) cubren preferiblemente todos los índices (N1, G1), (N2, G2), ..., (NB, GB) de direcciones de al menos B posiciones de supresión. En otras palabras, los B índices N1, N2, ..., NB se establecen en cualquiera de diferentes números enteros mayores o iguales a 1 y menores o iguales a I y los B índices G1, G2, ..., GB se establecen en cualquiera de diferentes números enteros mayores o iguales a 1 y menores o iguales a G.

El número Ξ de sonidos reflejados se establece en un número entero que satisface $1 \leq \Xi$. El número Ξ no está limitado y se puede establecer en un valor adecuado según la capacidad de cálculo y otros factores.

Para calcular vectores de dirección, se pueden usar las ecuaciones (125a), (125b), (126a) o (126b), por ejemplo. Obsérvese que las funciones de transferencia obtenidas mediante mediciones reales en un entorno real, por ejemplo, se pueden usar para diseñar los filtros en lugar de usar la ecuación (125).

Entonces, $W^{\rightarrow}(\omega, \theta_i, D_g)$ ($1 \leq i \leq I, 1 \leq g \leq G$) se obtiene según cualquiera de las ecuaciones (109), (109a), (132), (133), (136), (139), (140) y (141), por ejemplo, usando las funciones de transferencia $a^{\rightarrow}(\omega, \theta_i, D_g)$, excepto para el caso descrito en <Variación>. Obsérvese que si se usa la ecuación (109), (109a), (133), (136) o (139), la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (110b), excepto para el caso descrito con respecto a <5> el método de diseño de filtro que usa el método de máxima verosimilitud. Si la ecuación (109), (109a), (133), (136) o (139) se usa según <5> el método de diseño de filtro que usa el método de máxima verosimilitud descrito anteriormente, la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (110c) o (110d). Si se usa la ecuación (132), la matriz de correlación espacial $R_{nn}(\omega)$ se puede calcular según la ecuación (130). $I \times G \times |\Omega|$ filtros $W^{\rightarrow}(\omega, \theta_i, D_g)$ ($1 \leq i \leq I, 1 \leq g \leq G, \omega \in \Omega$) se almacenan en el almacenamiento 690, donde $|\Omega|$ representa el número de elementos del conjunto Ω .

[Paso S22]

Los M micrófonos 200-1, ..., 200-M que componen la agrupación de micrófonos se usan para captar sonidos, donde M es un número entero mayor o igual a 2.

No hay ninguna restricción en la disposición de los M micrófonos. Sin embargo, una disposición bidimensional o tridimensional de los M micrófonos tiene la ventaja de eliminar la incertidumbre de una dirección desde la cual llegan los sonidos a ser realzados. Es decir, una disposición plana o esférica de los micrófonos puede evitar el problema con una disposición lineal horizontal de los M micrófonos que un sonido que llega desde una dirección frontal no se pueda distinguir de un sonido que llega desde la derecha arriba, por ejemplo. Con el fin de proporcionar un intervalo amplio de direcciones que se pueden establecer como direcciones de captura de sonido, cada micrófono preferiblemente tiene una directividad capaz de captar sonidos con un cierto nivel de presión de sonido en direcciones de objetivo potenciales θ_s que son direcciones de captura de sonido. Por consiguiente, son preferibles micrófonos que tienen una directividad relativamente débil, tales como micrófonos omnidireccionales o micrófonos unidireccionales.

[Paso S23]

El convertidor 610 AD convierte las señales analógicas (señales de captura) captadas con los M micrófonos 200-1, ..., 200-M en señales digitales $x^{\rightarrow}(t) = [x_1(t), \dots, x_M(t)]^T$, donde t representa el índice de un tiempo discreto.

[Paso S24]

El generador 620 de tramas toma las entradas de las señales digitales $x^{\rightarrow}(t) = [x_1(t), \dots, x_M(t)]^T$ emitidas desde el convertidor 610 AD, almacena N muestras en un almacenador temporal sobre una base canal por canal, y emite señales digitales $x^{\rightarrow}(k) = [x^{\rightarrow}_1(k), \dots, x^{\rightarrow}_M(k)]^T$ en tramas, donde k es un índice de un número de tiempo de trama y $x^{\rightarrow}_m(k) = [x_m((k-1)N+1), \dots, x_m(kN)]$ ($1 \leq m \leq M$). N depende de la frecuencia de muestreo y 512 es adecuada para muestrear a 16 kHz.

[Paso S25]

La sección 630 de transformada en el dominio de frecuencia transforma las señales digitales $x^{\rightarrow}(k)$ en tramas a señales en el dominio de frecuencia $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y emite las señales en el dominio de frecuencia, donde ω es el índice de una frecuencia discreta. Una forma de transformar una señal en el dominio de tiempo a una señal en el dominio de frecuencia es la transformada de Fourier discreta rápida. Sin embargo, la forma de transformar la señal no se limita a ésta. Se puede usar otro método para transformar a una señal en el dominio de frecuencia. La señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k)$ se emite para cada frecuencia ω y trama k a la vez.

[Paso S26]

La sección 640 de aplicación de filtro aplica el filtro $W^{\rightarrow}(\omega, \theta_s, D_h)$ correspondiente a una posición (θ_s, D_h) a ser realzada a la señal en el dominio de frecuencia $X^{\rightarrow}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ en cada trama k para cada

frecuencia $\omega \in \Omega$ y emite una señal de salida $Y(\omega, k, \theta_s, D_h)$ (véase la ecuación (146)). Los índices s y h de la posición (θ_s, D_h) son $s \in \{1, \dots, l\}$ y $h \in \{1, \dots, G\}$ y el filtro $W^-(\omega, \theta_s, D_h)$ se almacena en el almacenamiento 690. Por lo tanto, la sección 640 de aplicación de filtro solamente tiene que recuperar el filtro $W^-(\omega, \theta_s, D_h)$ que corresponde a la posición (θ_s, D_h) a ser realizada del almacenamiento 690, por ejemplo, en el proceso en el paso S26. Si el índice s de la dirección θ_s no pertenece al conjunto $\{1, \dots, l\}$ o el índice h de dirección D_h no pertenece al conjunto $\{1, \dots, G\}$, es decir, el filtro $W^-(\omega, \theta_s, D_h)$ que corresponde a la posición (θ_s, D_h) no se ha calculado en el proceso en el paso S21, la sección 660 de diseño de filtro puede calcular en este momento el filtro $W^-(\omega, \theta_s, D_h)$ que corresponde a la posición (θ_s, D_h) o se puede usar el filtro $W^-(\omega, \theta_s, D_h)$ o $W^-(\omega, \theta_s, D_h)$ o $W^-(\omega, \theta_s, D_h)$ que corresponde a una dirección θ_s cerca de la dirección θ_s y/o una distancia D_h cerca de la distancia D_h .

$$Y(\omega, k, \theta_s, D_h) = \bar{W}^H(\omega, \theta_s, D_h) \bar{X}(\omega, k) \quad \forall \omega \in \Omega \quad (146)$$

[Paso S27]

La sección 650 de transformada en el dominio de tiempo transforma la señal de salida $Y(\omega, k, \theta_s, D_h)$ de cada frecuencia $\omega \in \Omega$ en un trama de orden k en un dominio de tiempo para obtener una señal $y(k)$ de trama en el dominio de tiempo en la trama de orden k, entonces combina las señales $y(k)$ en el dominio de tiempo de trama obtenidas en el orden del índice de número de tiempo de trama, y emite una señal $y(t)$ en el dominio de tiempo en la que se realiza el sonido de una posición (θ_s, D_h) . El método para transformar una señal en el dominio de frecuencia en una señal en el dominio de tiempo es una transformada inversa de la transformada usada en el proceso en el paso S25 y puede ser una transformada de Fourier inversa discreta rápida, por ejemplo.

Mientras que la primera realización se ha descrito aquí en la que los filtros $W^-(\omega, \theta_i, D_g)$ se calculan de antemano en el proceso en el paso S21, la sección 660 de diseño de filtro puede calcular el filtro $W^-(\omega, \theta_s, D_h)$ para cada frecuencia después de que se determina la posición (θ_s, D_h) , dependiendo de la capacidad de cálculo del aparato 3 de realce de punto de sonido.

<<Segunda realización de la técnica de realce de punto de sonido>>

Las Fig. 21 y 22 ilustran una configuración funcional y un flujo de proceso de la segunda realización de una técnica de realce de punto de sonido de la presente invención. Un aparato 4 de realce de punto de sonido de la segunda realización incluye un convertidor 610 AD, un generador 620 de tramas, una sección 630 de transformada en el dominio de frecuencia, una sección 640 de aplicación de filtro, una sección 650 de transformada en el dominio de tiempo, una sección 661 de cálculo de filtro y un almacenamiento 690.

[Paso S31]

M micrófonos 200-1, ..., 200-M que componen una agrupación de micrófonos se usan para captar sonidos, donde M es un número entero mayor o igual a 2. La disposición de los M micrófonos es como se describe en la primera realización.

[Paso S32]

El convertidor 610 AD convierte señales analógicas (señales de captura) captadas con los M micrófonos 200-1, ..., 200-M en señales digitales $x^-(t) = [x_1(t), \dots, x_M(t)]^T$, donde t representa el índice de un tiempo discreto.

[Paso S33]

El generador 620 de tramas toma las entradas de las señales digitales $x^-(t) = [x_1(t), \dots, x_M(t)]^T$ emitidas desde el convertidor 610 AD, almacena N muestras en un almacenador temporal sobre una base canal por canal, y emite las señales digitales $x^-(k) = [x_1^-(k), \dots, x_M^-(k)]^T$ en tramas, donde k es un índice de un número de tiempo de trama y $x_m^-(k) = [x_m((k-1)N+1), \dots, x_m(kN)]$ ($1 \leq m \leq M$). N depende de la frecuencia de muestreo y 512 es adecuada para muestrear a 16 kHz.

[Paso S34]

La sección 630 de transformada en el dominio de frecuencia transforma las señales digitales $x^-(k)$ en tramas en señales en el dominio de frecuencia $X^-(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ y emite las señales en el dominio de frecuencia, donde ω es un índice de una frecuencia discreta. Una forma de transformar una señal en el dominio de tiempo en una señal en el dominio de frecuencia es una transformada de Fourier discreta rápida. Sin embargo, la forma de transformar la señal no está limitada a ésta. Se puede usar otro método para transformar a una señal en el dominio de frecuencia. La señal en el dominio de frecuencia $X^-(\omega, k)$ se emite para cada frecuencia ω y trama k a la vez.

[Paso S35]

La sección 661 de cálculo de filtro calcula el filtro $W^{-1}(\omega, \theta_s, D_h, k)$ ($\omega \in \Omega$; Ω es un conjunto de frecuencias ω) que corresponde a la posición (θ_s, D_h) a ser usado en un trama actual de orden k .

Para hacerlo así, necesitan ser proporcionadas las funciones de transferencia $a^{-1}(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ ($\omega \in \Omega$). Las funciones de transferencia $a^{-1}(\omega, \theta_s, D_h) = [a_1(\omega, \theta_s, D_h), \dots, a_M(\omega, \theta_s, D_h)]^T$ se pueden calcular prácticamente según la ecuación (17a) (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_s y D se sustituye por D_h) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos y la información de entorno, tal como la relación de posición de un objeto reflector tal como un reflector, el suelo, las paredes o el techo a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y la reflectancia acústica del objeto reflector. Obsérvese que si se usa <3> el método de diseño de filtro que usa MVDR con uno o más puntos de supresión como condición de restricción, también necesitan ser obtenidas las funciones de transferencia $a^{-1}(\omega, \theta_{N_j}, D_{G_j})$ ($1 \leq j \leq B$, $\omega \in \Omega$). Las funciones de transferencia se pueden calcular prácticamente según la ecuación (125) (para ser precisos, por la ecuación (125) donde θ se sustituye por θ_{N_j} y D se sustituye por D_{G_j}) sobre la base de la disposición de los micrófonos en la agrupación de micrófonos y la información del entorno tal como la relación de posición de un objeto reflector, tal como un reflector, el suelo, una pared o el techo a la agrupación de micrófonos, la diferencia de tiempo de llegada entre un sonido directo y un sonido reflejado de orden ξ ($1 \leq \xi \leq \Xi$) y la reflectancia acústica del objeto reflector.

El número Ξ de sonidos reflejados se establece en un número entero que satisface $1 \leq \Xi$. El número Ξ no está limitado y se puede establecer en un valor adecuado según la capacidad de cálculo y otros factores.

Para calcular los vectores de dirección, se pueden usar la ecuación (125a), (125b), (126a) o (126b), por ejemplo. Obsérvese que las funciones de transferencia obtenidas mediante mediciones reales en un entorno real, por ejemplo, se pueden usar para diseñar los filtros en lugar de usar la ecuación (125).

Entonces, la sección 661 de cálculo de filtro calcula los filtros $W^{-1}(\omega, \theta_s, D_h, k)$ ($\omega \in \Omega$) según cualquiera de las ecuaciones (109m), (132m), (133m), (136m), (139m) y (141m) usando las funciones de transferencia $a^{-1}(\omega, \theta_s, D_h)$ ($\omega \in \Omega$) y, si se necesita, las funciones de transferencia $a^{-1}(\omega, \theta_{N_j}, D_{G_j})$ ($1 \leq j \leq B$, $\omega \in \Omega$). Obsérvese que la matriz de correlación espacial $Q(\omega)$ (o $R_{xx}(\omega)$) se puede calcular según la ecuación (144a) o (145a). En el cálculo de la matriz de correlación espacial $Q(\omega)$, se usan las señales en el dominio de frecuencia $X^{-1}(\omega, k - i)$ ($i = 0, 1, \dots, \zeta - 1$) de un total de ζ tramas actuales y pasadas almacenadas en el almacenamiento 690, por ejemplo.

[Paso S36]

La sección 640 de aplicación de filtro aplica el filtro $W^{-1}(\omega, \theta_s, D_h, k)$ correspondiente a la dirección del objetivo θ_s a ser realizada a la señal en el dominio de frecuencia $X^{-1}(\omega, k) = [X_1(\omega, k), \dots, X_M(\omega, k)]^T$ en cada trama k para cada frecuencia $\omega \in \Omega$ y emite una señal de salida $Y(\omega, k, \theta_s, D_h)$ (véase la ecuación (147)).

$$Y(\omega, k, \theta_s, D_h) = \bar{W}^H(\omega, \theta_s, D_h, k) \bar{X}(\omega, k) \quad \forall \omega \in \Omega \quad (147)$$

[Paso S37]

La sección 650 de transformada en el dominio del tiempo transforma la señal de salida $Y(\omega, k, \theta_s, D_h)$ de cada frecuencia $\omega \in \Omega$ de un trama de orden k en un dominio de tiempo para obtener una señal $y(k)$ de trama en el dominio de tiempo en la trama de orden k , entonces combina las señales $y(k)$ en el dominio de tiempo de tramas obtenidas en el orden del índice de número de tiempo de trama, y emite una señal $y(t)$ en el dominio de tiempo en la que se realiza el sonido de la posición (θ_s, D_h) . El método para transformar una señal en el dominio de frecuencia en una señal en el dominio de tiempo es una transformada inversa de la transformada usada en el proceso en el paso S34 y puede ser una transformada de Fourier inversa discreta rápida, por ejemplo.

Un filtro $W^{-1}(\omega, \theta_i)$ que corresponde a una dirección θ_i se puede calcular por $\sum_{g=1}^G \beta_g W^{-1}(\omega, \theta_i, D_g)$ en la técnica de realce de punto de sonido, donde β_g [$1 \leq g \leq G$] es un factor de ponderación, que satisface preferiblemente $\sum_{g=1}^G \beta_g = 1$ y preferiblemente $0 \leq \beta_g$ [$1 \leq g \leq G$]. Obsérvese que el filtro $W^{-1}(\omega, \theta_i, D_g)$ puede ser un filtro representado usando funciones de transferencia medidas en un entorno real.

