

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 430 251**

51 Int. Cl.:

H04R 3/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.09.2010 E 10765502 (9)**

97 Fecha y número de publicación de la concesión europea: **24.07.2013 EP 2478715**

54 Título: **Procedimiento para adquirir señales de audio y sistema de adquisición de audio del mismo**

30 Prioridad:

18.09.2009 IT TO20090713

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

19.11.2013

73 Titular/es:

**RAI RADIOTELEVISIONE ITALIANA S.P.A.
(100.0%)**

**Viale Mazzini 14
00195 Roma, IT**

72 Inventor/es:

**SCOPECE, LEONARDO y
FARINA, ANGELO**

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 430 251 T3

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

DESCRIPCIÓN

Procedimiento para adquirir señales de audio y sistema de adquisición de audio del mismo.

5 La presente invención se refiere a un procedimiento para adquirir señales de audio y un sistema de adquisición de audio que puede implementar dicho procedimiento.

10 En el ámbito del cine y la televisión y similares, hay una necesidad creciente de grabar sonidos con precisión en el entorno tridimensional donde tiene lugar el rodaje, de tal forma que estos se puedan reproducir fielmente en las dependencias del usuario.

La grabación de sonidos en un entorno tridimensional conlleva la necesidad de conocer la presión y la velocidad de las partículas de aire en un punto espacial determinado.

15 Para este fin, actualmente se conoce la forma de utilizar sondas microfónicas que comprenden varias cápsulas microfónicas dispuestas en una superficie (por ejemplo, una superficie esférica).

20 Un ejemplo de dichas sondas es la sonda microfónica comercializada con el nombre "*EigenMike32*" y fabricada por la empresa norteamericana "*mhAcoustics*". Se da a conocer una agrupación de micrófonos miniaturizados en el artículo "A miniaturized adaptive microphone array under directional constraint utilizing aggregated microphones" de Matsumoto y Hashimoto (The Journal of the Acoustical Society of America, vol. 119, nº 1, enero de 2006, páginas 352-359).

25 La figura 1 representa un ejemplo de sonda 11 que permite adquirir señales de audio de varias direcciones espaciales. Dicha sonda 11 comprende un número Y de cápsulas microfónicas B (treinta y dos, en este caso) dispuestas sobre una carcasa rígida y sustancialmente esférica C.

Cada una de las cápsulas B detecta una señal de audio procedente de una dirección espacial diferente.

30 Combinando adecuadamente estas señales, es posible obtener una señal que se correspondería con la señal que mediría un micrófono dotado de determinadas características deseadas.

35 Gracias a estas sondas, el usuario puede utilizar micrófonos "virtuales" dotados de las características de directividad (cardioide, supercardioide o similares) y posición (azimut, elevación, etc.) deseadas.

Las sondas de este tipo se utilizan generalmente en combinación con sistemas gráficos a fin de mostrar al usuario cualquier fuente de ruido y detectar cualquier defecto mecánico en una máquina (por ejemplo, un diente roto en una rueda dentada) o cualquier fuente de contaminación acústica.

40 Por este motivo, en las sondas conocidas, se concede mucha importancia a la directividad del micrófono, y se dedican muchos esfuerzos a definir filtros óptimos que puedan asegurar la mejor direccionalidad posible.

45 Una vez que se han encontrado los filtros teóricamente óptimos, se genera la señal de audio del micrófono virtual deseado por el usuario ponderando correctamente las salidas del filtro y aplicando a estos retardos y ganancias debidamente calculados y, a continuación, combinándolas para obtener determinadas formas de directividad del micrófono.

50 Un primer límite de estas sondas se relaciona con el hecho de que, aunque el uso de filtros teóricos predeterminados proporciona una buena directividad, a menudo no asegura una buena calidad de la señal de audio.

Por otra parte, otro límite de estas sondas conocidas es el hecho de que solo pueden proporcionar buena directividad hasta unas frecuencias determinadas, habitualmente en torno a los 4 kHz, mientras que por encima de este valor la directividad tiende a deteriorarse.

55 Estas sondas, por consiguiente, no son aptas para utilizar en los entornos de televisión o cinematográficos, en los que, además de la direccionalidad de los micrófonos, también es muy importante tener la capacidad de adquirir señales de audio de gran calidad.

60 El objetivo de la presente invención consiste en ofrecer un procedimiento para adquirir las señales de audio y un sistema de adquisición de audio relacionado que puedan superar los inconvenientes de la técnica anterior.

Este objetivo se alcanza a través de un procedimiento y un sistema que comprenden las características expuestas en las reivindicaciones adjuntas, que se conciben como partes integrantes de la presente descripción.

65 La presente invención se basa en la idea de procesar las señales adquiridas por las cápsulas de la sonda a partir de datos de la sonda reales medidos empíricamente durante una etapa de caracterización de la sonda.

En particular, se utilizan filtros que, en lugar de calcularse teóricamente, se determinan empíricamente durante una etapa de caracterización mediante sonda, en la que se detectan las respuestas a impulsos de las cápsulas a una o más señales de prueba predeterminadas.

5 Por lo tanto, en funcionamiento, el sistema permite detectar señales de audio de alta calidad, debido a que cualquier diferencia en el rendimiento de las cápsulas respecto de las especificaciones nominales no va a afectar a la calidad de la señal detectada.

10 Asimismo, también es posible tener en cuenta el efecto del soporte de la sonda, que rompe *de facto* la simetría perfecta de la sonda.

Además, la sonda puede mantener una buena directividad del micrófono virtual incluso a altas frecuencias por encima de 4 kHz, en la medida en que la señal de micrófono virtual no está basada en un procedimiento de filtrado teórico, sino en un procedimiento de filtrado que depende de las características concretas de la sonda y, en particular, de las respuestas a impulsos de las cápsulas, calculadas a partir de las señales de prueba determinadas de antemano durante una etapa de caracterización de la sonda.

15 Otros objetivos y ventajas de la presente invención se pondrán de manifiesto a partir de la descripción siguiente de una forma de realización de la misma representada en los dibujos adjuntos, proporcionada a título de ejemplo no limitativo, y en los que:

la figura 1 representa una sonda microfónica conocida como la descrita anteriormente;

25 la figura 2 representa esquemáticamente las etapas del procedimiento según la presente invención;

la figura 3 ilustra sintéticamente una operación de convolución utilizada por el procedimiento según la presente invención;

30 la figura 4 es un diagrama de bloques de una etapa del procedimiento según la presente invención;

la figura 5 es un diagrama de bloques de una etapa del procedimiento según la presente invención en la que se cambian los parámetros de un micrófono virtual;

35 la figura 6 ilustra un sistema de adquisición de audio 1 según la presente invención para implementar el procedimiento según la presente invención;

la figura 7 representa una primera variante del sistema de adquisición de audio según la presente invención;

40 la figura 8 representa una segunda variante del sistema de adquisición de audio según la presente invención.

Haciendo referencia a la figura 2, el procedimiento según la presente invención permite la ejecución preliminar de una primera etapa de caracterización de la sonda microfónica 11, denominada "caracterización de la sonda" en la figura 2, generando una matriz de IR (respuestas a impulsos) derivada de una medición de las respuestas de un número Y de cápsulas microfónicas de una sonda microfónica (como la sonda A descrita anteriormente) que se someten a una señal de prueba (preferentemente de tipo impulsivo) en una cámara anecoica, y de una segunda etapa (denominada "generación del filtro") de generación de una matriz de FIR (respuestas a impulsos finitas) basándose en la matriz de IR (respuestas a impulsos) y en los parámetros del micrófono virtual que puede establecer el operador.

50 En la primera etapa del procedimiento 200, la sonda microfónica 11 se coloca en una cámara anecoica (o un entorno similar) en la que se genera una o más señales de prueba, preferentemente por lo menos una señal sinusoidal cuya frecuencia se cambia a lo largo de sustancialmente todo el espectro de frecuencias audibles, es decir, se somete a un procedimiento denominado "*barrido senoidal logarítmico*", a partir de cuya convolución con una señal inversa (es decir "invertida" en el eje del tiempo) se obtiene la respuesta de la sonda al impulso. Aunque esta técnica sea conocida de por sí y por consiguiente no se describa en mayor detalle, cabe mencionar que también puede hallarse en las principales normas que definen las mediciones de las respuestas a impulsos (por ejemplo, la norma ISO 3382).

