

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 512 640**

51 Int. Cl.:

H04H 20/14 (2008.01)

H04H 60/58 (2008.01)

G10L 25/48 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.11.2008** **E 08873178 (1)**

97 Fecha y número de publicación de la concesión europea: **23.07.2014** **EP 2263335**

54 Título: **Métodos y aparato para generar firmas**

30 Prioridad:

05.03.2008 US 33992 P

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

24.10.2014

73 Titular/es:

THE NIELSEN COMPANY (US), LLC (100.0%)
150 North Martingale Road
Schaumburg, IL 60173, US

72 Inventor/es:

TOPCHY, ALEXANDER PAVLOVICH;
RAMAMSWAMY, ARUN y
SRINIVASAN, VENUGOPAL

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 512 640 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

Métodos y aparato para generar firmas

- 5 Esta solicitud reivindica el beneficio de la Solicitud Provisional de Estados Unidos 61/033.992, presentada el 05 de marzo 2008.

CAMPO DE LA DIVULGACIÓN

- 10 La presente divulgación se refiere, en general, a la monitorización de medios, a la búsqueda y recuperación de contenidos multimedia y, más especialmente, a unos métodos y aparatos para generar firmas para su uso en la información de los medios de identificación.

ANTECEDENTES

- 15 La información de los medios de identificación y, más específicamente, las señales de audio (por ejemplo, la información en los flujos de audio) que usan técnicas de emparejamiento de firma están bien establecidas. Las firmas se conocen también de forma equivalente, y con frecuencia se denominan como huellas dactilares. Las técnicas de emparejamiento de firma se usan a menudo en aplicaciones de medición de audiencia de televisión y radio y se implementan usando varios métodos para generar y emparejar firmas. Por ejemplo, en aplicaciones de medición de audiencia de televisión, las firmas se generan en los sitios de monitorización (por ejemplo, hogares monitorizados) y sitios de referencia. Los sitios de monitorización en general incluyen lugares tales como, por ejemplo, los hogares en los que se monitoriza el consumo de medios de los miembros de la audiencia. Por ejemplo, en un sitio de monitorización, las firmas monitorizadas pueden generarse en base a flujos de audio asociados con un canal seleccionado, una estación de radio, etc. A continuación, las firmas monitorizadas pueden enviarse a una instalación de recopilación de datos central para su análisis. En un sitio de referencia, las firmas, normalmente denominadas como firmas de referencia, se generan en base a programas conocidos que se proporcionan dentro de una región de difusión. Las firmas de referencia pueden almacenarse en el sitio de referencia y/o en una instalación de recopilación de datos central y se comparan con las firmas monitorizadas generadas en los sitios de monitorización. Puede encontrarse una firma monitorizada que se empareje con una firma de referencia y el programa conocido correspondiente a la firma de referencia emparejada puede identificarse como el programa que se ha presentado en el sitio de monitorización.

- 35 El número de publicación internacional WO 02/065782 A1 divulga un método que genera algoritmos robustos para el contenido multimedia, por ejemplo, clips de audio. El clip de audio se divide en tramas sucesivas (preferentemente emparejadas). Para cada trama, el espectro de frecuencias se divide en bandas. Una propiedad robusta de cada banda (por ejemplo, la energía) se calcula y se representa por un bit hash respectivo. Por lo tanto, un clip de audio se representa por una concatenación de palabras hash binarias, una para cada trama. Para identificar una posible señal de audio comprimido, un bloque de palabras hash derivado de la misma se emparejan mediante un ordenador con una gran base de datos. También se describen tales estrategias de emparejamiento. En una realización ventajosa, el proceso de extracción también proporciona información en cuanto a cuales de los bits hash son los menos fiables. Rechazando estos bits se mejora considerablemente la velocidad y el rendimiento del proceso de emparejamiento.

SUMARIO

La invención se dirige a métodos y aparatos para generar firmas como se define en las reivindicaciones adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

- 50 Las figuras 1A y 1B ilustran un ejemplo de sistemas de identificación de flujo de audio para generar firmas e identificar flujos de audio.
- La figura 2 es un diagrama de flujo que ilustra un proceso de generación de firma de ejemplo.
- 55 La figura 3 es una representación del dominio de tiempo de un flujo de audio monitorizado de ejemplo.
- La figura 4 es una representación gráfica de un ejemplo de una parte del flujo de audio monitorizado (es decir, un bloque de audio) que es una senoide.
- 60 La figura 5 es una representación gráfica de una ventana de ejemplo que puede aplicarse al bloque de audio de la figura 4.
- La figura 6 es una representación gráfica de un bloque de audio de ventana resultante de una aplicación de la ventana de la figura 5 al bloque de audio de la figura 4.
- 65

La figura 7 es una representación gráfica de una segunda ventana de ejemplo que puede aplicarse al bloque de audio de la figura 4.

5 La figura 8 es una representación gráfica de un bloque de audio de ventana resultante de una aplicación de la ventana de la figura 7 al bloque de audio de la figura 4.

La figura 9 es una representación gráfica de la ventana de la figura 5 que muestra la respuesta de frecuencia correspondiente de la ventana.