[Ejemplo experimental de técnica de realce de punto de sonido]

Se describirán los resultados de ejemplos experimentales sobre el realce de punto de sonido según la primera realización de la técnica de realce de punto de sonido de la presente invención (el método de respuesta sin distorsión de varianza mínima (MVDR) bajo una única condición de restricción). Los experimentos se dirigieron en el mismo entorno ilustrado en la Fig. 9. Como se ilustra en la Fig. 9, están dispuestos linealmente 24 micrófonos y está colocado un reflector 300 de modo que la dirección a lo largo de la cual los micrófonos en la agrupación de micrófonos lineales es normal al reflector 300. Si bien no hay ninguna restricción en la forma del reflector 300, se usó un reflector plano rígido semigrueso que tenía un tamaño de 1,0 m x 1,0 m. La distancia entre micrófonos adyacentes era 4 cm y la reflectancia α del reflector 300 era 0,8. Una fuente de sonido se situó en una dirección θ_s de 45 grados a una distancia D_h de 1,13 m. La Fig. 23A muestra la directividad (en un dominio bidimensional) de un conformador de haces de varianza mínima obtenido como resultado del experimento donde no se colocó un reflector 300; la Fig. 23B muestra la directividad (en un dominio bidimensional) de un conformador de haces de varianza mínima obtenido como resultado del experimento donde se colocó un reflector 300. La presión de sonido [en dB] se representa como sombras, donde las regiones más blancas representan presiones más altas de los sonidos captados. Idealmente, si solamente la posición en una dirección de 45 grados a una distancia de 1,13 m es blanca y las otras regiones están más cerca de negras, se puede decir que se ha logrado un realce de punto de los sonidos deseados. La comparación entre los resultados experimentales en las Fig. 23A y 23B muestra que el realce de punto de los sonidos deseados no se puede lograr suficientemente sin un reflector 300 y el realce de punto de los sonidos deseados se puede lograr con un reflector 300.

<Aplicaciones de ejemplo>

Hablando en sentido figurado, la técnica de realce de punto de sonido es equivalente a la generación de una imagen clara a partir de una imagen difusa, borrosa y es útil para obtener información detallada acerca de un campo acústico. La siguiente es una descripción de ejemplos de servicios donde es útil la técnica de realce de punto de sonido de la presente invención.

Un primer ejemplo es la creación de contenidos que son combinación de audio y video. El uso de una realización de la técnica de realce de punto de sonido de la presente invención permite que el sonido del objetivo desde una gran distancia sea realzado claramente incluso en un entorno ruidoso con sonidos de ruido (sonidos distintos de los sonidos del objetivo). Por lo tanto, por ejemplo, sonidos en un área particular correspondiente a una imagen en movimiento acercada de un jugador de fútbol que dribla que se filmó desde el exterior del campo se pueden añadir a la imagen en movimiento.

Un segundo ejemplo es una aplicación para una videoconferencia (o una teleconferencia de audio). Cuando se mantiene una conferencia en una sala pequeña, la voz de un orador humano se puede realzar hasta un cierto grado con varios micrófonos según una técnica convencional. Sin embargo, en una sala de conferencias grande (por ejemplo, un espacio grande donde hay oradores humanos a una distancia de 5 m o más desde los micrófonos), es difícil realzar claramente la voz de un orador humano a una distancia con las técnicas convencionales mediante el método convencional y necesita ser colocado un micrófono enfrente a cada orador humano. Por el contrario, el uso de una realización de la técnica de realce de punto de sonido de la presente invención es capaz de realzar claramente sonidos de un área particular más alejada de un área particular y por lo tanto permite la construcción de un sistema de videoconferencia que es utilizable en una sala de conferencias grande sin tener que colocar un micrófono enfrente de cada orador humano. Además, dado que se pueden realzar los sonidos de un área en particular, se pueden relajar las restricciones en las ubicaciones de los participantes en la conferencia con respecto a las ubicaciones de los micrófonos.

<Configuraciones de la implementación de la técnica de realce de sonido>

Las configuraciones ejemplares de implementaciones de las técnicas de realce de sonido de la presente invención se describirán a continuación con referencia a las Fig. 24 a 28. Si bien las agrupaciones de micrófonos en los ejemplos se representan como agrupaciones de micrófonos lineales, las agrupaciones de micrófonos no están limitadas a las configuraciones de agrupaciones de micrófonos lineales.

En una configuración ejemplar de una implementación ilustrada en las Fig. 24A, 24B y 24C, M micrófonos 200-1, ..., 200-M que componen una agrupación de micrófonos lineales están fijados a una placa 400 de soporte plana rectangular y en este estado el orificio de captura de sonido de cada micrófono está colocado en una superficie plana (en lo sucesivo conocida como la superficie de abertura) de la placa 400 de soporte (M = 13 en los ejemplos representados). Obsérvese que las líneas de cableado conectadas a los micrófonos 200-1, ..., 200-M no se representan. Un reflector 300 de placa plana rectangular está fijado en un borde de la placa 400 de soporte de tal manera que la dirección en la que están dispuestos los micrófonos 200-1, ..., 200-M es normal al reflector 300. La superficie de abertura de la placa 400 de soporte es un ángulo de 90 grados al reflector 300. En la configuración ejemplar ilustrada en las Fig. 24A, 24B y 24C, las propiedades preferibles del reflector 300 son las mismas que las del reflector descrito anteriormente. No hay restricciones sobre las propiedades de la placa 400 de soporte; es esencial solamente que la placa 400 de soporte sea lo suficientemente rígida para fijar firmemente los micrófonos 200-1, ..., 200-M.

En una configuración ejemplar ilustrada en la Fig. 25A, un eje 410 está fijado a un borde de la placa 400 de soporte y un reflector 300 está unido de manera giratoria al eje 410. En esta configuración ejemplar, se puede cambiar la colocación geométrica del reflector 300 en la agrupación de micrófonos.

5 En una configuración ejemplar ilustrada en la Fig. 25B, se añaden dos reflectores 310 y 320 adicionales a la configuración ilustrada en las Fig. 24A, 24B y 24C. Los dos reflectores 310 y 320 adicionales pueden tener las mismas propiedades que el reflector 300 o tener propiedades diferentes de las propiedades del reflector 300. El reflector 310 puede tener las mismas propiedades que el reflector 320 o tener propiedades diferentes de las propiedades del reflector 320. El reflector 300 se conoce en lo sucesivo como el reflector 300 fijo. Un eje 510 está fijado a un borde del reflector 300 fijo (el borde opuesto al borde del reflector 300 fijo que está fijado a la placa 400 de soporte) y el reflector 310 está unido de manera giratoria al eje 510. Un eje 520 está fijado a un borde de la placa 400 de soporte (el borde opuesto al borde de la placa 400 de soporte en la que está fijado el reflector 300 fijo) y el reflector 320 está unido de manera giratoria al eje 520. Los reflectores 310 y 320 se conocerán en lo sucesivo como reflectores 310 y 320 móviles. Cuando el reflector 310 móvil está colocado de modo que la superficie reflectora del reflector 310 móvil está al ras con la superficie reflectora del reflector 300 fijo en la configuración ilustrada en la Fig. 25B, la combinación del reflector 300 fijo y del reflector 310 móvil funciona como un reflector que tiene una superficie reflectora mayor que el reflector 300 fijo. Además, en la configuración ejemplar ilustrada en la Fig. 25B, cuando los reflectores 310 y 320 móviles se establecen en posiciones adecuadas, un sonido se puede reflejar repetidamente en un espacio encerrado por la placa 400 de soporte y los reflectores 300 fijos, los reflectores 310 y 320 móviles como se representa en la Fig. 26, por ejemplo, por ello, se puede controlar el número Ξ de sonidos reflejados. Obsérvese que la placa 400 de soporte en la configuración ejemplar ilustrada en la Fig. 25B funciona como un objeto reflector y, por lo tanto, tiene preferiblemente las mismas propiedades que el objeto reflector descrito anteriormente.

Una configuración ejemplar de una implementación ilustrada en las Fig. 27A, 27B y 27C difiere de la configuración ejemplar ilustrada en las Fig. 24A, 24B y 24C en que una agrupación de micrófonos (una agrupación de micrófonos lineales en el ejemplo representado) también se proporciona en el reflector 300. Mientras que la dirección en la que están dispuestos los M micrófonos fijados a la placa 400 de soporte y la dirección en la que están dispuestos los M' micrófonos fijados al reflector 300 están en el mismo plano en la configuración ejemplar ilustrada en las Fig. 27A, 27B y 27C, los micrófonos no están limitados a esta disposición (M' = 13 en el ejemplo representado). Por ejemplo, los M' micrófonos pueden estar dispuestos y fijados al reflector 300 en la dirección ortogonal a la dirección en la que los M micrófonos están dispuestos y fijados a la placa 400 de soporte. En la configuración ejemplar ilustrada en las Fig. 27A, 27B y 27C, la combinación de la agrupación de micrófonos proporcionada en la placa 400 de soporte y el reflector 300 (el reflector 300 se usa como un objeto reflector sin usar la agrupación de micrófonos proporcionada en el reflector 300) se pueden usar para implementar la técnica de realce de sonido de la presente invención o la combinación de la placa 400 de soporte (la placa 400 de soporte se usa como objeto reflector sin usar la agrupación de micrófonos provista en la placa 400 de soporte) y la agrupación de micrófonos provista en el reflector 300 para implementar la técnica de realce de sonido de la presente invención.

En una configuración ejemplar extendida ilustrada en las Fig. 27A, 27B y 27C, se pueden añadir dos reflectores 310 y 320 adicionales a la configuración ejemplar ilustrada en las Fig. 27A, 27B y 27C como en la configuración ejemplar ilustrada en la Fig. 25B (véase la Fig. 28). Aunque no se representa, se puede proporcionar una agrupación de micrófonos en al menos uno de los reflectores 310 y 320 móviles. El orificio de captura de sonido de cada uno de los micrófonos de la agrupación de micrófonos proporcionada en el reflector 310 móvil se puede colocar en una superficie (la superficie de abertura) del reflector 310 móvil que es opuesta a la superficie de abertura de la placa 400 de soporte, por ejemplo. El orificio de captura de sonido de cada uno de los micrófonos de la agrupación de micrófonos proporcionada en el reflector 320 móvil se puede colocar en una superficie plana (la superficie de abertura) que puede formar el mismo plano que la superficie de abertura de la placa 400 de soporte, por ejemplo. Esta configuración ejemplar se puede usar de la misma forma que la configuración ejemplar ilustrada en la Fig. 25B. Además, en esta configuración ejemplar, cuando el reflector 320 móvil está colocado de modo que la superficie de abertura del reflector 320 móvil está al ras con la superficie de abertura de la placa 400 de soporte, la combinación de la placa 400 de soporte y el reflector 320 móvil funcionan como una agrupación de micrófonos más grande que la agrupación de micrófonos proporcionada en la placa 400 de soporte. Tanto la configuración ejemplar ilustrada en la Fig. 28 como la configuración ejemplar en la que se proporciona una agrupación de micrófonos al menos uno de los reflectores 310 y 320 móviles se puede usar de la misma forma que la configuración ejemplar ilustrada en la Fig. 26. En ambas de la configuración ejemplar ilustrada en la Fig. 28 y la configuración ejemplar en la que se proporciona una agrupación de micrófonos en al menos uno de los reflectores 310 y 320 móviles, los reflectores 310 y 320 móviles se pueden usar como objetos reflectores ordinarios y la agrupación de micrófonos proporcionada en la placa 400 de soporte y la agrupación de micrófonos proporcionada en el reflector 300 fijo se pueden usar como una agrupación combinada de micrófonos. Esto es equivalente a una configuración ejemplar que usa una agrupación de micrófonos compuesta por (M + M') micrófonos y dos objetos reflectores.

Si se proporciona una agrupación de micrófonos en el reflector 310 móvil, la agrupación de micrófonos se puede colocar en el reflector 310 móvil de modo que el orificio de captura de sonido de cada uno de los micrófonos de la agrupación de micrófonos proporcionada en el reflector 310 móvil está colocado en la superficie plana (la superficie de abertura) opuesta a la superficie plana del reflector 310 móvil que es opuesta a la superficie de abertura de la placa 400 de soporte. Si se proporciona una agrupación de micrófonos en el reflector 320 móvil, la agrupación de

micrófonos se puede colocar en el reflector 320 móvil de modo que el orificio de captura de sonido de cada uno de los micrófonos de la agrupación de micrófonos proporcionada en el reflector 320 móvil está colocado en la superficie plana (la superficie de abertura) opuesta a la superficie plana del reflector 320 móvil que puede formar el mismo plano que la superficie de abertura de la placa 400 de soporte. Por supuesto, se puede proporcionar una agrupación de micrófonos en al menos uno de los reflectores 310 y 320 móviles de modo que ambas superficies del reflector 310 y/o 320 móvil sean superficies de abertura.

[A] Si se proporciona una agrupación de micrófonos en al menos uno de los reflectores 310 y 320 móviles y, además, la superficie de abertura del reflector 310 móvil es una superficie plana opuesta a la superficie de abertura de la placa 400 de soporte o la superficie de abertura del reflector 320 móvil es una superficie plana que puede formar el mismo plano que la superficie de abertura de la placa 400 de soporte, colocar el reflector 310 móvil y/o el reflector 320 móvil de tal manera que la superficie de abertura del reflector 310 móvil y/o el reflector 320 móvil es invisible desde la dirección de la vista en la forma ilustrada en las Fig. 24A, 24B y 24C puede proporcionar el mismo efecto que aumentar el tamaño de la agrupación a través del uso de la agrupación de micrófonos proporcionada en el reflector 310 móvil y/o el reflector 320 móvil, aunque el tamaño de la agrupación aparente como se ve desde la dirección de la vista disminuye.

[B] Si se proporciona una agrupación de micrófonos en al menos uno de los reflectores 310 y 320 móviles y, además, la superficie de abertura del reflector 310 móvil es una superficie plana opuesta a la superficie opuesta a la superficie de abertura de la placa 400 de soporte o la superficie de abertura del reflector 320 móvil es una superficie plana opuesta a la superficie que puede formar el mismo plano que la superficie de abertura de la placa 400 de soporte, se puede proporcionar el mismo efecto que aumentar el tamaño de la agrupación en la forma ilustrada en las Fig. 24A, 24B y 24C, mientras que el tamaño de la agrupación aparente como se ve desde la dirección de la vista se mantiene igual.

Proporcionar una agrupación de micrófonos en ambas superficies de al menos uno de los reflectores 310 y 320 móviles de modo que ambas superficies del reflector 310 y/o 320 móvil sean superficies de abertura, puede proporcionar los mismos efectos que ambos de [A] y [B].

<Referencias>

(Referencia 1) Simon Haykin, "Adaptive Filter Theory", traducido por Hiroshi Suzuki et. al, primera edición, Kagaku Gijutsu Shuppan, 2001, páginas 66 - 73, 248 - 255

(Referencia 2) Nobuyoshi Kikuma, "Adaptive Antenna Technology", Primera edición, Ohmsha, 2003, páginas 35-90, ISBN4-27403611-1

(Referencia 3) Futoshi Asano, "Array signal processing – sound source localization/tracking and separation", editado por la Sociedad Acústica de Japón, serie de tecnología acústica 16, primera edición, Corona Publishing, páginas 88-89, 259 -261, ISBN978-4-339-01116-6

(Referencia 4) Yutaka Kaneda, "Directivity characteristics of adaptative microphone-array for noise reduction (AMNOR)", El Diario de la Sociedad Acústica de Japón, Vol. 44, Nº 1, 1988, páginas 23 -30

<Configuración de hardware ejemplar de aparato de realce de sonido>

Un aparato de realce de sonido relacionado con las realizaciones descritas anteriormente incluye una sección de entrada a la cual se puede conectar un teclado y similares, una sección de salida a la que se puede conectar un visualizador de cristal líquido y similares, una CPU (Unidad de Procesamiento Central) (que puede incluir una memoria como memoria caché), memorias tales como una RAM (Memoria de Acceso Aleatorio) y una ROM (Memoria de Sólo Lectura), un almacenamiento externo, que es un disco duro, y un bus que interconecta la sección de entrada, la sección de salida, la CPU, la RAM, la ROM y el almacenamiento externo de tal manera que puedan intercambiar datos. Un dispositivo (unidad) capaz de leer y escribir datos en un medio de grabación tal como un CD-ROM se puede proporcionar en el aparato de realce de sonido según sea necesario. Una entidad física que incluye estos recursos de hardware puede ser un ordenador de propósito general.