60 Para cada señal de prueba, se registran las respuestas a impulsos de cada cápsula B variando en etapas regulares (acción esquematizada en el bloque 201) el azimut y la elevación de la dirección de procedencia de la señal de prueba. En la figura 2, el azimut y la elevación relativas al centro de coordenadas (que coincide con el centro geométrico de la sonda 11) se identifican mediante las referencias M y K.

65 Esto genera un conjunto de funciones de transferencia entre cada cápsula y altavoz (que genera la señal) individuales para cada dirección en torno al centro de la sonda.

La sonda, por lo tanto, se caracteriza a lo largo de las tres dimensiones espaciales mediante un número de funciones de transferencia igual a $Y \times M \times K$, donde:

5 Y es el número de cápsulas microfónicas de la sonda microfónica 11,

M es el azimut de la señal de prueba relativa a un centro de coordenadas esféricas derivado del centro de la sonda A,

10 K es la elevación de la señal de prueba relativa a dicho sistema de coordenadas.

Estas funciones de transferencia se expresan en forma de matriz por medio de la matriz de las respuestas a impulsos IR, que se almacenan en una área de la memoria del sistema de adquisición de audio asociado a la sonda.

15 Una dimensión de la matriz de IR (el número de filas, por ejemplo) es igual a Y, mientras que la otra dimensión de la matriz de IR (el número de columnas, por ejemplo) es igual a $M \times K$.

La matriz de IR contiene datos que caracterizan las cápsulas de la sonda y, puesto que se han medido empíricamente, estos datos no son los datos nominales.

20 De esta manera, se detectan ventajosamente las características reales de la sonda 11, siendo posible adquirir, en funcionamiento, una señal de mejor calidad, ya que se tiene en cuenta el hecho de que cada una de las Y cápsulas microfónicas B puede comportarse de forma diferente a las demás, así como el hecho de que la sonda no es perfectamente esférica, debido por lo menos a la presencia de un soporte.

25 Una vez que se ha realizado esta primera etapa de caracterización de la sonda y, por consiguiente, definido la matriz de IR, es posible utilizar la sonda microfónica 11 a fin de adquirir sonido, o señales de audio, en un entorno.

30 En un entorno tridimensional, las señales recibidas por las Y cápsulas pueden proceder de varias fuentes distribuidas espacialmente.

Para elegir qué fuente debe escuchar y grabar la sonda, es necesario sintetizar un micrófono virtual a partir de las señales detectadas por las Y cápsulas microfónicas.

35 En otras palabras, las señales de audio captadas por las cápsulas reales B de la sonda microfónica 11 se procesan de tal manera que se obtiene una señal que corresponde idealmente a la que captaría un micrófono cuyos parámetros, más concretamente la dirección hacia la que apunta y la directividad, podrían ser elegidos libremente por el operador.

40 Por "directividad del micrófono" se entiende la forma en que la sensibilidad del micrófono varía cuando el ángulo de incidencia del sonido varía, que puede ser, por ejemplo, cardioide, supercardioide, cardioide de 3.^{er} orden o similar.

45 Los demás parámetros de un micrófono son, de forma más general, la sensibilidad, la curva de respuesta, el ruido, la distorsión, el rango dinámico, la impedancia y la respuesta transitoria. Sin embargo, en la presente memoria, solo se tendrán en cuenta la dirección de apuntamiento y la directividad como parámetros del micrófono virtual, y se prescindirá del resto de los parámetros indicados anteriormente.

50 El operador elige, por lo tanto, los parámetros de uno o más micrófonos virtuales que se van a utilizar en el entorno donde debe captarse el campo de sonido, por ejemplo, para concentrarse en ciertas áreas del entorno que se va a detectar con micrófonos (virtuales) que presentan una determinada directividad.

La definición de los parámetros de los micrófonos virtuales se representa esquemáticamente mediante el bloque 202 de la figura 2.

55 Según lo expuesto en la presente invención, los micrófonos virtuales se generan en la etapa de procedimiento designada en la figura 2 como "generación del filtro" (número de referencia 203), que conlleva la generación de una matriz de filtros FIR que se utiliza (como se explica con mayor detalle a continuación) para filtrar la señal captada por las cápsulas microfónicas reales B de la sonda 11.

60 Como se describirá con mayor detalle a continuación, el operador que interactúa con el sistema de adquisición de audio define los parámetros del (los) micrófono(s) virtual(es) facilitando entradas al sistema, por ejemplo, moviendo una palanca de mando y seleccionando en tiempo real una área del entorno que se va a escuchar.

65 Basándose en las entradas del operador, el sistema genera (etapa 204 de la figura 4) una matriz denominada "función objetivo" A de tamaño $(M \times K)$, que depende de las características del (los) micrófono(s) virtual(es), correspondiente a las entradas recibidas del operador.

La matriz A es, por lo tanto, una matriz que representa el modelo de directividad del micrófono virtual, es decir, una figura espacial hacia la cual debe tender el micrófono virtual.

5 Los elementos a_{ij} adoptan por lo general un valor comprendido preferentemente entre 0 y 1, que depende de las coordenadas espaciales (azimut y elevación) y la directividad del micrófono virtual deseado.

La expresión matemática de la directividad (por ejemplo, cardioide, supercardioide, cardioide de 3.^{er} orden, etc.) es conocida de por sí y se describe por medio de funciones conocidas en la literatura; por consiguiente, los expertos en la materia están capacitadas para crear la matriz A correspondiente al (los) micrófono(s) deseado(s).

A continuación, se genera la matriz H de filtros FIR (etapa 203 de las figuras 2 y 4) mediante el conocido algoritmo de Kirkeby (en notación "Matlab"):

$$H = A \cdot \frac{\text{Conj}[IRs(\omega)]}{\text{Conj}[IRs(\omega)] \cdot IRs(\omega) + \varepsilon(\omega)} \quad (1)$$

15

es decir (en notación estándar):

$$H = A \cdot [IRs(\omega)]^T * ([IRs(\omega)]^T \cdot IRs(\omega) + \varepsilon(\omega))^{-1} \quad (2)$$

20

donde:

$IRs(\omega)$ es la matriz de respuestas a impulsos generada en la etapa de caracterización descrita previamente,

25

A es la "función objetivo" generada basándose en los parámetros del micrófono virtual elegidos por el operador,

$\varepsilon(\omega)$ es un parámetro de "regularización" cuya función es evitar que el procedimiento de filtrado produzca artefactos de baja frecuencia y de alta frecuencia indeseables, constituido por una matriz de tamaño N x N, cuyos elementos de la diagonal son iguales a un mismo valor $\bar{\varepsilon}$, siendo N el número de micrófonos virtuales,

30

$\text{Conj}[IRs(\omega)]$ es una operación que genera la matriz traspuesta conjugada de la matriz $IRs(\omega)$,

H es una matriz de tamaño Y x N.

35

La elección del valor del parámetro de regularización ε en el algoritmo de Kirkeby se lleva a cabo preferentemente de forma empírica durante la etapa de caracterización de la sonda, cuando se graban las señales detectadas por la sonda mientras se miden las respuestas a impulsos de las cápsulas.

40

En esta etapa, se va cambiando el valor de ε hasta que se obtiene una señal grabada de alta calidad.

El efecto del filtrado es, en realidad, la modificación, frecuencia a frecuencia, de las amplitudes de las señales recibidas por las cápsulas, de tal forma que la suma de estas genera la señal del micrófono virtual deseada en la salida.

45

En esta etapa, es necesario ampliar algunas frecuencias de las señales procedentes de las cápsulas, por ejemplo, para llenar los huecos espectrales, y disminuir otras frecuencias para que no estén demasiado acentuadas en la señal del micrófono virtual.

50

Dependiendo del parámetro ε elegido, la matriz de filtros calculada por medio del algoritmo de Kirkeby realizará una compensación diferente para las frecuencias de las señales procedentes de las cápsulas Y y, como consecuencia, la calidad de la señal del micrófono virtual cambiará. En particular, a frecuencias bajas o altas, es necesario utilizar un parámetro de regularización diferente del que se utiliza en la banda central, a fin de limitar la inversión provocada por la fórmula de Kirkeby e impedir que el filtro calculado se desestabilice y aparezcan molestos artefactos durante la fase de escucha.