10 La figura 10 es una representación gráfica de la ventana de la figura 7 que muestra la respuesta de frecuencia correspondiente de la ventana.

La figura 11 es una representación gráfica de una segunda ventana de ejemplo alternativa y la respuesta de frecuencia correspondiente a la misma.

15 La figura 12 es una representación gráfica de una tercera ventana de ejemplo alternativa y la respuesta de frecuencia correspondiente a la misma.

La figura 13 es un diagrama de flujo de un proceso de emparejamiento de firma de ejemplo.

20 La figura 14 es un diagrama que muestra cómo pueden compararse unas firmas de acuerdo con el diagrama de flujo de la figura 13.

La figura 15 es un diagrama de bloques de un sistema de generación de firma de ejemplo para generar firmas en base a flujos de audio o bloques de audio.

25 La figura 16 es un diagrama de bloques de un sistema de comparación de firmas de ejemplo para comparar firmas.

La figura 17 es un diagrama de bloques de un sistema de procesador de ejemplo que puede usarse para implementar los métodos y los aparatos descritos en el presente documento.

DESCRIPCIÓN DETALLADA

35 Aunque lo siguiente divulga unos sistemas de ejemplo implementados usando, entre otros componentes, software ejecutado en hardware, debería señalarse que dichos sistemas son meramente ilustrativos y no deberían considerarse como limitativos. Por ejemplo, se contempla que cualquiera o todos estos componentes de hardware y de software podrían realizarse exclusivamente en hardware, exclusivamente en software, o en cualquier combinación de hardware y software.

40 En consecuencia, mientras que lo siguiente describe sistemas de ejemplo, los expertos en la materia apreciarán fácilmente que los ejemplos proporcionados no son la única manera de implementar tales sistemas.

Los métodos y aparatos descritos en el presente documento se refieren en general a la generación de firmas digitales que pueden usarse para identificar la información de los medios. Una firma digital, o huella digital, es un descriptor de señal que caracteriza con precisión las señales con el fin de emparejar, indexar o buscar y recuperar en una base de datos. En particular, los métodos y aparatos divulgados se describen con respecto a la generación de firmas de audio digitales en base a flujos de audio o bloques de audio (por ejemplo, información de audio). Sin embargo, los métodos y aparatos descritos en el presente documento también pueden usarse para generar firmas digitales en base a cualquier otro tipo de señales, datos de series de tiempo e información de los medios, tales como, por ejemplo, información de vídeo, páginas web, imágenes fijas, datos informáticos, etc. Además, la información de los medios puede estar asociada con información de difusión (por ejemplo, información de televisión, información de radio, etc.), información reproducida a partir de cualquier medio de almacenamiento (por ejemplo, discos compactos (CD), discos versátiles digitales (DVD), etc.), o cualquier otra información que se asocie con un flujo de audio, un flujo de vídeo, o cualquier otra información de medio para el que se generan las firmas digitales. En un ejemplo específico, los flujos de audio se identifican en base a firmas digitales, que incluyen firmas monitorizadas digitales generadas en un sitio de monitorización (por ejemplo, un hogar monitorizado) y firmas digitales de referencia generadas y/o almacenadas en un sitio de referencia y/o una instalación de recopilación de datos central.

60 Como se describe en detalle a continuación, los métodos y aparatos descritos en el presente documento identifican la información de los medios, incluyendo los flujos de audio o cualquier otro medio, en base a las firmas digitales. Las técnicas de ejemplo descritas en el presente documento calculan una firma en un momento determinado usando, por ejemplo, un único bloque de audio de muestras de audio, pero procesa el bloque de audio usando dos o más funciones de ventana para resultar en dos o más bloques de audio de ventana. El tratamiento posterior de los bloques de audio de ventana permite la detección de los efectos de ventana en el espectro de audio del bloque de audio. Los valores de firma únicos o sustancialmente únicos para el bloque de audio se derivan a partir de los efectos de las dos o más funciones de ventana en el bloque de audio. Es decir, las técnicas de ejemplo descritas en

el presente documento permiten el cálculo o la determinación de las firmas de audio sin el uso de bloques de audio desplazados en tiempo. Por supuesto, puede ajustarse la selección de las funciones de ventana, como puede ser el tipo de transformaciones, sus parámetros, y/o las resoluciones usadas para determinar las firmas.

5 Como se describe en detalle a continuación, después de la aplicación de las funciones de ventana al bloque de muestras de audio, se generan los componentes de frecuencia de los bloques de audio de ventana transformando los bloques de audio de ventana del dominio de tiempo al dominio de frecuencia usando, por ejemplo, una transformada de Fourier discreta (DFT) o cualquier otra transformación adecuada (por ejemplo, la transformada de coseno discreta (DCT), la transformada de coseno discreta modificada (MDCT), la transformada de Haar, la transformada de Walsh, etc.), esté o no basada en la transformada de Fourier. La transformada puede usarse para analizar los componentes de frecuencia en los bloques de audio de ventana y para identificar la potencia espectral de cada componente de frecuencia. A continuación, las potencias espectrales pueden usarse para generar firmas digitales.