Los programas para realizar sonidos en un intervalo estrecho y los datos requeridos para el procesamiento por los programas se almacenan en el almacenamiento externo del aparato de realce de sonido (el almacenamiento no está limitado a un almacenamiento externo; por ejemplo, los programas se pueden almacenar en un dispositivo de almacenamiento de sólo lectura tal como una ROM). Los datos obtenidos a través del procesamiento de los programas se almacenan en la RAM o en el dispositivo de almacenamiento externo según sea adecuado. Un dispositivo de almacenamiento que almacena datos y direcciones de sus ubicaciones de almacenamiento se conocerá en lo sucesivo simplemente como "almacenamiento".

El almacenamiento del aparato de realce de sonido almacena un programa para obtener un filtro para cada frecuencia usando una matriz de correlación espacial, un programa para convertir una señal analógica en una señal digital, un programa para generar tramas, un programa para transformar una señal digital en cada trama a una señal en el dominio de frecuencia en el dominio de frecuencia, un programa para aplicar un filtro correspondiente a una

dirección o posición que es un objetivo de realce de sonido a una señal en el dominio de frecuencia en cada frecuencia para obtener una señal de salida, y un programa para transformar la salida única a una señal en el dominio de tiempo.

5 En el aparato de realce de sonido, los programas almacenados en el almacenamiento y los datos requeridos para el procesamiento de los programas se cargan en la RAM según sea necesario y se interpretan y ejecutan o procesan por la CPU. Como resultado, la CPU implementa funciones dadas (la sección de diseño de trama, el convertidor AD, el generador de tramas, la sección de transformación en el dominio de frecuencia, la sección de aplicación de filtro y la sección de transformación en el dominio de tiempo) para implementar el realce de sonido.

<Apéndice>

10 Los procesos descritos en las realizaciones se pueden realizar no solamente en una secuencia temporal como está escrita o se pueden realizar en paralelo unos con otros o individualmente, dependiendo del flujo máximo de los aparatos que realizan los procesos o requerimientos.

15 Si las funciones de procesamiento de cualquiera de las entidades de hardware (aparato de realce de sonido) descritas en las realizaciones se implementan mediante un ordenador, el procesamiento de las funciones que las entidades de hardware deberían incluir se describe en los programas. El programa se ejecuta en el ordenador para implementar las funciones de procesamiento de la entidad de hardware en el ordenador.

20 Los programas que describen el procesamiento pueden estar grabados en un medio de grabación legible por ordenador. El medio de grabación legible por ordenador puede ser cualquier medio de grabación tal como un dispositivo de grabación magnética, un disco óptico, un medio de grabación magnetoóptico y una memoria de semiconductores. Específicamente, por ejemplo, un dispositivo de disco duro, un disco flexible o una cinta magnética se pueden usar como dispositivo de grabación magnética, un DVD (Disco Versátil Digital), un DVD-RAM (Memoria de Acceso Aleatorio), un CD-ROM (Memoria de Sólo Lectura en Disco Compacto) o un CD-R (Grabable)/RW (Reescribible) se pueden usar como disco óptico, MO (disco magnetoóptico) se puede usar como medio de grabación magnetoóptico, y una EEPROM (Memoria de Sólo Lectura Borrable y Programable Electrónicamente) se puede usar como memoria de semiconductores.

25 El programa se distribuye vendiendo, transfiriendo o prestando un medio de grabación portátil en el que está grabado el programa, tal como un DVD o un CD-ROM. El programa puede estar almacenado en un dispositivo de almacenamiento de un ordenador servidor y ser transferido desde el ordenador servidor a otros ordenadores sobre una red, distribuyendo por ello el programa.

30 Un ordenador que ejecuta el programa primero almacena el programa grabado en un medio de grabación portátil o se transfiere desde un ordenador servidor a un dispositivo de almacenamiento del ordenador. Cuando el ordenador ejecuta los procesos, el ordenador lee el programa almacenado en el medio de grabación del ordenador y ejecuta los procesos según el programa leído. En otro modo de ejecución del programa, el ordenador puede leer el programa directamente de un medio de grabación portátil y ejecutar los procesos según el programa o puede ejecutar los procesos según el programa cada vez que se transfiere el programa desde el ordenador servidor al ordenador. Alternativamente, los procesos se pueden ejecutar usando un servicio denominado ASP (Proveedor de Servicios de Aplicaciones) en el que el programa no se transfiere desde un ordenador servidor al ordenador, sino que las funciones de proceso se implementan mediante instrucciones para ejecutar el programa y la adquisición de los resultados de la ejecución. Obsérvese que el programa en este modo abarca información que se proporciona para el procesamiento por un ordenador electrónico y es equivalente al programa (tal como datos que no son comandos directos a un ordenador pero que tienen la naturaleza que define el procesamiento del ordenador).

35 40 Mientras que las entidades de hardware están configuradas haciendo que un ordenador ejecute un programa predeterminado en las realizaciones descritas anteriormente, en al menos algunos de los procesos se pueden implementar por hardware.

45

REIVINDICACIONES

1. Un método de realce de sonido de obtención de una señal de salida en el dominio de frecuencia en el que un sonido de una posición deseada determinada por una dirección y una distancia se realiza aplicando, para cada frecuencia, un filtro que realiza el sonido de la posición deseada a señales en el dominio de la frecuencia transformadas a partir de M sonidos captados, captados con M micrófonos (200-1, ..., 200-M), donde M es un número entero mayor o igual a dos, en donde cada una de las funciones de transferencia $a_{i,g}$ se obtiene mediante la suma de una función de transferencia de un sonido directo que viene de la posición determinada por la dirección i y la distancia g y llega directamente a los M micrófonos y una función de transferencia de uno o más sonidos reflejados cuyas descomposiciones debidas a la reflexión y diferencias de tiempo de llegada con respecto al sonido directo se corrigen, siendo el uno o más sonidos reflejados producidos por la reflexión del sonido directo fuera de un objeto reflector y llegando a los M micrófonos,
- comprendiendo el método:
- un paso de diseño de filtro para obtener uno o una pluralidad de filtros; y
- un paso (S26, S36) de aplicación de filtro de aplicación, para cada frecuencia, para una posición deseada que es un objetivo de un realce de sonido, de un filtro obtenido en el paso de diseño de filtro a señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con los M micrófonos para obtener una señal de salida en el dominio de frecuencia en el que se realiza un sonido de la posición deseada;
- en donde el paso (S21, S35) de diseño de filtro usa la función de transferencia $a_{i,g}$ de un sonido que viene de cada una de una o de una pluralidad de posiciones predeterminadas que se supone que son fuentes de sonido y llega a cada uno de los micrófonos para obtener, para cada frecuencia y para cada una o una pluralidad de posiciones predeterminadas, un filtro respectivo para la posición predeterminada respectiva como objetivo de un realce de sonido antes de captar los M sonidos captados con los M micrófonos (200-1, ..., 200-M), donde i denota una dirección y g denota una distancia para identificar cada una de las posiciones.
2. El método de realce de sonido según la reivindicación 1,
- en donde cada una de las funciones de transferencia $a_{i,g}$ se obtiene mediante medición en un entorno real.
3. El método de realce de sonido según la reivindicación 1 o 2,
- en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de los sonidos de posiciones distintas de la posición que es el objetivo del realce de sonido.
4. El método de realce de sonido según la reivindicación 1 o 2,
- en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de maximización de la relación señal a ruido de un sonido de la posición que es el objetivo del realce de sonido.
5. El método de realce de sonido según la reivindicación 1 o 2,
- en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de los sonidos de posiciones distintas de la una o la pluralidad de posiciones que se supone que son posiciones de fuente de sonido mientras que un coeficiente de filtro para uno de los M micrófonos se fija a un valor constante.
6. El método de realce de sonido según la reivindicación 1 o 2,
- en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de los sonidos de posiciones distintas de la posición que es el objetivo de realce de sonido y uno o más puntos de supresión cuyas direcciones en las que la ganancia de ruido se suprime en condiciones que (1) el filtro pasa sonidos en todas las bandas de frecuencia de la posición que es el objetivo de realce de sonido y que (2) el filtro suprime los sonidos en todas las bandas de frecuencia del uno o más puntos de supresión.
7. El método de realce de sonido según la reivindicación 1 o 2,
- en donde el paso de diseño de filtro normaliza una función de transferencia $a_{s,h}$ de un sonido de la posición en una dirección $i = s$ a una distancia $g = h$ que es el objetivo de realce de sonido para obtener el filtro W para cada frecuencia, como $W = a_{s,h} / (a_{s,h}^H a_{s,h})$ donde H representa la transposición hermitiana.
8. El método de realce de sonido según la reivindicación 1 o 2,

en donde el paso de diseño de filtro usa una matriz de correlación espacial representada por las funciones de transferencia $a_{i,g}$ correspondientes a las posiciones distintas de la posición que es el objetivo del realce de sonido para obtener el filtro para cada frecuencia.

9. El método de realce de sonido según la reivindicación 1 o 2,

5 en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de los sonidos de posiciones distintas de la posición que es el objetivo de realce de sonido bajo la condición de que el filtro reduce la descomposición de un sonido de la posición que es el objetivo del realce de sonido a una cantidad predeterminada o menos.

10. El método de realce de sonido según la reivindicación 1 o 2,

10 en donde el paso de diseño de filtro usa una matriz de correlación espacial representada por una señal en el dominio de frecuencia para obtener el filtro para cada frecuencia, siendo la señal en el dominio de frecuencia obtenida transformando una señal obtenida mediante observación con una agrupación de micrófonos a un dominio de frecuencia.

11. El método de realce de sonido según la reivindicación 1 o 2,

15 en donde el paso de diseño de filtro usa una matriz de correlación espacial representada por las funciones de transferencia $a_{i,g}$ correspondientes a cada posición incluida en una o una pluralidad de posiciones que se supone que son posiciones de fuente de sonido para obtener el filtro para cada frecuencia

12. El método de realce de sonido según cualquiera de las reivindicaciones 1 a 11,

en donde los M sonidos captados incluyen sonidos reflejados de uno o más objetos (300) reflectores colocados.

20 13. Un aparato de realce de sonido que obtiene una señal de salida en el dominio de frecuencia en el que un sonido de una posición deseada determinada por una dirección y una distancia se realza aplicando, para cada frecuencia, un filtro que realza el sonido de la posición deseada a las señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con M micrófonos (200-1, ..., 200-M), donde M es un número entero mayor o igual a dos, en donde cada una de las funciones de transferencia $a_{i,g}$ se obtiene mediante la suma de una función de transferencia de un sonido directo que viene de la posición determinada por la dirección i y la distancia g y llega directamente a los M micrófonos y una función de transferencia de uno o más sonidos reflejados cuyas descomposiciones debidas a reflexión y diferencias de tiempo de llegada con respecto al sonido directo se corrigen, siendo el uno o más sonidos reflejados producidos por la reflexión del sonido directo fuera de un objeto reflector y llegando a los M micrófonos,

30 comprendiendo el aparato:

una sección de diseño de filtro para obtener uno o una pluralidad de filtros; y

35 una sección (640) de aplicación de filtro aplicando, para cada frecuencia, para una posición deseada que es un objetivo de un realce de sonido, un filtro obtenido por la sección de diseño de filtro a señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con los M micrófonos para obtener una señal de salida en el dominio de frecuencia en la que se realza un sonido de la posición deseada;

40 en donde la sección (660, 661) de diseño de filtro usa la función de transferencia $a_{i,g}$ de un sonido que viene de cada una de una o una pluralidad de posiciones predeterminadas que se supone que son fuentes de sonido y llega a cada uno de los micrófonos para obtener, para cada frecuencia y para cada una de una o una pluralidad de posiciones predeterminadas, un filtro respectivo para la posición predeterminada respectiva como objetivo de un realce de sonido antes de la captación de los M sonidos captados con los M micrófonos (200-1, ..., 200-M), donde i denota una dirección y g denota una distancia para identificar cada una de las posiciones; y

el objeto reflector existe en un entorno o está comprendido por el aparato de realce de sonido.

14. El aparato de realce de sonido según la reivindicación 13, comprendiendo además uno o más objetos (300) reflectores que proporcionan cada uno de los sonidos reflejados a los M micrófonos.

45 15. Un método de realce de sonido de obtención de una señal de salida en el dominio de frecuencia en el que un sonido de una dirección deseada se realza aplicando, para cada frecuencia, un filtro que realza el sonido de la dirección deseada a señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con M micrófonos (200-1, ..., 200-M), donde M es un número entero mayor o igual a dos, en donde cada una de las funciones de transferencia a_{ϕ} se obtiene mediante la suma de una función de transferencia de un sonido directo que viene de la dirección Φ y llega directamente a los M micrófonos y una función de transferencia de uno o más sonidos reflejados cuyas descomposiciones debidas a la reflexión y a diferencias de tiempo de llegada con respecto al sonido directo se corrigen, siendo el uno o más sonidos reflejados producidos por reflexión del sonido directo fuera de un objeto reflector y que llega a los M micrófonos,

50

comprendiendo el método:

un paso de diseño de filtro para obtener uno o una pluralidad de filtros; y

5 un paso (S6, S16) de aplicación de filtro de aplicación, para cada frecuencia, para una dirección deseada que es un objetivo de un realce de sonido, de un filtro obtenido en el paso de diseño de filtro a señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con los M micrófonos para obtener una señal de salida en el dominio de frecuencia en la que se realiza un sonido de la dirección deseada;

10 en donde el paso (S1, S15) de diseño de filtro usa la función de transferencia a_ϕ de un sonido que viene de cada una de una o una pluralidad de direcciones predeterminadas Φ que se supone que son direcciones desde las cuales vienen los sonidos y llegan a cada uno de los micrófonos para obtener, para cada frecuencia y para cada una de una o una pluralidad de direcciones predeterminadas, un filtro respectivo para la dirección predeterminada respectiva como objetivo de un realce de sonido antes de la captación de los M sonidos captados con los M micrófonos (200-1, ..., 200-M).

16. El método de realce de sonido según la reivindicación 15,

en donde cada una de las funciones de transferencia a_ϕ se obtiene mediante medición en un entorno real.

15 17. El método de realce de sonido según la reivindicación 15 o 16,

en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de sonidos de direcciones distintas de la dirección que es el objetivo del realce de sonido.

18. El método de realce de sonido según la reivindicación 15 o 16,

20 en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de maximización de la relación señal a ruido de un sonido de la dirección que es el objetivo del realce de sonido.

19. El método de realce de sonido según la reivindicación 15 o 16,

25 en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de sonidos a partir de la una o la pluralidad de direcciones que se supone que son direcciones desde las cuales vienen los sonidos, mientras que un coeficiente de filtro para uno de los M micrófonos se fija a un valor constante.

20. El método de realce de sonido según la reivindicación 15 o 16,

30 en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de sonidos de las direcciones distintas de la dirección que es el objetivo de realce de sonido y una o más direcciones nulas en condiciones que (1) el filtro pasa sonidos en todas las bandas de frecuencia de la dirección que es el objetivo de realce de sonido y que (2) el filtro suprime sonidos en todas las bandas de frecuencia de la una o más direcciones nulas.

21. El método de realce de sonido según la reivindicación 15 o 16,

35 en donde el paso de diseño de filtro normaliza una función de transferencia a_s de un sonido de la posición en la dirección $\Phi = s$ que es el objetivo de realce de sonido para obtener el filtro W para cada frecuencia, como $W = a_s / (a_s^H a_s)$ donde H representa la transposición hermitiana.

22. El método de realce de sonido según la reivindicación 15 o 16,

40 en donde el paso de diseño de filtro usa una matriz de correlación espacial representada por las funciones de transferencia a_ϕ correspondientes a las direcciones distintas de las direcciones que son el objetivo del realce de sonido para obtener el filtro para cada frecuencia.

23. El método de realce de sonido según la reivindicación 15 o 16,

45 en donde el paso de diseño de filtro implica, para cada frecuencia, el diseño del filtro basado en el criterio de minimización de la potencia de los sonidos de direcciones distintas de la dirección que es el objetivo de realce de sonido bajo la condición de que el filtro reduzca la descomposición de un sonido de la dirección que es el objetivo del realce de sonido a una cantidad predeterminada o menos.