55

En particular, para obtener una señal virtual de buena calidad, el parámetro de regularización ε debe elegirse esencialmente de tal manera que sea suficientemente alto a altas frecuencias (en particular, superior a 14 kHz) y a bajas frecuencias (en particular, inferior a 100 Hz), y al mismo tiempo suficientemente bajo en la banda de frecuencias centrales, para que de ese modo la amplificación o el amortiguamiento de la frecuencia obtenidos por medio del filtrado realizado con el algoritmo de Kirkeby sea inferior a altas y bajas frecuencias y superior en la banda

60

de frecuencias centrales.

Los valores de $\bar{\epsilon}$ preferidos son los siguientes:

- 5 $0,09 \leq \bar{\epsilon} \leq 10$, más preferentemente $0,1 \leq \bar{\epsilon} \leq 3$, para frecuencias superiores a 14 kHz o inferiores a 100 Hz;
 $0,001 \leq \bar{\epsilon} \leq 0,089$, más preferentemente $0,002 \leq \bar{\epsilon} \leq 0,05$, para frecuencias entre 100 Hz y 14 kHz.

10 Con referencia nuevamente a la ecuación de la matriz (1), puede observarse que la matriz de filtros H generada se ve afectada tanto por las elecciones del operador (que influyen en la determinación de la función objetivo A) y por la caracterización de la sonda real (que influye en la determinación de la matriz de IR, bloque 206 en la figura 4).

15 Esto resulta ventajoso en la medida en que permite obtener, a partir del procedimiento de filtrado de las señales recibidas por las cápsulas reales B, un resultado sumamente natural del campo acústico del entorno, que reproducirá con fidelidad las características del entorno y al mismo tiempo aportará flexibilidad conforme a los parámetros establecidos por el operador.

20 Una vez que se ha determinado de ese modo la matriz H, se sintetizan los micrófonos virtuales filtrando las señales captadas por las cápsulas a través de los filtros determinados conforme a el procedimiento descrito.

 La señal procedente de cada cápsula se combina (etapa 207), por medio de una operación de convolución, con un filtro adecuado y, a continuación, se añade a las otras señales a fin de obtener la señal del micrófono virtual deseado:

$$\left\{ \begin{array}{l} Virtual_Mic_1 = \sum_{i=1}^y FIR_{i,1} \otimes Ch_i \\ \cdot \\ \cdot \\ Virtual_Mic_N = \sum_{i=1}^y FIR_{i,N} \otimes Ch_i \end{array} \right.$$

25 donde:

30 Virtual_Mic_1..N indica la señal de audio detectada por cada micrófono virtual,

 FIR_{i,1..N} indica el elemento i,1..N de la matriz H,

 Ch_i indica la señal captada por la i-ésima cápsula microfónica de la sonda.

35 Se representa asimismo un diagrama de dicha convolución en la figura 3, mientras que la segunda etapa del procedimiento, denominada "generación del filtro", se representa también en el flujo de información de la figura 4.

40 El procedimiento descrito anteriormente permite cambiar ventajosamente los parámetros del micrófono virtual en tiempo real.

 El operador puede cambiar los parámetros del micrófono virtual durante su uso (por ejemplo, con el propósito de seguir a un actor en una escena cinematográfica o la acción que tiene lugar en un punto determinado del entorno) manipulando una consola de control dedicada.

45 Después de recibir una entrada correspondiente a un cambio de los parámetros de uno de los micrófonos virtuales o una petición de añadir o eliminar un micrófono virtual, el sistema volverá a calcular la matriz de filtros H.

 El diagrama de flujo de esta operación se representa en la figura 5.

50 Después de encender un micrófono virtual (etapa 500), se comprueba si ha llegado una entrada que requiera un cambio de acimut (etapa 501); de no ser así, se comprueba si ha llegado una entrada que requiera un cambio de elevación (etapa 502) y, si esta comprobación también da un resultado negativo, se comprueba si ha llegado una entrada que requiera un cambio de directividad (etapa 503).

55 Si esta última comprobación resulta también negativa, el procedimiento vuelve a la etapa 501.

De lo contrario, si cualquiera de las comprobaciones realizadas en las etapas 501 a 503 da un resultado positivo, los coeficientes de la función objetivo A se recalculan conforme a la nueva entrada (etapa 504).

5 Una vez que se han cambiado los coeficientes, el procesador puede utilizarlos para generar la matriz de filtros H.

El algoritmo esquematizado en la figura 5 permite comprobar si el micrófono está aún activo o no (etapa 505) después de que se hayan actualizado los coeficientes de la matriz A. Si el micrófono sigue activo, el procedimiento vuelve a la etapa 501 y los parámetros del micrófono virtual se comprueban de nuevo; si el micrófono no está activo, se interrumpe el algoritmo (etapa 506).

Por consiguiente, dicho brevemente, cuando el operador varía el azimut y/o la elevación y/o la directividad del micrófono virtual (y, por lo tanto, los parámetros de este), los coeficientes de la matriz de la función objetivo A se modifican como corresponde y la matriz H se recalcula.

15 Según otro perfeccionamiento, también es posible cambiar el micrófono virtual sin generar una sensación de movimiento "brusco" provocado por perturbaciones o ruido de masa. Esto puede lograrse ejecutando un "fundido encadenado" entre el audio del micrófono virtual que se está utilizando y el del micrófono virtual hasta el cual el operador desea desplazarse.

20 En esencia, cuando el operador cambia el micrófono virtual en uso y elige un segundo micrófono, el cambio entre una primera matriz H correspondiente a un primer micrófono (el micrófono en uso) y una segunda matriz H correspondiente a un segundo micrófono (el micrófono al cual quiere cambiar el operador) se lleva a cabo gradualmente por medio de un conjunto ordenado de matrices de transacción (o sea, filtros de transacción). El sonido captado por las cápsulas B se filtra con las matrices de transacción según el orden de estas. De forma más detallada, el conjunto ordenado de matrices de transacción $T_1, T_2, T_3, \dots, T_n$ permite alternar entre la primera matriz y la segunda matriz de la siguiente manera: al principio, el sonido se filtra mediante la primera matriz, a continuación se filtra mediante la matriz de transacción T_1 , a continuación mediante la matriz de transacción T_2 , a continuación, mediante la matriz de transacción T_3 y así sucesivamente hasta llegar a la segunda matriz.

30 Cada una de las matrices de transacción $T_1, T_2, T_3, \dots, T_n$ comprende submatrices correspondientes a submatrices que pertenecen a la primera matriz o la segunda matriz. En particular, la matriz de transacción T_k (correspondiente a la k-ésima matriz del conjunto ordenado de matrices de transacción, siendo $k = 2 \dots n$) comprende un número de submatrices correspondiente a las submatrices de la segunda matriz que es superior al número de submatrices que comprende una matriz de transacción anterior T_{k-1} . Por otra parte, la matriz de transacción T_k comprende un número de submatrices correspondiente a las submatrices de la primera matriz que es inferior al número de submatrices que comprende la matriz de transacción anterior T_{k-1} .

40 Entonces, utilizando una sintaxis matemática, las matrices de transacción comprenden submatrices que cumplen la siguiente condición:

$$\#S2_k > \#S2_{k-1} \text{ y } \#S1_k < \#S1_{k-1}, k = 2 \dots n$$

donde:

45 $\#S2_k$ indica el número de submatrices de la matriz de transacción T_k que corresponden a las submatrices de la segunda matriz,

50 $\#S2_{k-1}$ indica el número de submatrices de la matriz de transacción T_{k-1} que corresponden a submatrices de la segunda matriz,

$\#S1_k$ indica el número de submatrices de la matriz de transacción T_k que corresponden a submatrices de la primera matriz,

55 $\#S1_{k-1}$ indica el número de submatrices de la matriz de transacción T_{k-1} que corresponden a submatrices de la primera matriz, siendo el índice k un valor entero entre 2 y n, y siendo n el número de matrices de transacción.

Por consiguiente, la matriz de transacción T_1 es la más parecida a la primera matriz, mientras que la matriz de transacción T_n es la más parecida a la segunda matriz.

60 En una forma de realización preferida, todas las submatrices presentan las mismas dimensiones y, en particular, una de las dimensiones (fila o columna) es igual a N.

65 La alternancia entre diferentes filtros (es decir, diferentes matrices) puede realizarse mediante un "fundido encadenado" estándar (es decir, una disminución del nivel de la señal de audio correspondiente a un filtro, acompañada de un incremento de la señal de audio correspondiente a otro filtro) entre el audio que procede de un

filtro en uso y el que procede de un filtro siguiente. A continuación, la señal del filtro en uso y la del siguiente filtro se mezclan para realizar un fundido gradual del volumen de la primera hasta cero e incrementar progresivamente el volumen de esta última hasta el valor máximo, con lo cual se logra transmitir al usuario una sensación de gran suavidad.

5 Con referencia a la figura 6, se representa un sistema de adquisición de audio 1 para implementar el procedimiento descrito anteriormente.

10 El sistema 1 permite captar señales de audio procedentes de un entorno.

El sistema 1 comprende una sonda microfónica 11 que comprende una pluralidad de cápsulas (por ejemplo, una sonda microfónica de 32 canales denominada "*em32 Eigenmike*", comercializada por la empresa *mhAcoustics*), cuyas señales se preamplifican y convierten en señales digitales.

15 La sonda 11 está conectada a un ordenador electrónico 3 equipado con una interfaz de audio 2 (por ejemplo, una interfaz de audio EMIB Firewire), que recibe las señales de la sonda y las transmite, tal vez después de haberlas procesado, a un procesador 300, por ejemplo, un procesador de señales digitales (DSP), programado para ejecutar el procedimiento de adquisición de audio descrito anteriormente.

20 El sistema 1 comprende además una unidad de entrada de datos o mandatos 4 conectada también al ordenador 3, por ejemplo, a través de un puerto bus serie universal (USB), por medio de la cual un operador puede facilitar información acerca del área donde debe adquirirse el sonido o introducir directamente los parámetros de uno o más micrófonos virtuales (por ejemplo, seleccionando formas de directividad predefinidas por medio de botones).

25 La unidad de entrada de datos o mandatos 4 puede ser, por ejemplo, una consola de control provista de una palanca para controlar el apuntamiento de los micrófonos virtuales.

El sistema 1 comprende además una grabadora 5 y/o una salida analógica 6 y/o una salida digital 7, a través de las cuales se puede grabar o transmitir la señal captada por el (los) micrófonos virtual(es).

30 Aunque en el ejemplo de la figura 6 la grabadora 5, la salida analógica 6 y la salida digital 7 están instaladas dentro del ordenador 3, como alternativa, la grabadora 5 puede hallarse en el exterior del ordenador 3 y estar conectada al mismo.

35 La figura 7 representa una versión perfeccionada del sistema 1, designada por 1'. El sistema perfeccionado permite adquirir señales de audio de un entorno y sincronizarlas con imágenes de vídeo del mismo entorno.

40 Además de las partes que se designan por los mismos números de referencia en la figura 6 y que tienen las mismas funciones, el sistema 1' comprende también una cámara de vídeo 8 que filma el entorno cuyas señales de audio van a detectarse mediante la sonda 11, unos medios de interfaz gráfica 9 y un temporizador 10 (preferentemente situado en el interior del ordenador 3 y conectado al procesador 300) para sincronizar el audio captado por la sonda 11 con el vídeo captado por la cámara de vídeo 8.

45 La cámara de vídeo 8 ajusta la imagen del entorno donde tiene lugar la escena cuyo audio se desea adquirir y, por este motivo, la cámara de vídeo 8 es una cámara de vídeo gran angular, por ejemplo, del tipo "abovedado" utilizado normalmente para fines de vigilancia o similares.

50 La cámara de vídeo 8 transmite la señal de vídeo adquirida a los medios de interfaz gráfica 9, que comprenden un monitor para visualizar las imágenes tomadas por la cámara de vídeo 8.

Los mismos medios de interfaz gráfica 9 están operativamente conectados a la unidad de entrada de datos o mandatos 4 y, por consiguiente, reciben información sobre el (los) micrófono(s) virtual(es) seleccionados por el operador.

55 Los medios de interfaz gráfica 9 procesan esta información y la plasman gráficamente; en particular, presentan, superpuesto a las imágenes tomadas por la cámara de vídeo 8, un puntero móvil que indica el área que se está escuchando mediante el micrófono virtual elegido por el operador.

60 Preferentemente, la forma y el tamaño del puntero están relacionados con la directividad y la orientación del micrófono, para reflejar los parámetros del micrófono en uso y permitir al operador llevar un control más intuitivo de este.

65 La unidad de entrada de datos o mandatos 4 puede estar ventajosamente provista de una palanca de control o un interruptor de mando deslizante o similar para permitir al operador acercar o alejar el campo de sonido del micrófono virtual de una manera rápida e intuitiva.

A través de la unidad de entrada de datos o mandatos 4, el operador mueve, por lo tanto, el micrófono dentro de la escena filmada y puede escuchar por separado diferentes fuentes de sonido comprendidas en las imágenes tomadas.

5 Al mover la palanca de mando, el operador mueve el micrófono virtual y puede seguir el movimiento de este, gracias a las imágenes visualizadas por los medios de interfaz gráfica 9. El operador puede controlar la directividad a través del interruptor deslizante, y el puntero adopta el tamaño adecuado.

10 En otra forma de realización alternativa, el puntero puede reemplazarse por unas áreas de colores correspondientes a las zonas escuchadas mediante el micrófono; por ejemplo, el área que mejor se recibe puede aparecer en color rojo, y las demás áreas pueden aparecer en colores más fríos conforme a su calidad de recepción. Cuando se mueve el micrófono virtual o se cambia la directividad de este, el color de las imágenes también cambia.

15 La figura 8 representa una variante del sistema de la figura 7.

En este ejemplo, el operador tiene la posibilidad de ajustar los parámetros del micrófono virtual a través de la unidad de entrada de datos o mandatos 4 o la interfaz gráfica 90, para apuntar de ese modo el micrófono virtual (en términos de azimut y elevación) y seleccionar su directividad (cardioide, supercardioide, cardioide de 3.^{er} orden, etc.).

20 Con este propósito, los medios de interfaz gráfica 90 de la figura 8 comprenden una pantalla táctil en la que se visualizan las imágenes procedentes de la cámara de vídeo 8 y el puntero del micrófono, tal como se ha indicado anteriormente con referencia a la figura 7.

25 Por medio de su interacción con la pantalla táctil, el operador puede mover el micrófono o cambiar la extensión del espacio que se va a escuchar, es decir, cambiar la orientación y la directividad del micrófono.

30 Los datos del micrófono virtual definidos de esta manera por el usuario se envían al procesador 300, donde la ejecución de algunas partes del código permite la generación de la función objetivo A mencionada anteriormente y el cálculo del algoritmo de Kirkeby, que se lleva a cabo mediante la matriz de IR de respuestas a impulsos (medidas en la etapa de caracterización de la sonda mencionada anteriormente) precargada en la memoria y relativa a la sonda microfónica 11.

A continuación, se genera la matriz de filtros H tal como se ha descrito anteriormente.

35 El archivo que contiene los coeficientes de filtro FIR se utiliza entonces para realizar el procedimiento de filtrado con los datos de audio procedentes de la sonda microfónica 11.

40 La señal del micrófono virtual sintetizada mediante dicho procedimiento de filtrado se transmite a una interfaz Jack 15, que a continuación puede transmitirla a las salidas digitales (ADAT) de la tarjeta EMIB o desviarla hacia una tarjeta de memoria.

45 Cada vez que se cambian los parámetros del micrófono virtual (por ejemplo, cuando se cambia la directividad), se ejecuta de nuevo el algoritmo de Kirkeby y se calcula otra matriz H, de tal forma que el cambio se produce en tiempo real.

En este sentido, el procesador 3 o el procesador 300 comprenden preferentemente un área de memoria (por ejemplo, una memoria flash) en la que se almacena la matriz

$$\Gamma = \frac{\text{Conj}[IRs(\omega)]}{\text{Conj}[IRs(\omega)] \cdot IRs(\omega) + \varepsilon(\omega)}$$

50 calculada durante la etapa de caracterización de la sonda y, por consiguiente, dependiente de las respuestas a impulsos de las cápsulas calculadas mediante las señales de prueba predeterminadas y conocidas.