24. El método de realce de sonido según la reivindicación 15 o 16,

en donde el paso de diseño de filtro usa una matriz de correlación espacial representada por una señal en el dominio de frecuencia para obtener el filtro para cada frecuencia, siendo la señal en el dominio de frecuencia obtenida

transformando una señal obtenida mediante observación con una agrupación de micrófonos a un dominio de frecuencia.

25. El método de realce de sonido según la reivindicación 15 o 16,

en donde los M sonidos captados incluyen sonidos reflejados de uno o más objetos (300) reflectores colocados.

5 26. Un aparato de realce de sonido que obtiene una señal de salida en el dominio de frecuencia en la que un sonido de una dirección deseada se realza aplicando, para cada frecuencia, un filtro que realza el sonido de la dirección deseada a las señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con M micrófonos (200-1, ..., 200-M), donde M es un número entero mayor o igual a dos, en donde cada una de las funciones de transferencia a_{ϕ} se obtiene mediante la suma de una función de transferencia de un sonido directo que viene de la dirección Φ y llega directamente a los M micrófonos y una función de transferencia de uno o más sonidos reflejados cuyas descomposiciones debidas a la reflexión y a las diferencias de tiempo de llegada con respecto al sonido directo se corrigen, siendo el uno o más sonidos reflejados producidos por reflexión del sonido directo fuera de un objeto reflector y que llegan a los M micrófonos,

10 comprendiendo el método:

15 una sección de diseño de filtro para obtener uno o una pluralidad de filtros; y

una sección (240) de aplicación de filtro que aplica, para cada frecuencia, para una dirección deseada que es un objetivo de un realce de sonido, un filtro obtenido por la sección de diseño de filtro a señales en el dominio de frecuencia transformadas a partir de M sonidos captados, captados con los M micrófonos para obtener una señal de salida en el dominio de frecuencia en la que se realza un sonido de la dirección deseada;

20 en donde la sección (260, 261) de diseño de filtro usa la función de transferencia a_{ϕ} de un sonido que viene de cada una de una o una pluralidad de direcciones predeterminadas Φ que se supone que son direcciones desde las cuales vienen los sonidos y llegan a cada uno de los micrófonos para obtener, para cada frecuencia y para cada una de una o una pluralidad de direcciones predeterminadas, un filtro respectivo para la dirección predeterminada respectiva como objetivo de un realce de sonido antes de la captación de los M sonidos captados con los M micrófonos (200-1, ..., 200-M); y

25 el objeto reflector existe en un entorno o está comprendido por el aparato de realce de sonido.

27. El aparato de realce de sonido según la reivindicación 26 comprendiendo además uno o más objetos (300) reflectores que proporcionan cada uno de los sonidos reflejados a los M micrófonos.

30 28. Un programa de ordenador que hace que un ordenador ejecute el procesamiento del método de realce de sonido según la reivindicación 1 o 15.

29. Un medio de grabación legible por ordenador que tiene grabado en el mismo un programa de ordenador para hacer que un ordenador ejecute los pasos del método de realce de sonido según la reivindicación 1 o 15.

FIG. 1A

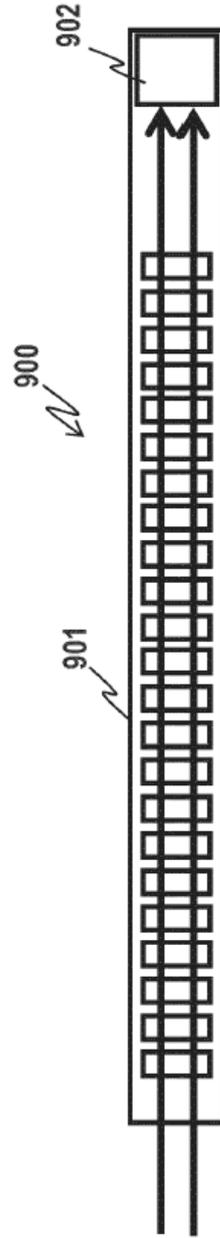


FIG. 1B

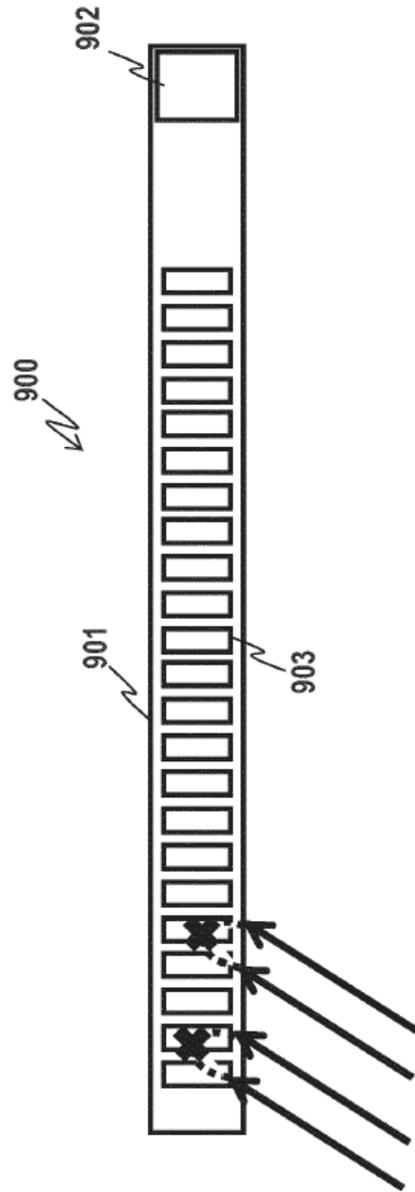


FIG. 2A

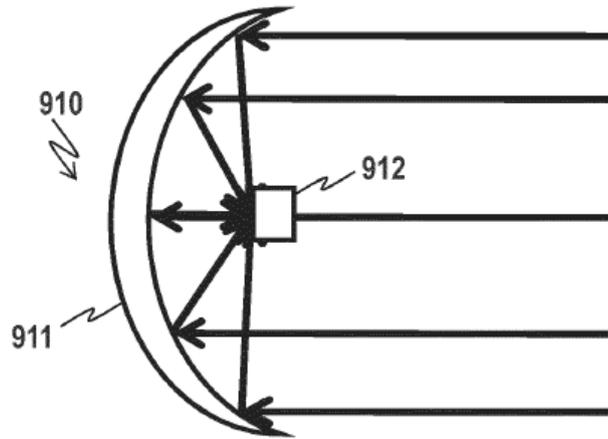
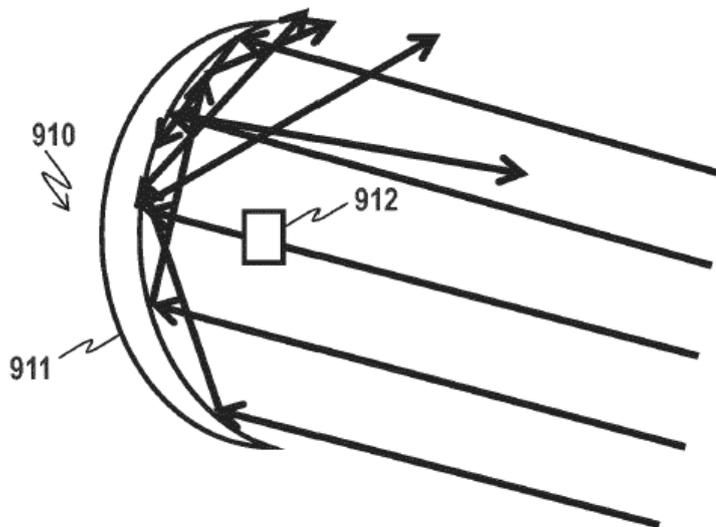
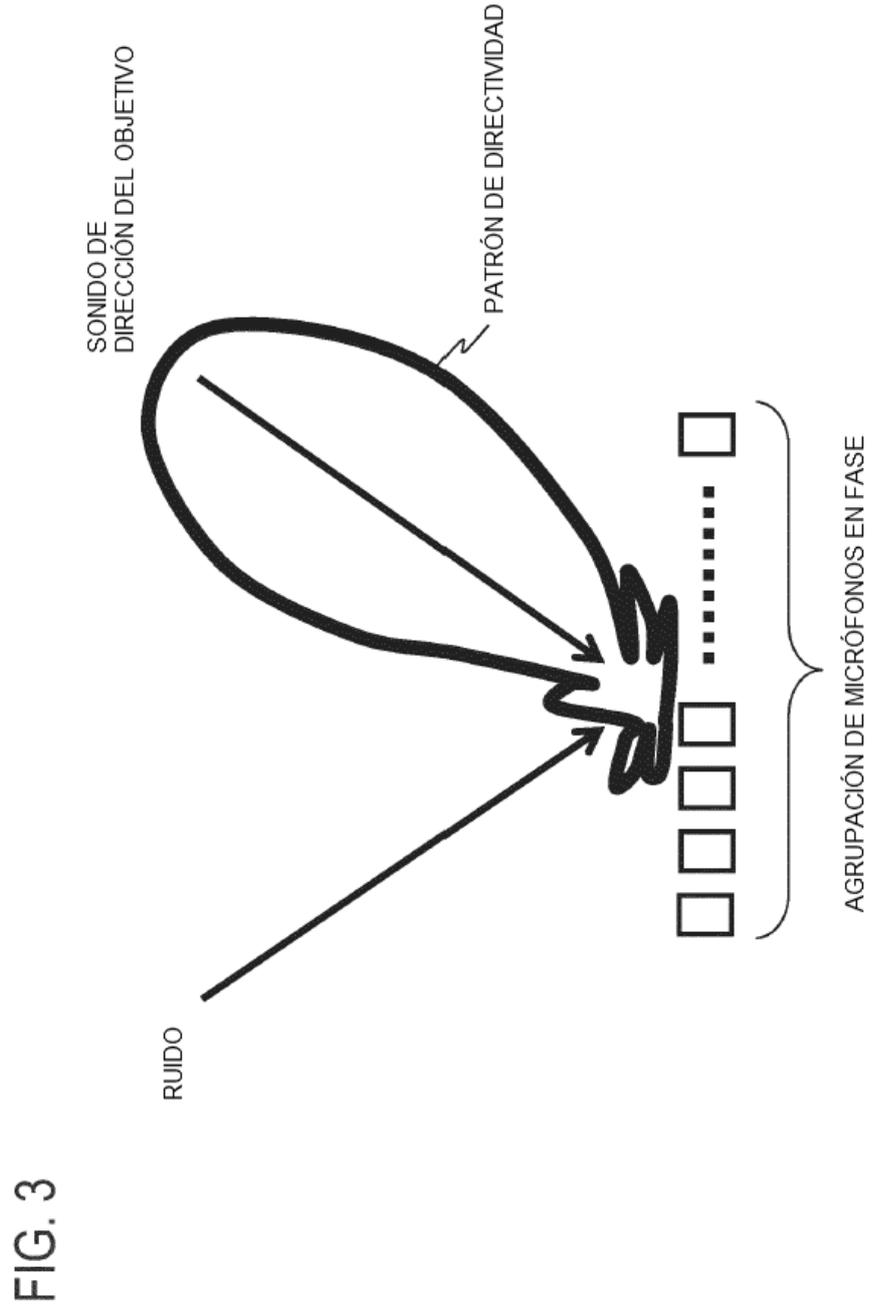