55 Esta solución permite reducir el coste computacional del procedimiento de captación de audio descrito anteriormente, en la medida en que no es necesario recalcularse Γ para recalcularse la matriz H, sino solo el producto de las matrices A y Γ .

60 Aunque en la presente memoria se ha descrito la presente invención haciendo referencia a algunas formas de realización preferidas, como resultará evidente para los expertos en la materia, es posible realizar algunos cambios en el sistema de adquisición de audio y el procedimiento de adquisición de audio descritos anteriormente.

En particular, los diversos elementos y bloques lógicos del sistema de adquisición de audio pueden presentar muchas composiciones y distribuciones diferentes y, al mismo tiempo, desempeñar en su conjunto las mismas

funciones descritas en la presente invención o funciones equivalentes a estas.

REIVINDICACIONES

1. Procedimiento para adquirir señales de audio, en el que una sonda microfónica (11) provista de una pluralidad de cápsulas microfónicas (B) detecta una pluralidad de señales de audio, y en el que dichas señales de audio detectadas se combinan a fin de obtener una señal de un micrófono virtual (207),

en el que dicha señal de un micrófono virtual se genera en función de unos datos de sonda característicos (206) medidos durante una etapa de caracterización de la sonda, en el que se miden las señales detectadas por cada cápsula microfónica (B) después de una señal de prueba predeterminada correspondiente, en el que cada cambio en los parámetros de dicho micrófono virtual es seguido por una nueva generación de filtros (203) que pueden utilizarse para filtrar las señales recibidas por dicha pluralidad de cápsulas y generar una nueva señal de audio de dicho micrófono virtual,

caracterizado porque, cuando los parámetros de dicho micrófono virtual se cambian para conmutar de un primer micrófono virtual, correspondiente a un primer filtro, a un segundo micrófono virtual, tiene lugar lo siguiente:

- se calcula un segundo filtro correspondiente al segundo micrófono virtual;
- se calcula un conjunto ordenado de filtros de transacción, en el que cada uno de dichos filtros de transacción comprende submatrices correspondientes a submatrices de dicho primer filtro o dicho segundo filtro, en el que el número de submatrices del segundo filtro de dicho filtro de transacción es mayor que el número de submatrices del segundo filtro de un filtro de transacción anterior,
- y en el que el número de submatrices del primer filtro de dicho filtro de transacción es menor que el número de submatrices del primer filtro de un filtro de transacción anterior;
- la señal captada por dichas cápsulas (B) se filtra a través de dichos filtros de transacción según el orden de dicho conjunto de filtros de transacción;
- después del último filtro de transacción de dicho conjunto, la señal captada por dichas cápsulas (B) se filtra a través de dicho segundo filtro.

2. Procedimiento según la reivindicación 1, en el que la etapa de caracterización de la sonda comprende por lo menos las etapas siguientes:

- someter a dicha sonda (11), a varias señales de prueba, cuyas coordenadas de emisión M, K relativas a la sonda (11) son conocidas,
- detectar las señales captadas por cada cápsula microfónica (B) de dicha sonda (11) en dichas señales de prueba,
- generar una matriz de las respuestas a impulsos de dichas cápsulas.

3. Procedimiento según la reivindicación 1 o 2, en el que dicha señal de un micrófono virtual se calcula basándose en los parámetros deseados, en particular, la orientación y la directividad, del micrófono virtual.

4. Procedimiento según la reivindicación 2 o 3, en el que dicha señal de un micrófono virtual se genera filtrando las señales recibidas por dicha pluralidad de cápsulas (Y) a través de un filtro H calculado según la fórmula siguiente:

$$H = A \cdot \frac{Conj[IRs(\omega)]}{Conj[IRs(\omega)] \cdot IRs(\omega) + \epsilon(\omega)}$$

en la que:

IRs(ω) es la matriz de las respuestas a impulsos de cada cápsula microfónica (B) como respuesta a dicha señal de prueba predeterminada,

A es una matriz denominada "de función objetivo" generada basándose en dichos parámetros de dicho micrófono virtual,

ε(ω) es un parámetro de ajuste predefinido.

5. Procedimiento según cualquiera de las reivindicaciones 1 a 4, en el que, para conmutar de un filtro en uso al filtro

que sigue a dicho filtro en uso, tiene lugar lo siguiente:

- se calcula dicho filtro que sigue a dicho filtro en uso;
- 5 - se filtra la señal captada por dichas cápsulas (B), a través de dicho filtro que sigue a dicho filtro en uso;
- se mezclan entre sí las señales de dicho filtro en uso y de dicho filtro que sigue a dicho filtro en uso;
- 10 - se reduce el nivel de la señal de dicho filtro en uso en proporción con el incremento del nivel de la señal de dicho filtro que sigue a dicho filtro en uso.

6. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que una cámara de vídeo (8) toma imágenes de un área en la que van a adquirirse señales de audio por medio de dicho micrófono virtual, en el que dichas imágenes tomadas se visualizan en un monitor, y en el que dicho por lo menos un elemento gráfico, en particular, un puntero cuyos forma y/o tamaño dependen de las características de dicho micrófono virtual, se superpone sobre dichas imágenes visualizadas.

7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que un operador establece las características de orientación y/o directividad de dicho micrófono virtual.

8. Sistema de adquisición de audio (1,1'), que comprende por lo menos una sonda microfónica (11) provista de una pluralidad de cápsulas microfónicas (B) para detectar una pluralidad de señales de audio, y por lo menos un procesador (300) adaptado para combinar las señales recibidas por dicha pluralidad de cápsulas (B) a fin de obtener una señal de un micrófono virtual,

25 que comprende además un área de memoria que almacena datos característicos de dichas cápsulas medidos después de una señal de prueba predeterminada, en el que dicho procesador comprende unas partes de código que, una vez ejecutadas, permiten que se genere dicha señal de un micrófono virtual basándose en dichos datos característicos de las cápsulas, caracterizado porque dicho sistema de adquisición de audio (1,1') está adaptado para implementar el procedimiento según cualquiera de las reivindicaciones 1 a 7.

9. Sistema (1,1') según la reivindicación 8, que comprende además unos medios (4, 9) practicables por un operador de dicho sistema para establecer los parámetros de por lo menos un micrófono virtual.

35 10. Sistema (1,1') según la reivindicación 9, en el que dichos medios practicables por un operador comprenden una pantalla táctil.

40 11. Sistema (1,1') según la reivindicación 8 o 9 o 10, caracterizado porque comprende además una grabadora (5) y/o una salida analógica (6) y/o una salida digital (7) para grabar y/o transmitir la señal captada por dicho por lo menos un micrófono virtual.

45 12. Sistema (1') según cualquiera de las reivindicaciones 8 a 11, en el que dicho sistema comprende una cámara de vídeo (8) conectada funcionalmente a unos medios de interfaz gráfica (9) adaptados para visualizar en un monitor las imágenes tomadas por dicha cámara de vídeo, y en el que dicho procesador (300) está adaptado para transmitir información acerca de las características de dicho micrófono virtual a dichos medios de interfaz gráfica, de tal manera que dichos medios de interfaz gráfica puedan generar un elemento gráfico adaptado para superponerse a dichas imágenes visualizadas en dicho monitor, y representativo de dicho micrófono virtual.

50 13. Sistema según cualquiera de las reivindicaciones 8 a 12, en el que dicho sistema comprende una cámara de vídeo (8) conectada funcionalmente a los medios de interfaz gráfica (9) adaptados para visualizar en un monitor las imágenes tomadas por dicha cámara de vídeo (8), y en el que dicho sistema comprende un temporizador (10) para sincronizar el audio captado por la sonda (11) con el vídeo captado por la cámara de vídeo (8).