FIG. 2B





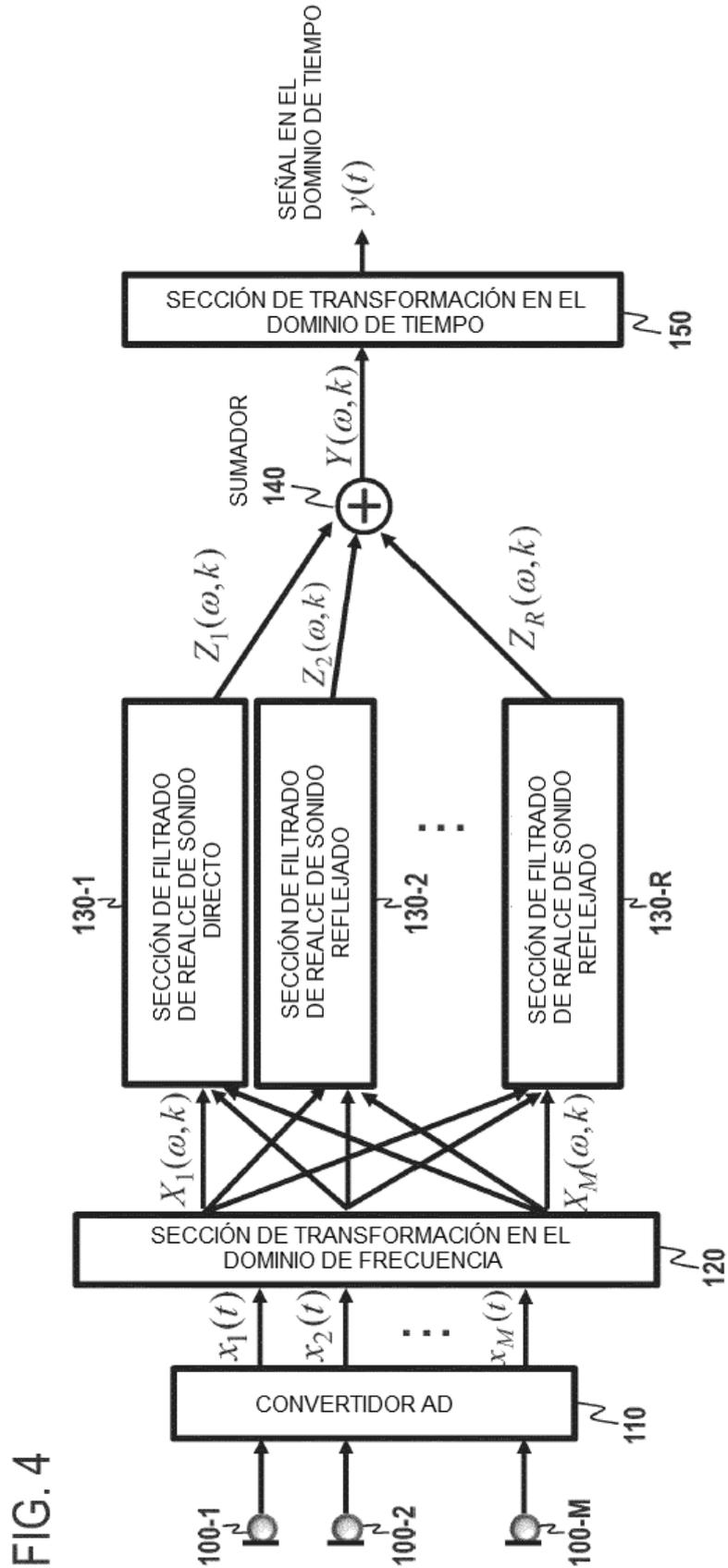


FIG. 4

FIG. 5A

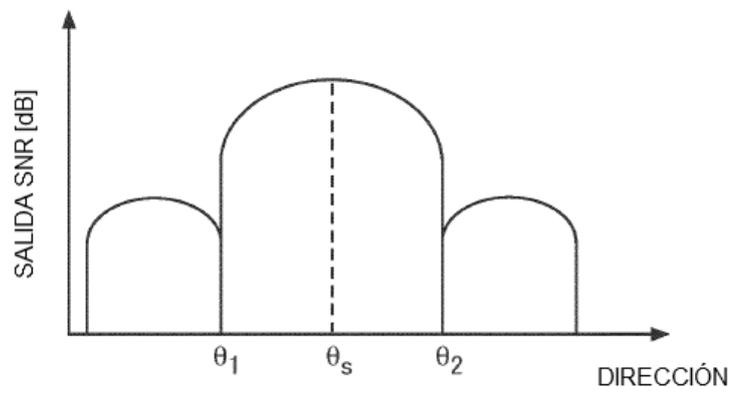


FIG. 5B

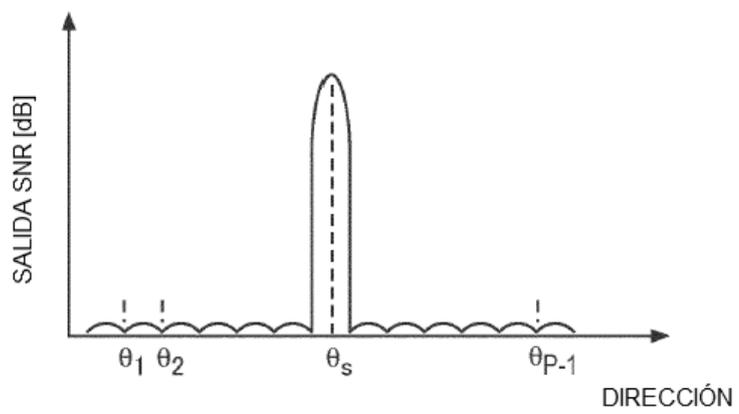
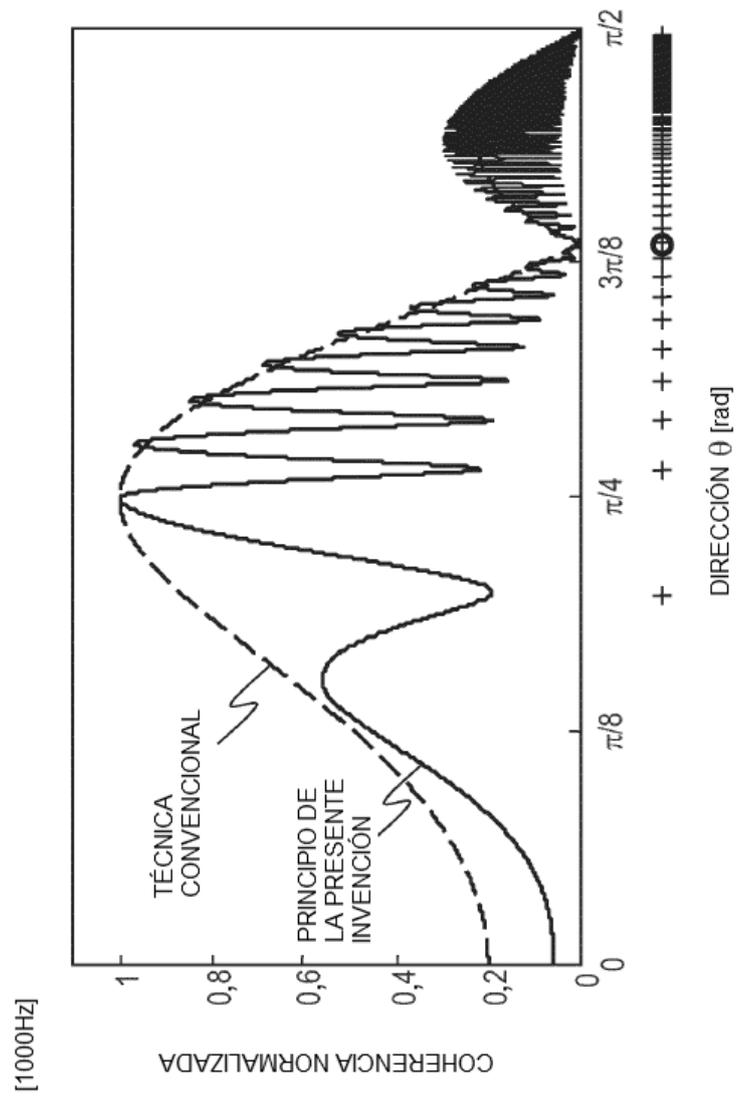


FIG. 6



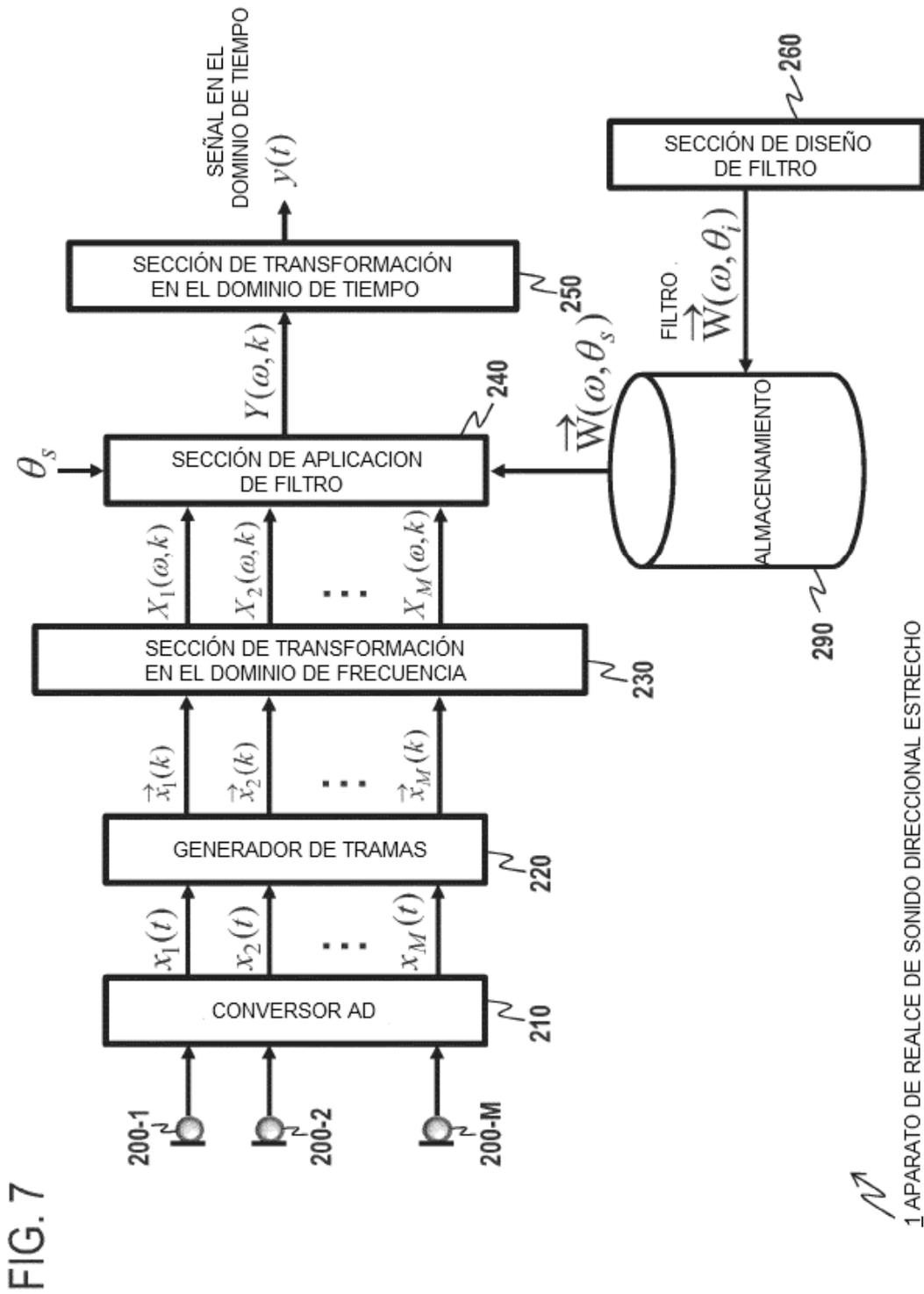


FIG. 8

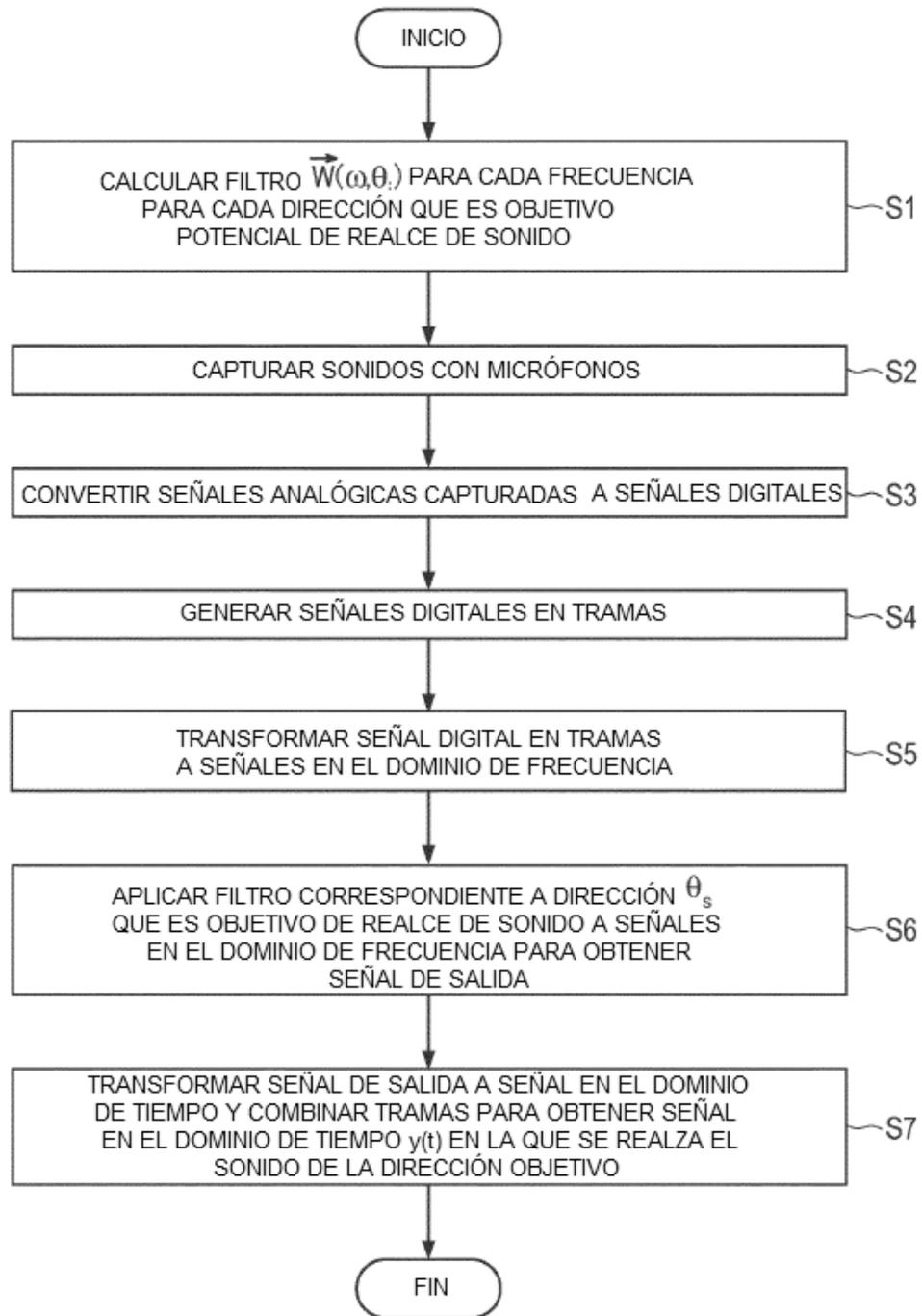
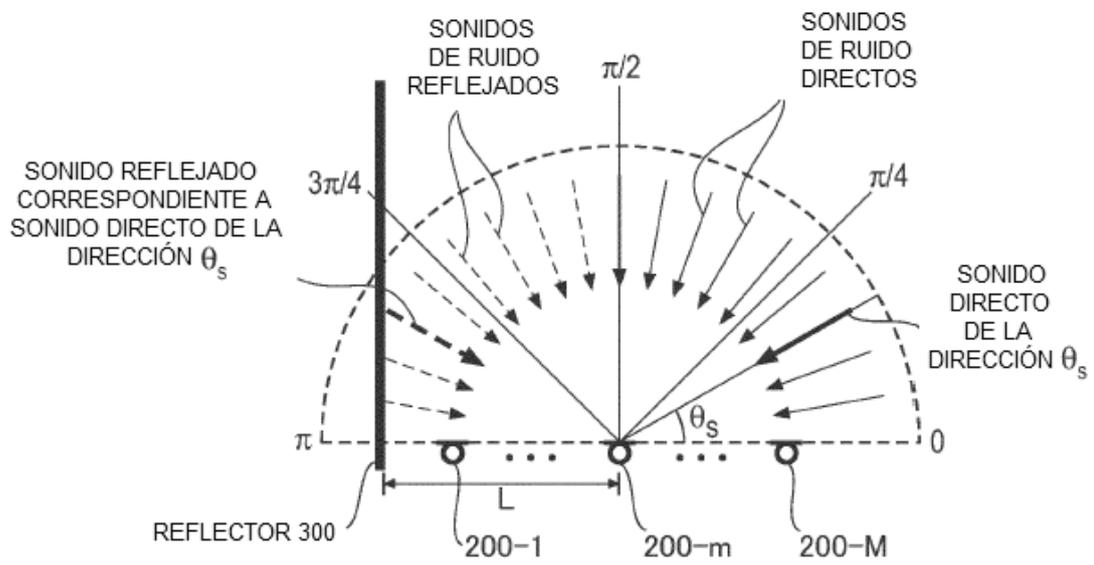