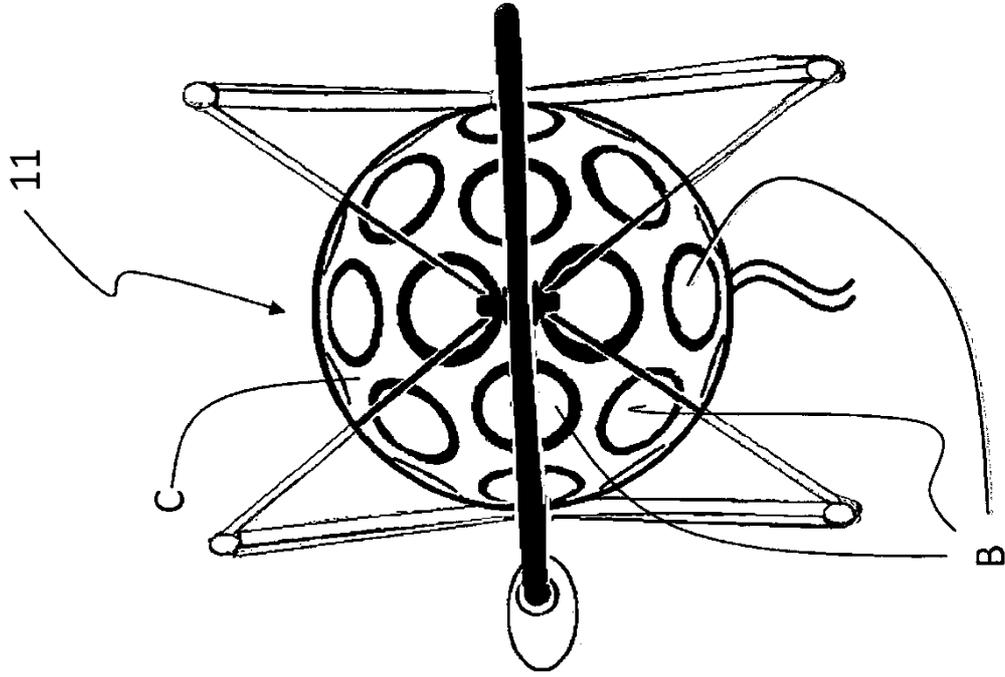


Fig.1

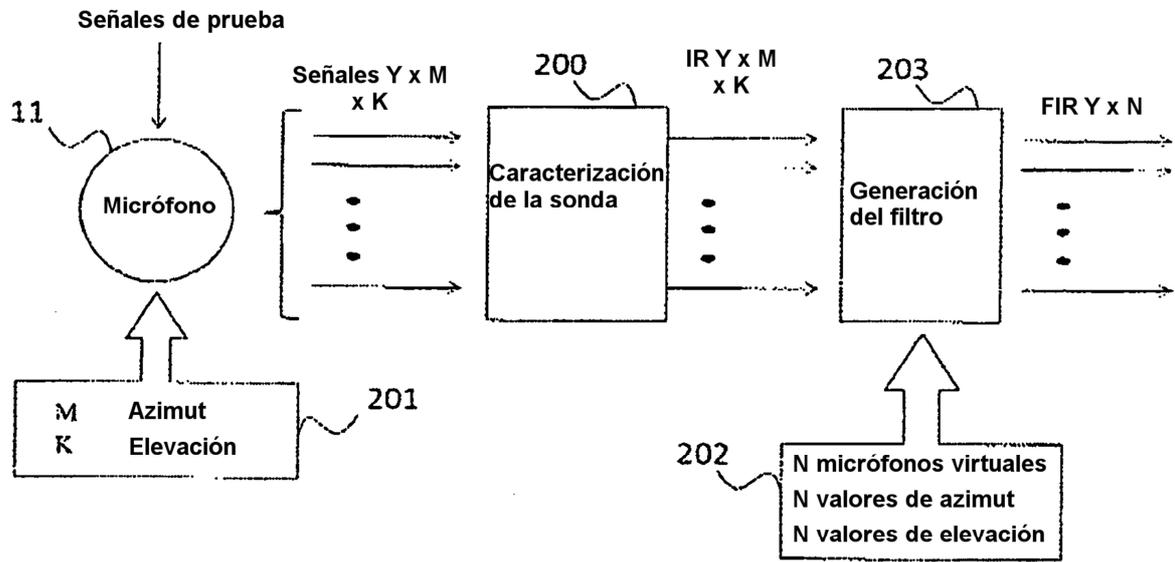


Fig.2

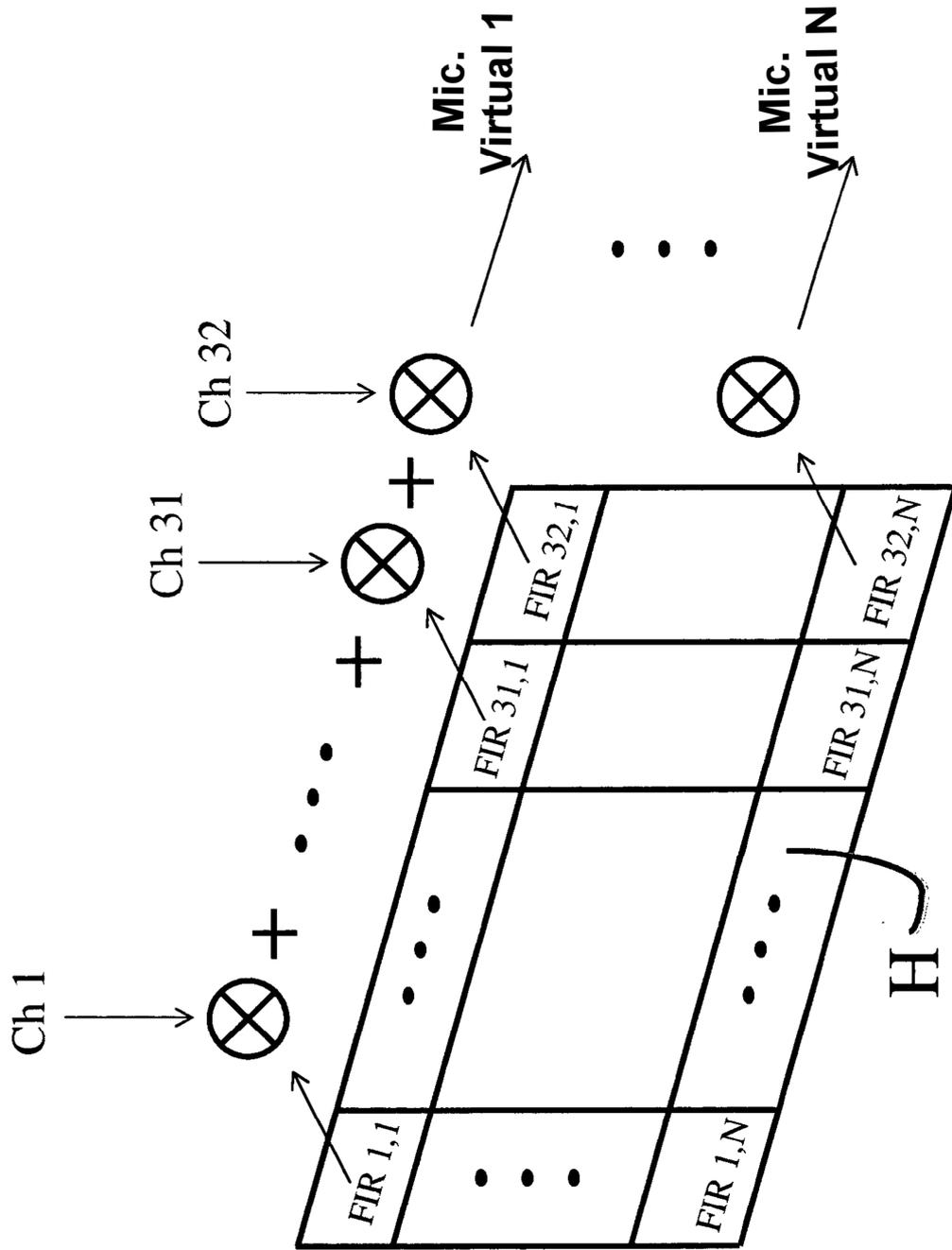


Fig.3

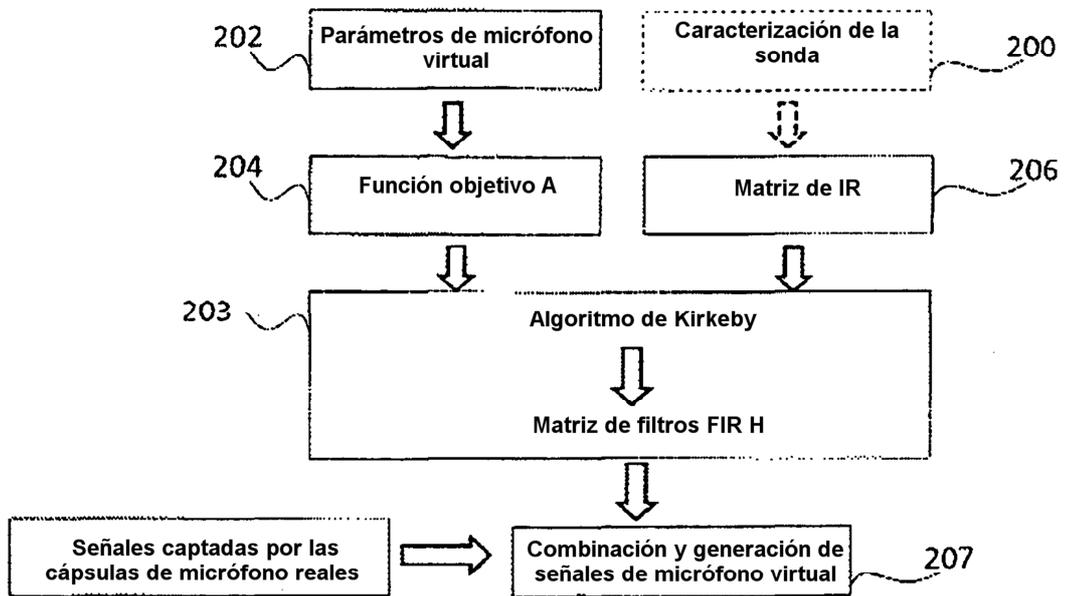


Fig. 4

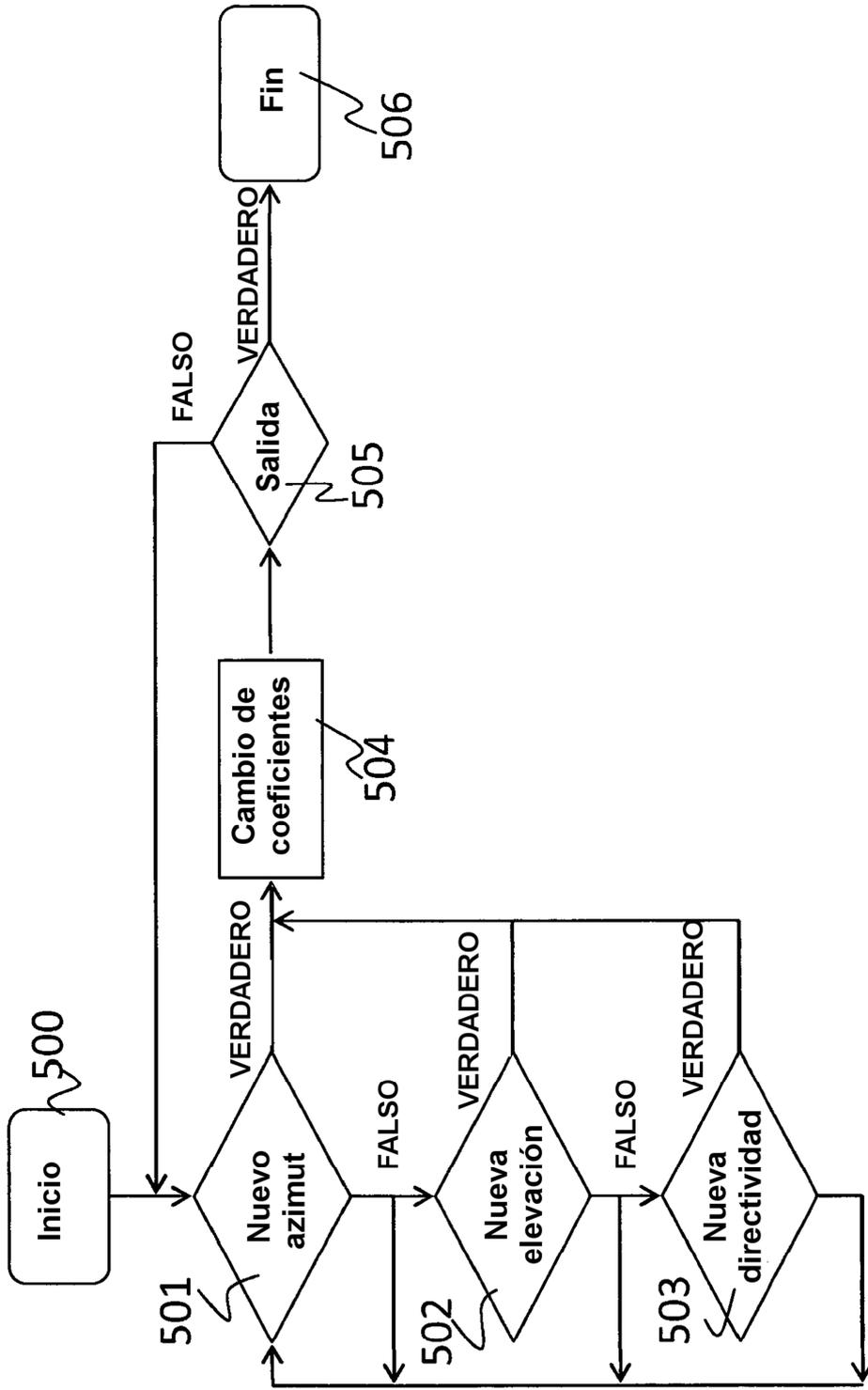


Fig.5