FIG. 9



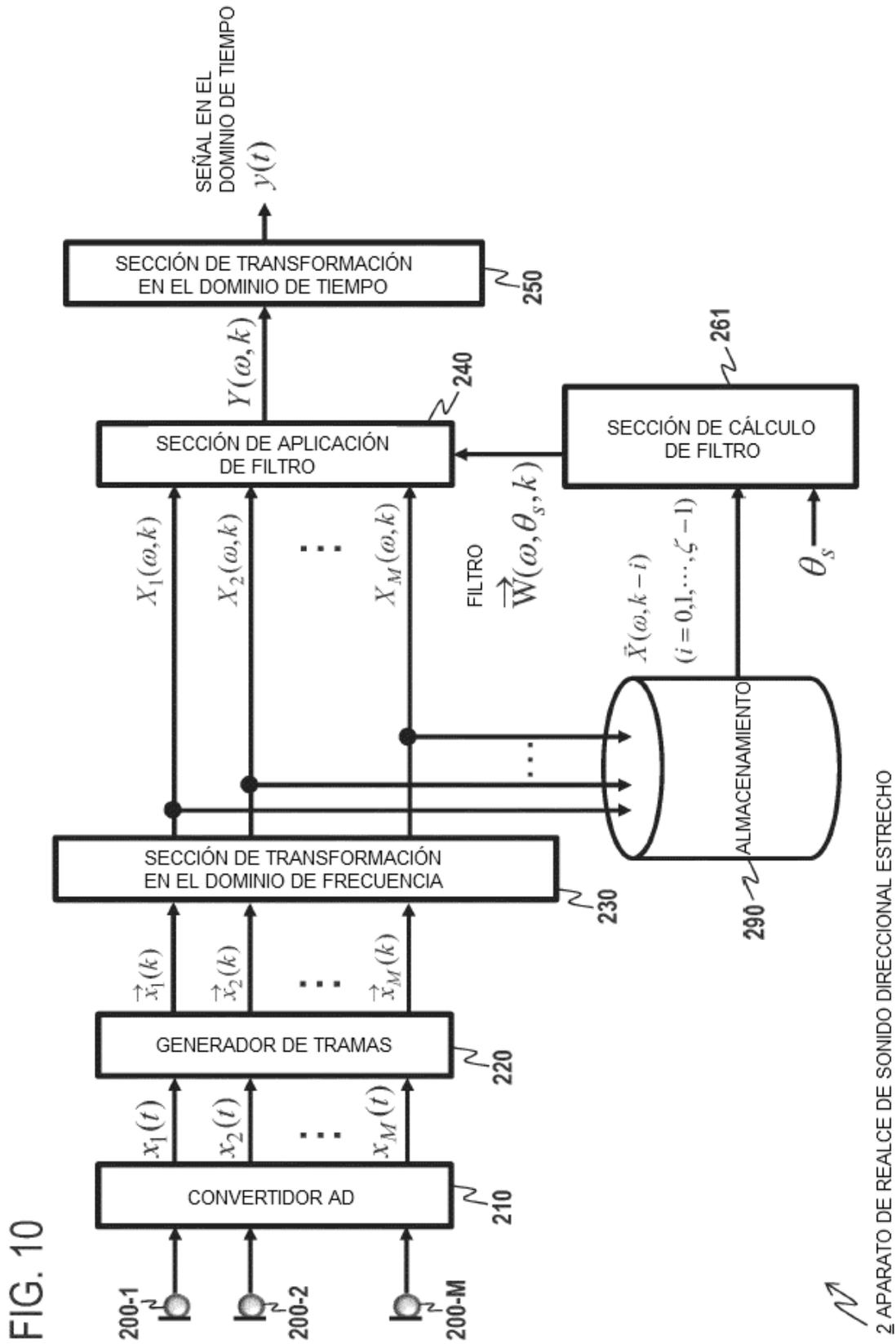


FIG. 10

FIG. 11

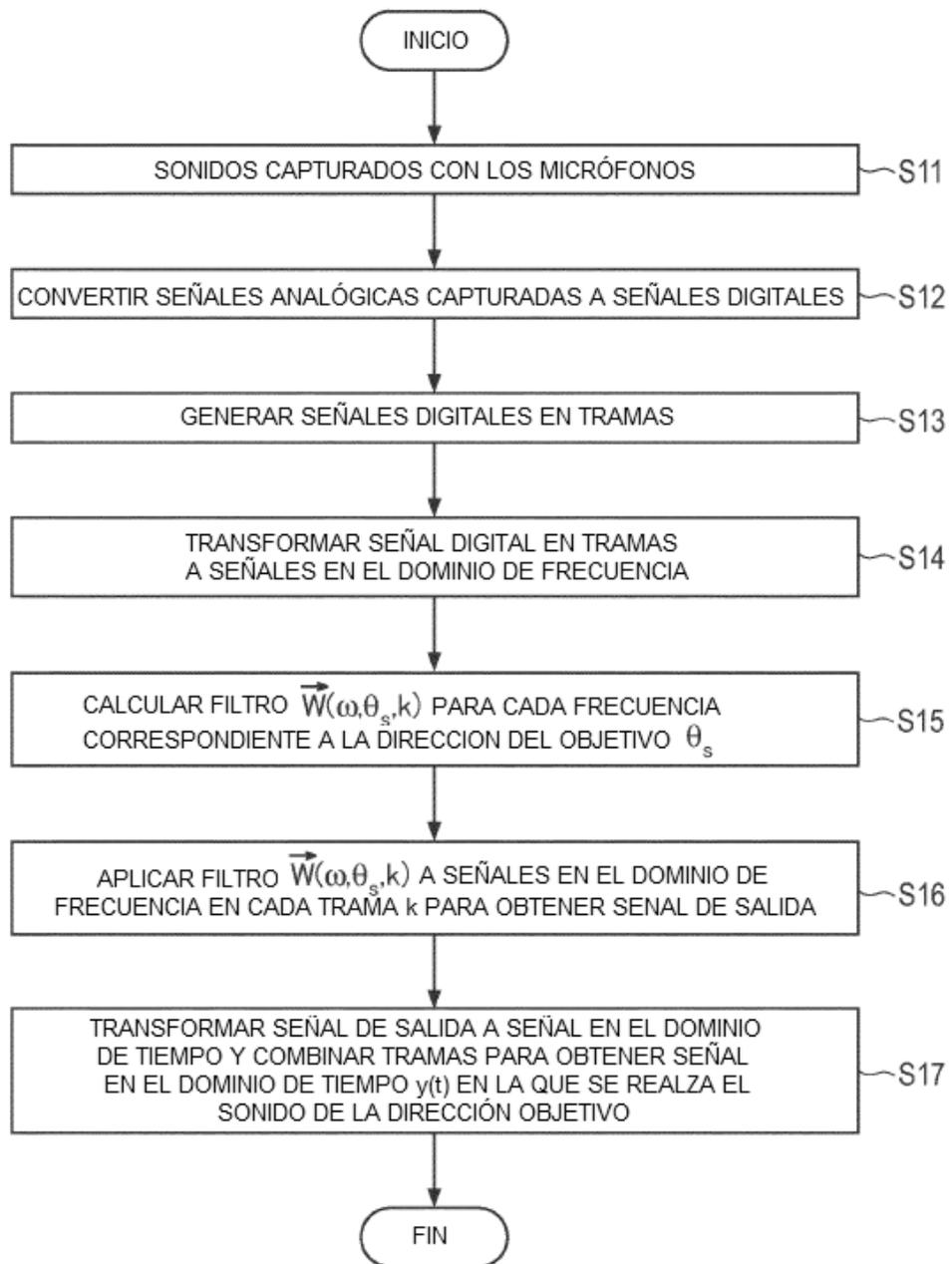


FIG. 12

MÉTODO CONVENCIONAL 1 - - - - - MÉTODO MVDR (SIN REFLECTOR)
 MÉTODO CONVENCIONAL 2 ······ MÉTODO DE CONFORMACIÓN DE HACES DE RETARDO Y SUMA (CON REFLECTOR)
 PRIMERA REALIZACIÓN ——— MÉTODO MVDR (CON REFLECTOR)

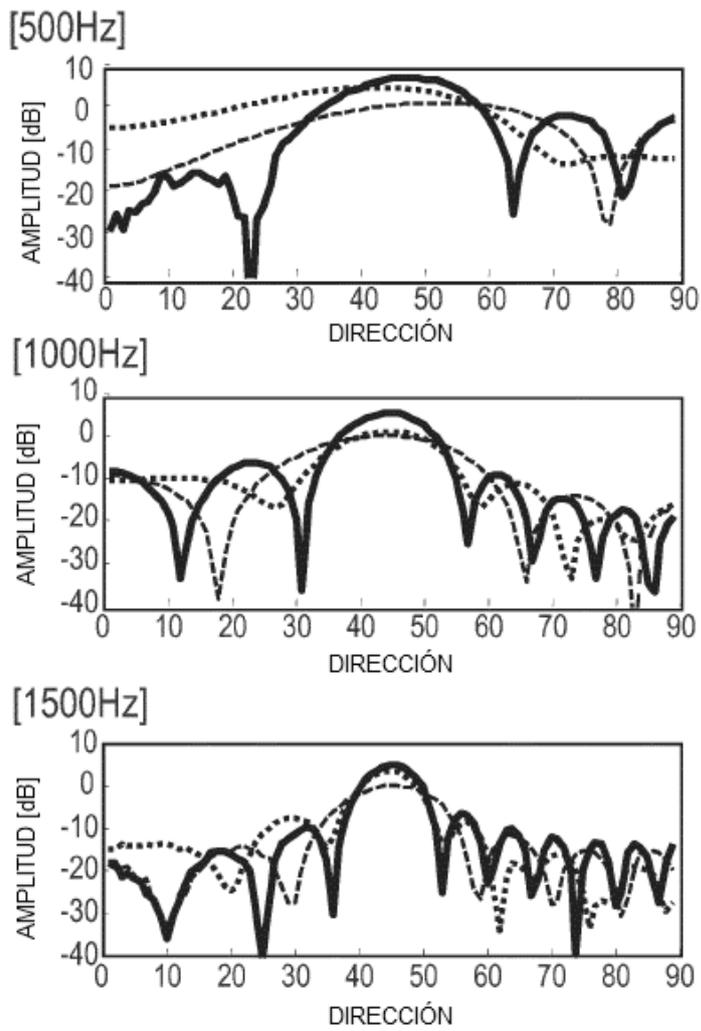


FIG. 13

MÉTODO CONVENCIONAL 1 - - - - - MÉTODO MVDR (SIN REFLECTOR)
 MÉTODO CONVENCIONAL 2 ······ MÉTODO DE CONFORMACIÓN DE HACES DE RETARDO Y SUMA (CON REFLECTOR)
 PRIMERA REALIZACIÓN ————— MÉTODO MVDR (CON REFLECTOR)

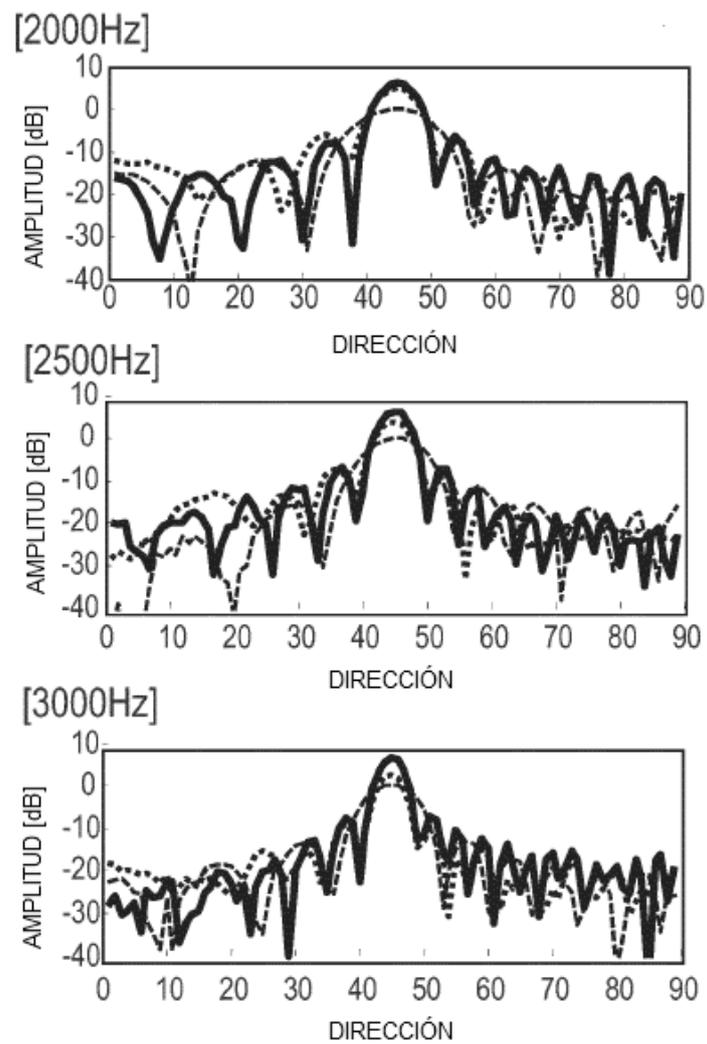


FIG. 14

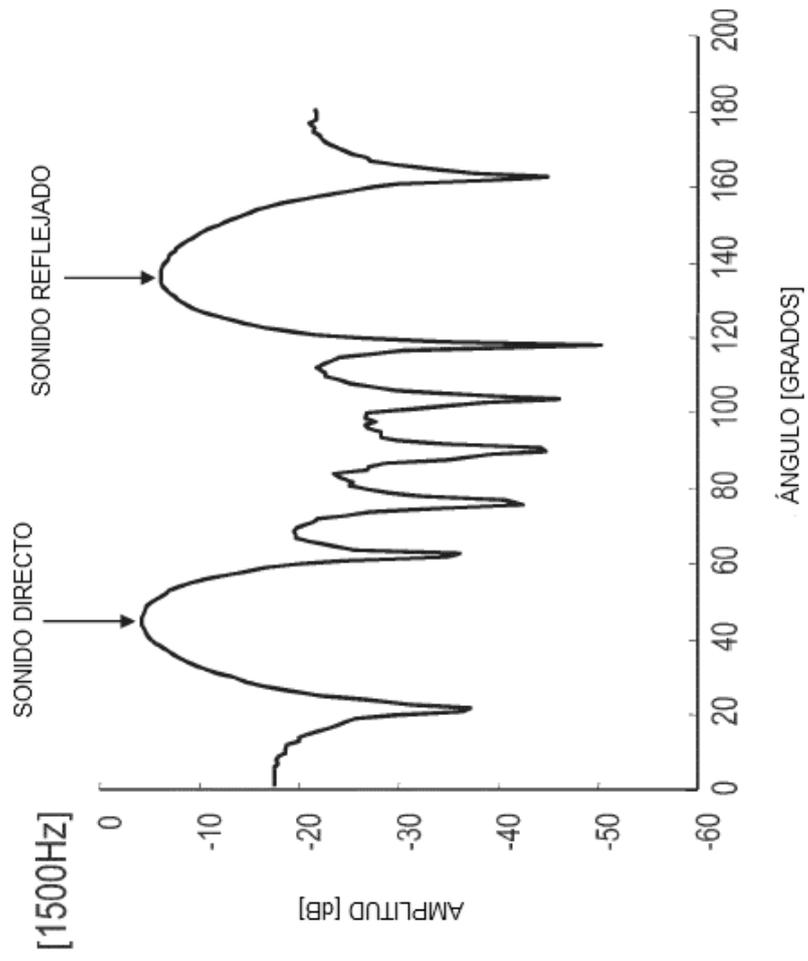


FIG. 15

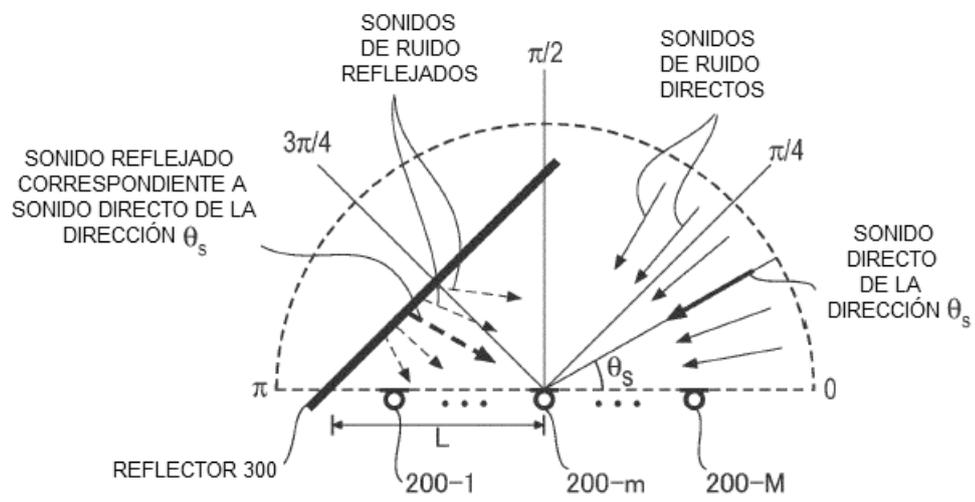


FIG. 16

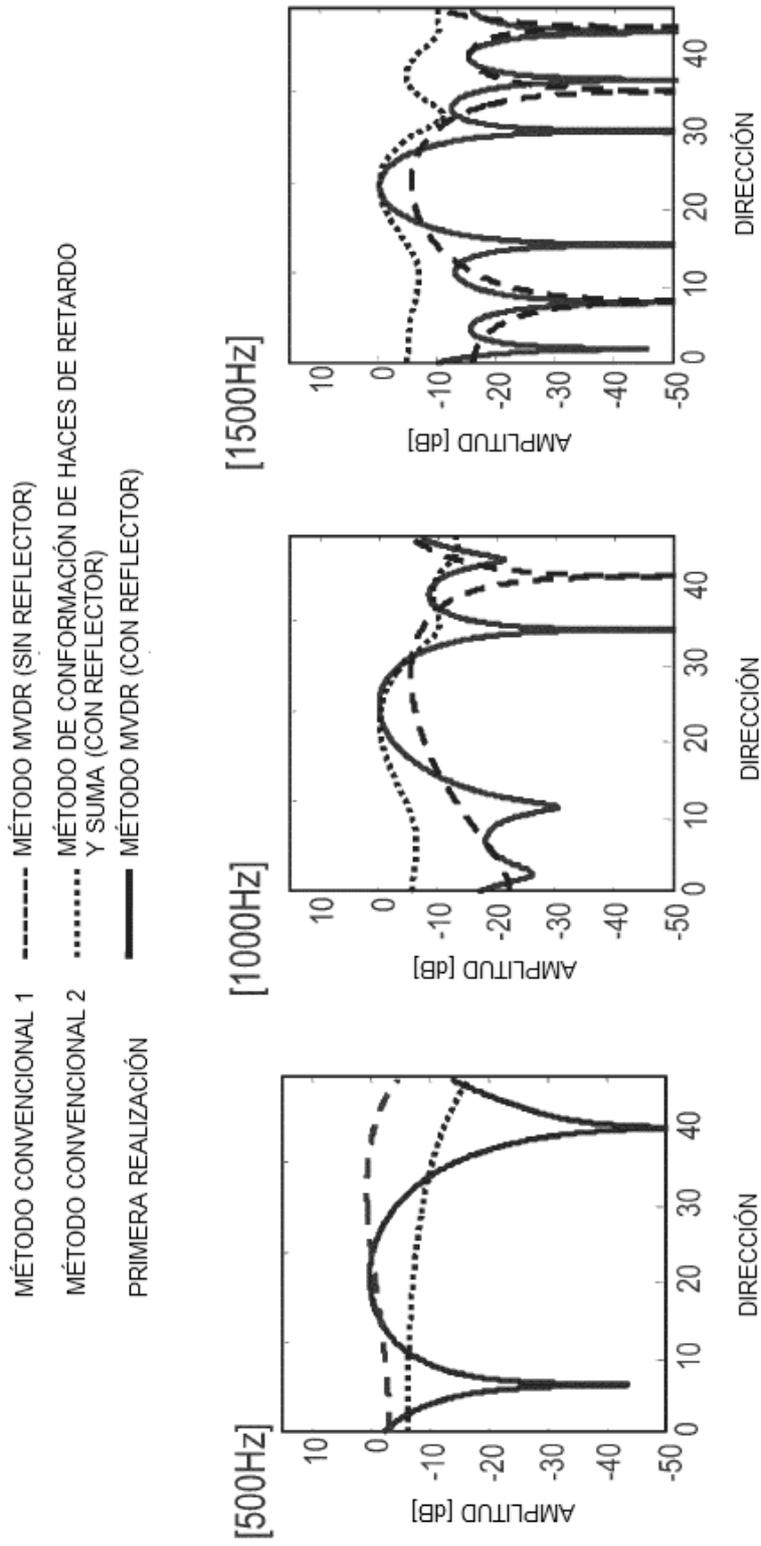


FIG. 17

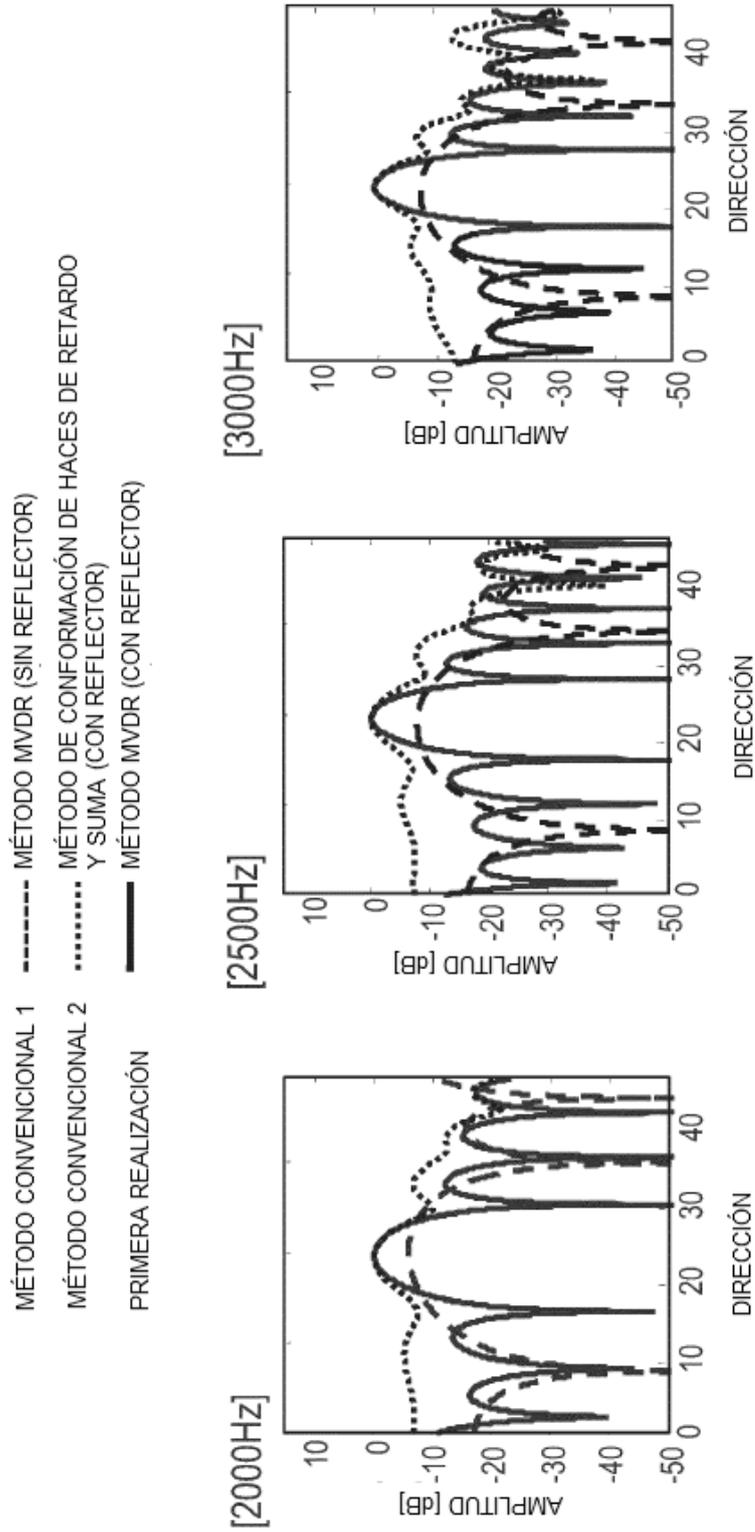


FIG. 18A

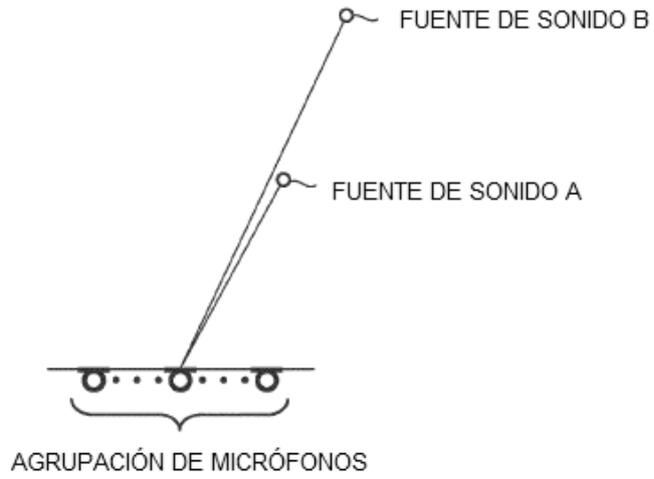
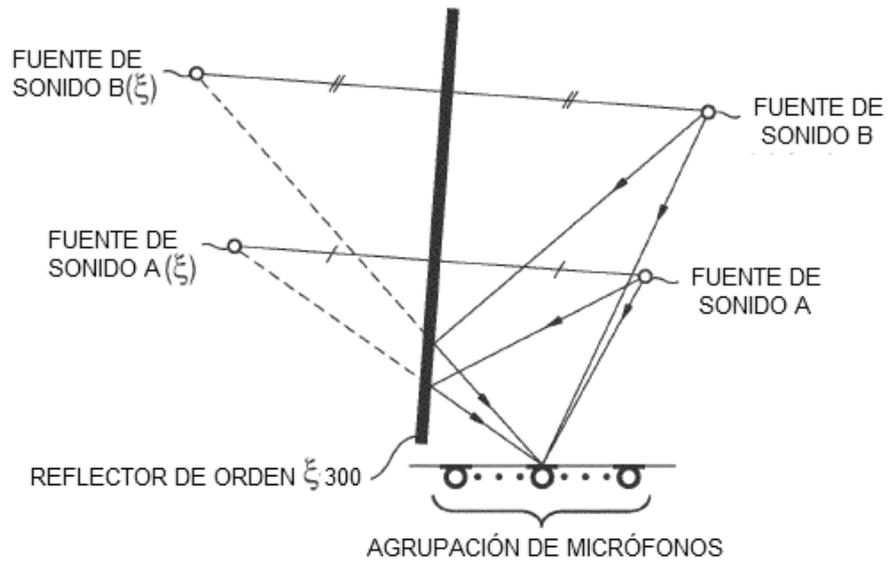


FIG. 18B



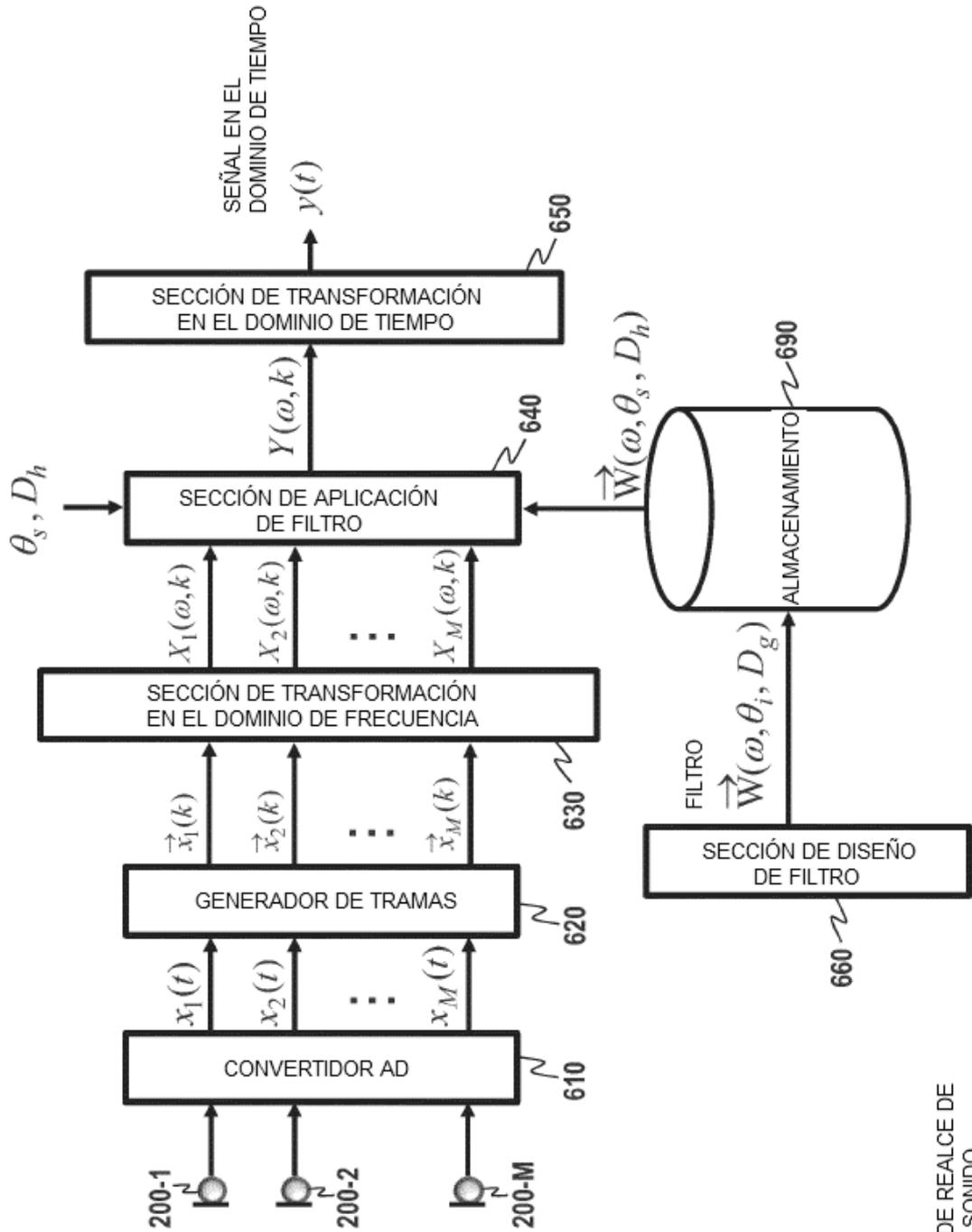
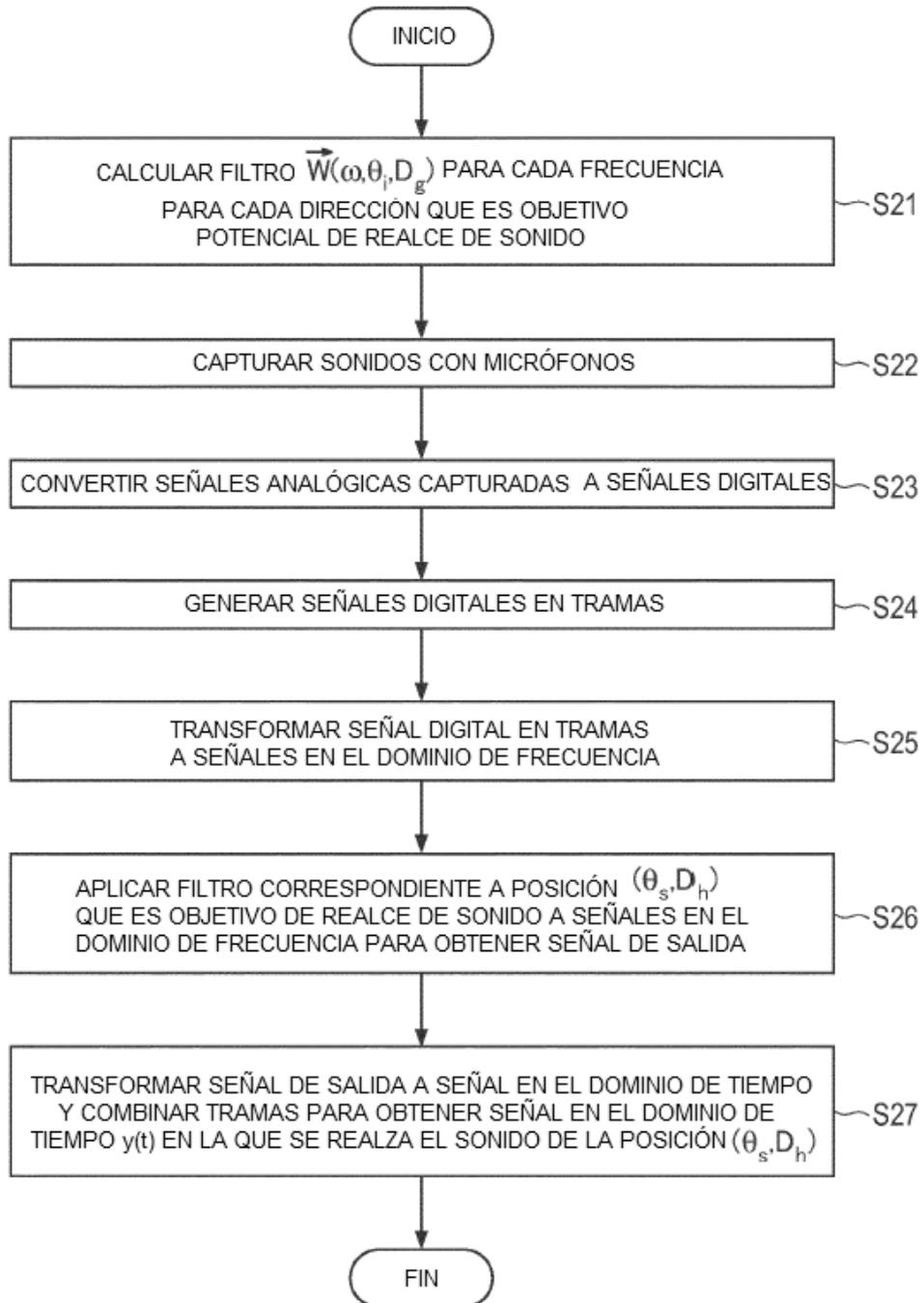