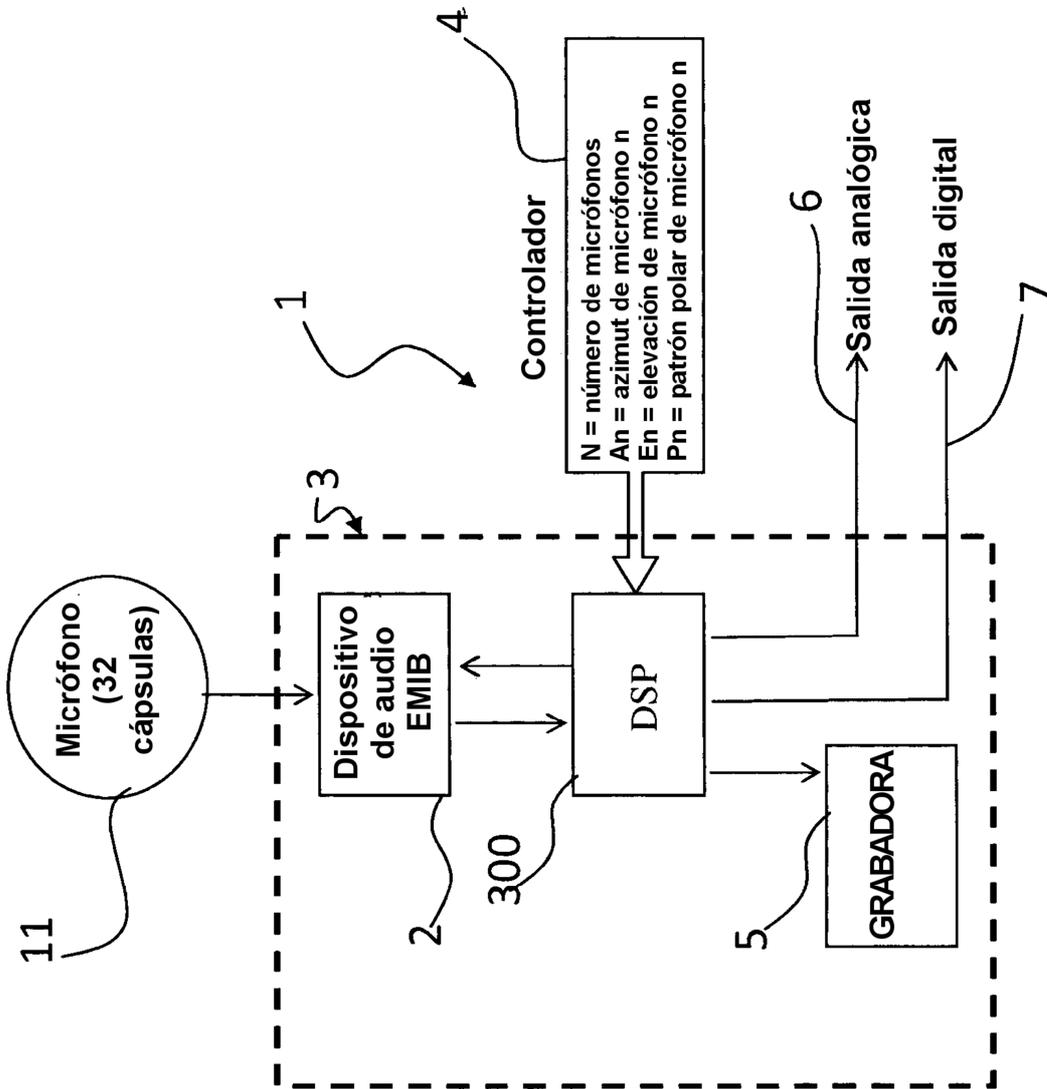


Fig.6

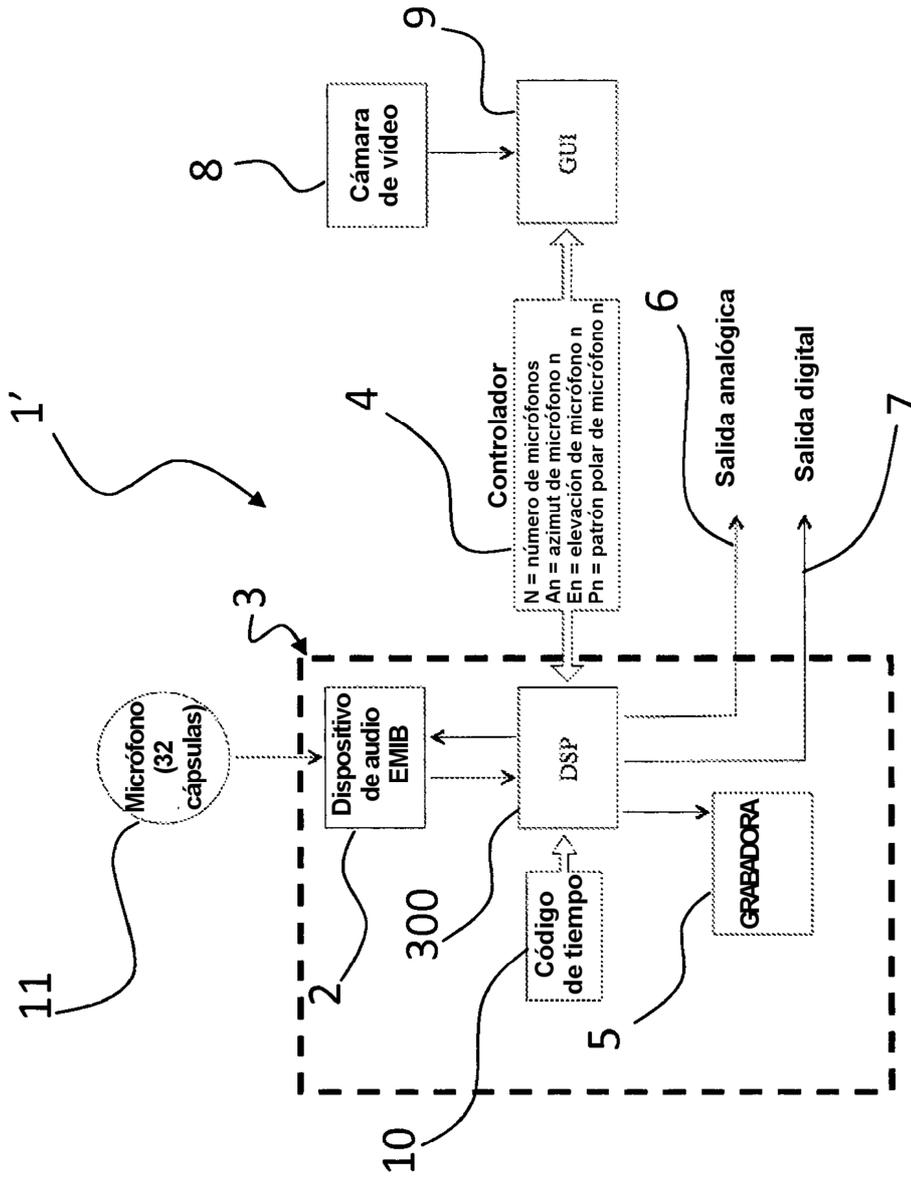


Fig.7

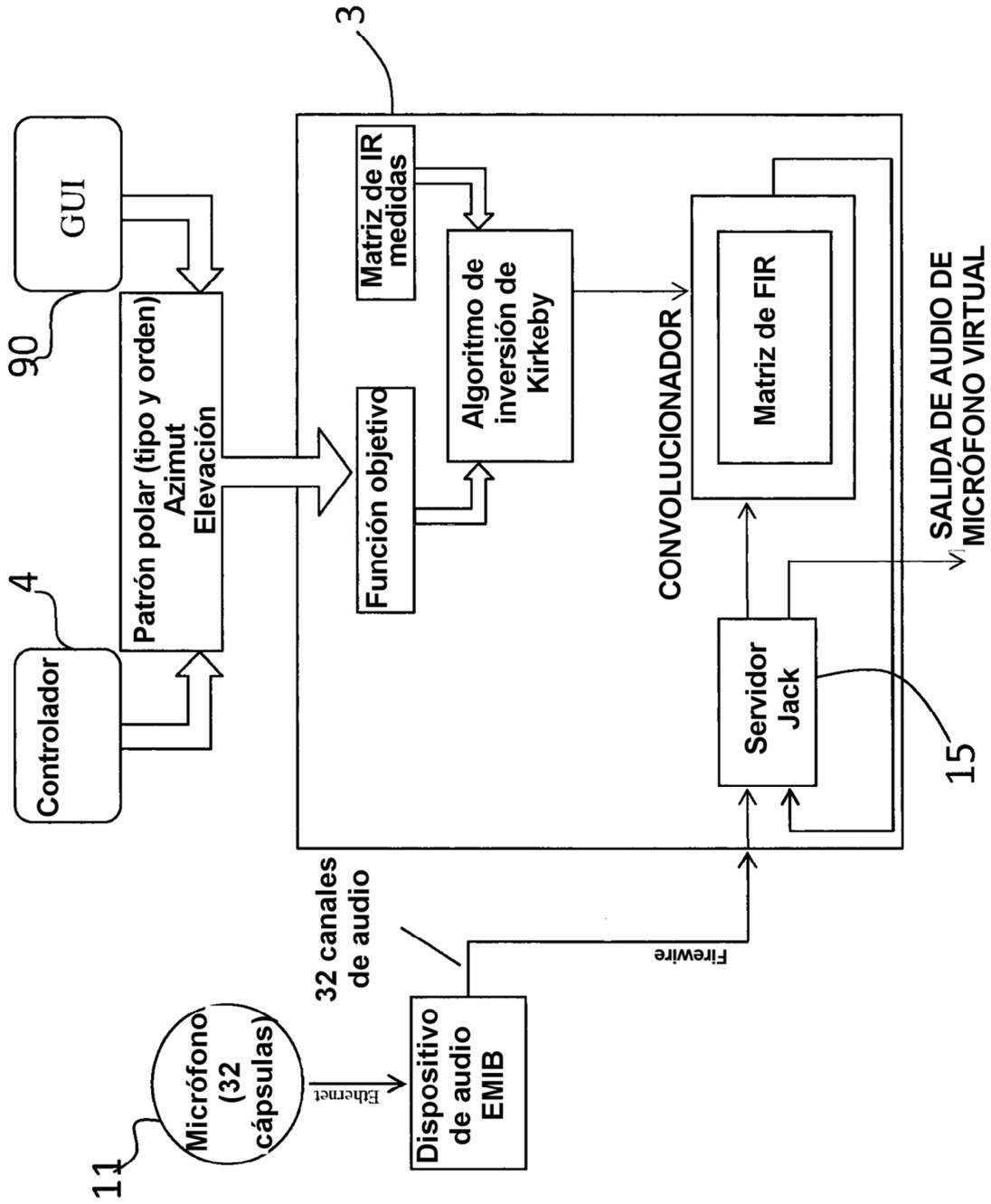


Fig.8