FIG. 19

3 APARATO DE REALCE DE PUNTO DE SONIDO

FIG. 20



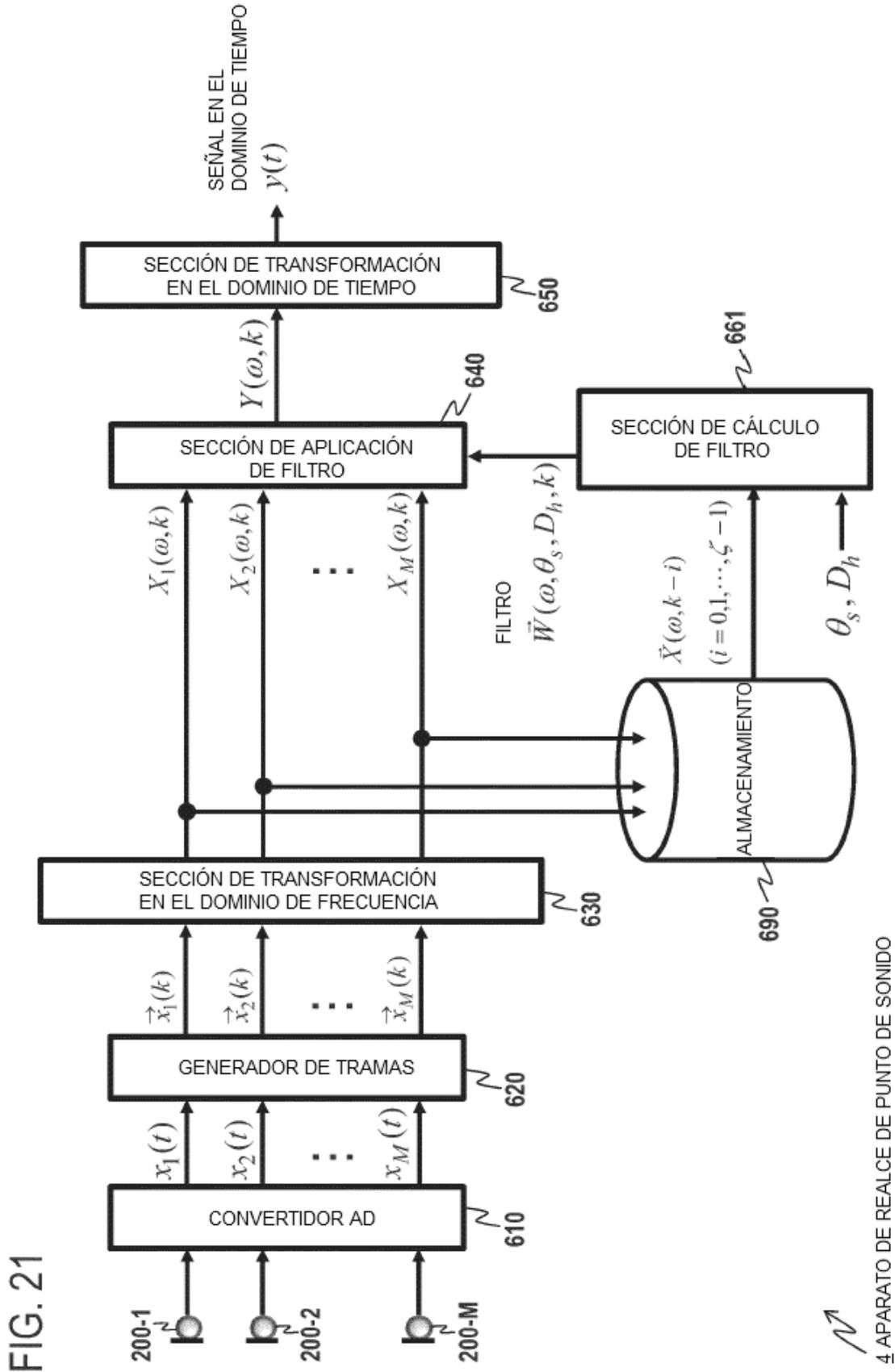


FIG. 21

FIG. 22

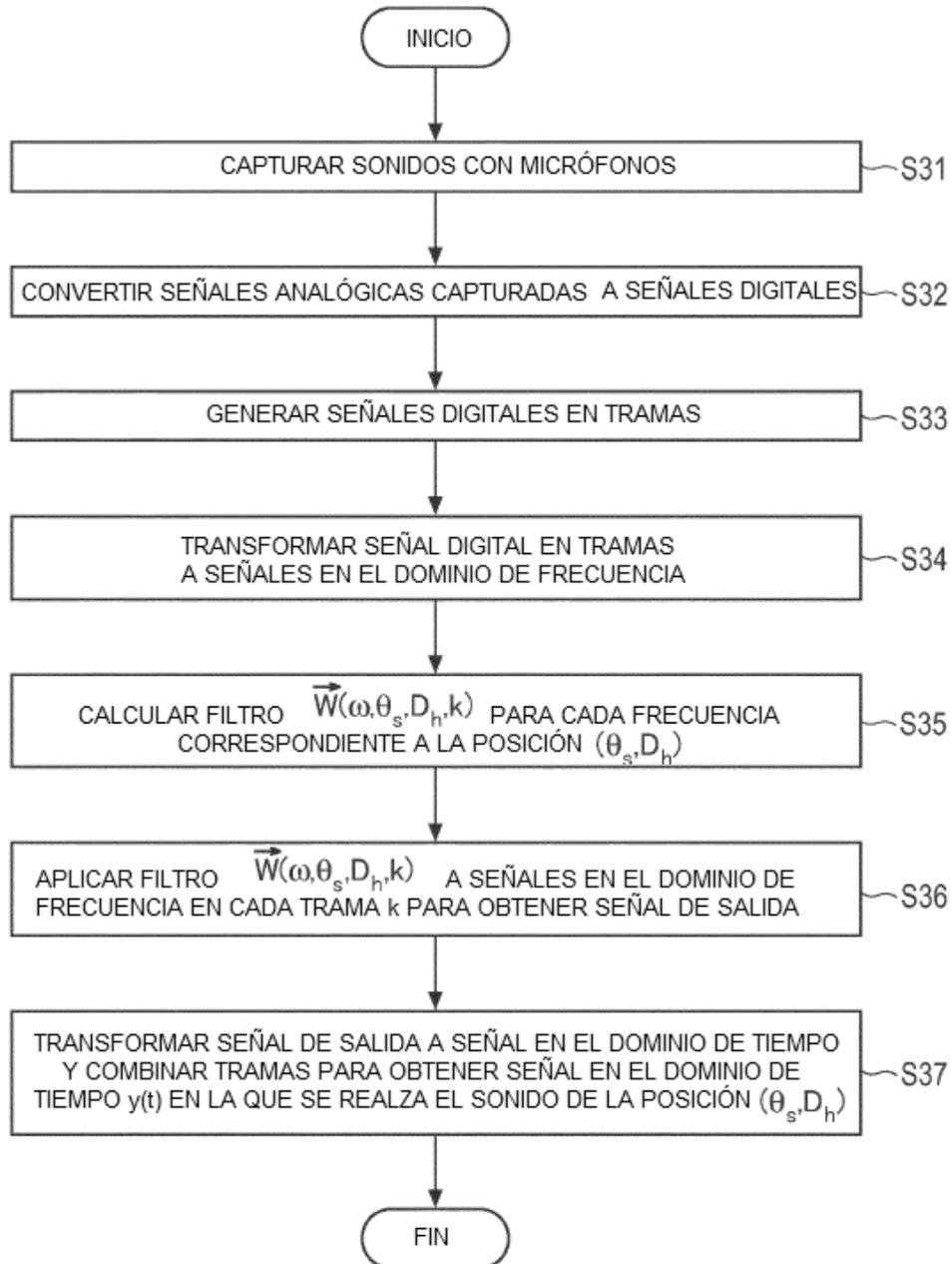


FIG. 23A

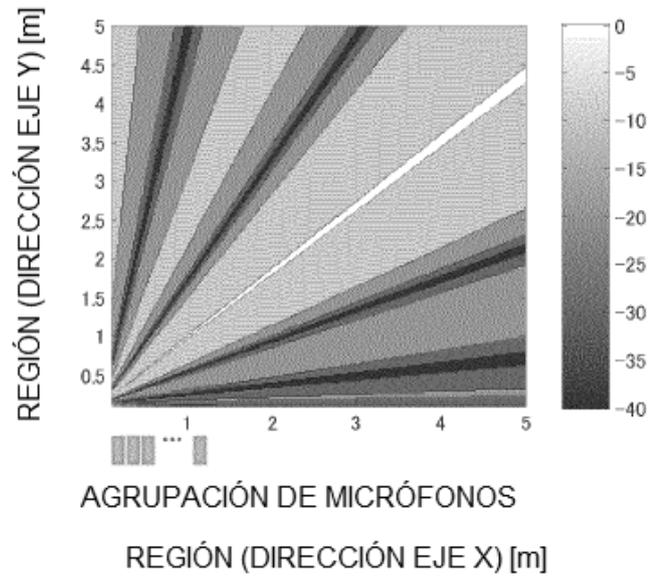


FIG. 23B

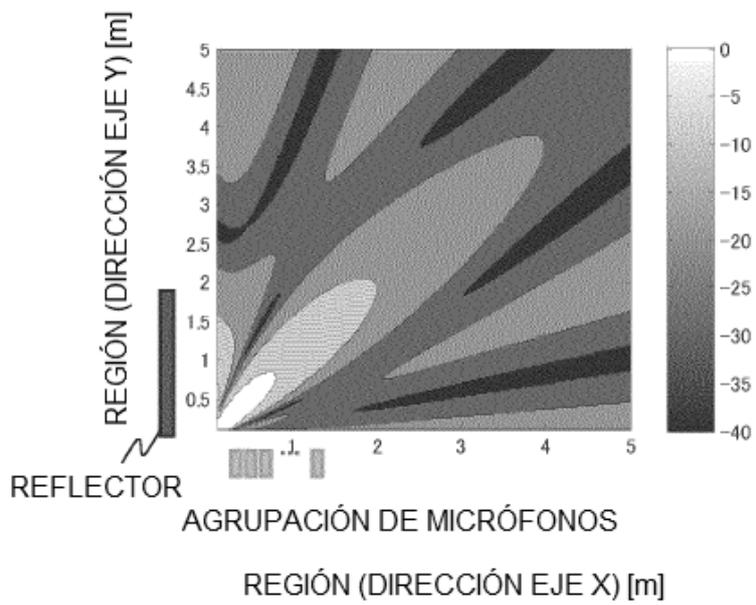


FIG. 24A

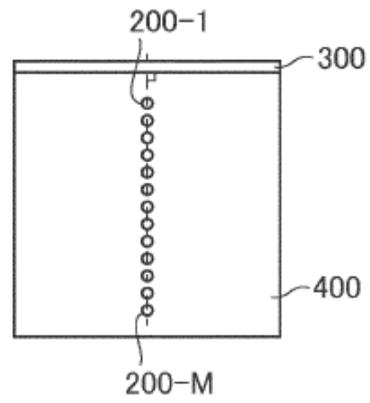


FIG. 24B

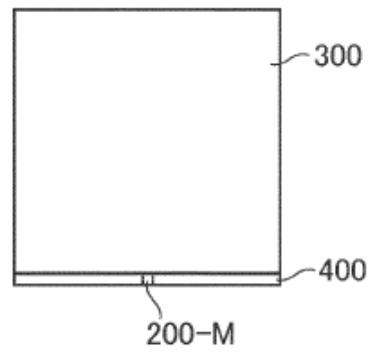


FIG. 24C

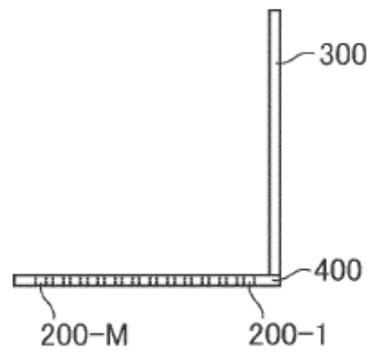


FIG. 25A

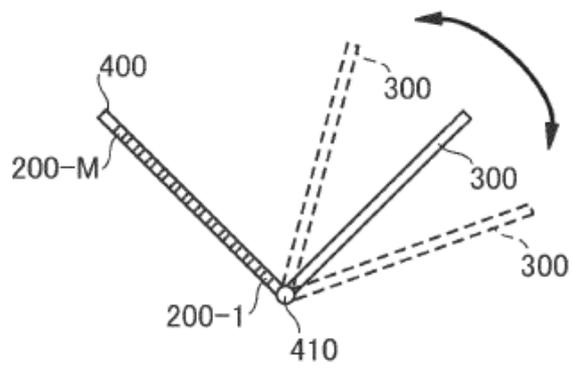


FIG. 25B

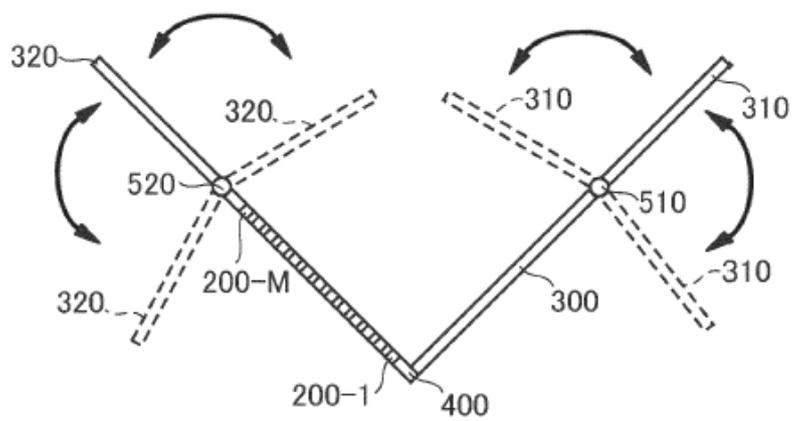


FIG. 26

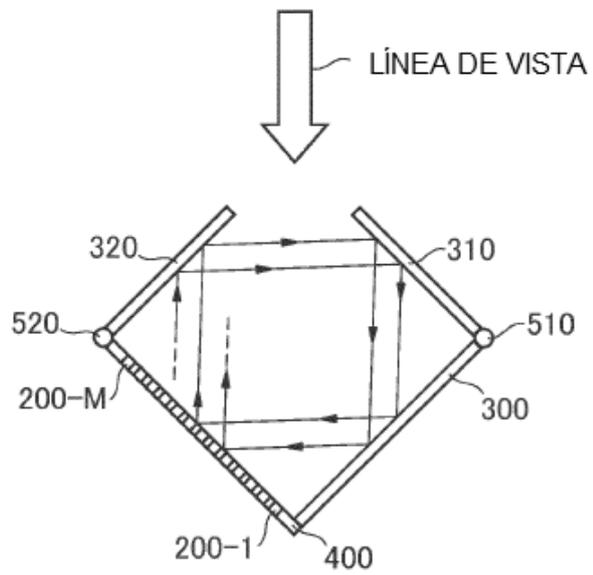


FIG. 27A

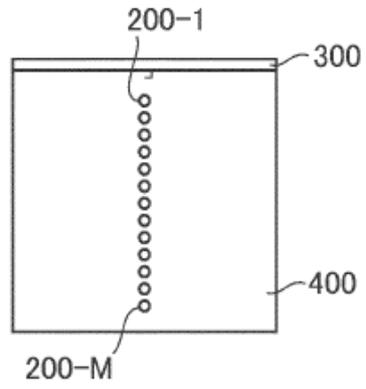


FIG. 27B

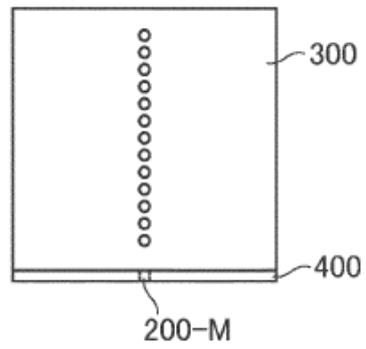


FIG. 27C

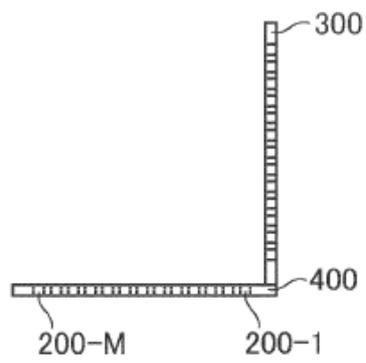


FIG. 28