15 Pueden usarse otras técnicas después de la aplicación de las funciones de ventana a los bloques de audio. Por ejemplo, los bloques de audio de ventana pueden procesarse usando las transformadas de wavelet que transforman los datos de audio a partir del dominio de tiempo al dominio wavelet. En general, las transformadas de wavelet pueden usarse para descomponer los bloques o las tramas de datos (por ejemplo, las muestras de audio del dominio de tiempo) en múltiples sub-bandas, permitiendo de este modo que los conjuntos de datos se analicen a diferentes escalas y/o resoluciones. Separando los datos en múltiples sub-bandas, puede usarse una transformada wavelet para analizar cada intervalo de tiempo de los datos en una escala o resolución deseada.

25 Como alternativa, en lugar de aplicar las funciones de ventana en el dominio de tiempo a los bloques del dominio de tiempo de las muestras de audio, la visualización podría hacerse en el dominio de frecuencia, en el que una respuesta de frecuencia correspondiente a una ventana del dominio de tiempo puede convolucionarse con el espectro de frecuencia de un bloque de audio. Si se usa el procesamiento del dominio de frecuencia que incluye una convolución, puede realizarse una conversión del bloque de audio al dominio de frecuencia usando una transformada de Fourier, en la que los ajustes se hacen entre los bloques de audio para tener en cuenta la discontinuidad. Además, si el procesamiento y la aplicación de las ventanas se hacen en el dominio de frecuencia, puede seleccionarse una ventana del dominio de tiempo que tenga una característica de frecuencia con un número de elementos distinto de cero (por ejemplo, 3-5 elementos distintos de cero).

35 Las firmas monitorizadas pueden generarse usando las técnicas anteriores en un sitio de monitorización en base a los flujos de audio asociados a la información de los medios (por ejemplo, un flujo de audio monitorizado) que se consume por una audiencia o para que se exponga a una audiencia. Por ejemplo, una firma monitorizada puede generarse en base a la pista de audio de un programa de televisión o cualquier otro medio presentado en un sitio de monitorización. A continuación, la firma monitorizada puede comunicarse a una instalación de recopilación de datos central para la comparación de una o más firmas de referencia.

40 Las firmas de referencia se generan en un sitio de referencia y/o en una instalación de recopilación de datos central usando las técnicas anteriores en los flujos de audio asociados con la información de los medios conocida. La información de los medios conocida puede incluir medios que se difunden dentro de una región, medios que se reproducen dentro de un hogar, medios que se reciben a través de Internet, etc. Cada firma de referencia se almacena en una memoria con una información de identificación de medios tales como, por ejemplo, un título de una canción, un título de una película, etc. Cuando se recibe una firma monitorizada en la instalación de recopilación de datos central, la firma monitorizada se compara con uno o más firmas de referencia hasta que se encuentra un emparejamiento. A continuación, esta información de emparejamiento puede usarse para identificar la información de los medios (por ejemplo, el flujo de audio monitorizado) a partir del cual se ha generado la firma monitorizada. Por ejemplo, puede hacerse referencia a una tabla de consulta o a una base de datos para recuperar un título de medios, una identidad de programa, un número de episodio, etc. que corresponde a la información de los medios a partir de la que se ha generado la firma monitorizada.

55 En un ejemplo, las velocidades a las que se generan las firmas monitorizadas y las firmas de referencia pueden ser diferentes. Por ejemplo, para el procesamiento y otros asuntos, una firma monitorizada puede ser un 25% de la velocidad de datos de una firma de referencia. Por ejemplo, una firma de referencia de 48 bits puede generarse cada 0,032 segundos, lo que resulta en una velocidad de datos de referencia de $48 \text{ bits} \times 31,25 / \text{segundo}$ o 187,5 bytes/segundo. En tal disposición, una firma monitorizada de 48 bits puede generarse cada 0,128 segundos, lo que resulta en una velocidad de datos monitorizados de $48 \text{ bits} \times 7,8125 / \text{segundos}$ o 46,875 bytes/segundo. Por supuesto, en una disposición en la que las velocidades de datos de las firmas monitorizadas y de referencia son diferentes, debe tenerse en cuenta esta diferencia cuando se comparen las firmas monitorizadas con las firmas de referencia. Por ejemplo, si la velocidad de monitorización es el 25% de la velocidad de referencia, cada firma monitorizada consecutiva corresponderá a cada cuatro firmas de referencia.

65 Las figuras 1A y 1B ilustran unos sistemas 100 y 150 de identificación del flujo de audio de ejemplo para generar firmas espectrales digitales e identificar los flujos de audio. Los sistemas 100 y 150 de identificación del flujo de audio de ejemplo pueden implementarse como un sistema de identificación de información de difusión de televisión y

un sistema de identificación de información de difusión de radio, respectivamente. El sistema 100 de identificación de flujo de audio incluye un sitio 102 de monitorización (por ejemplo, un hogar monitorizado), un sitio 104 de referencia, y una instalación 106 de recopilación de datos central.

La monitorización de la información de difusión de televisión implica la generación de firmas monitorizadas en el sitio 102 de monitorización en base a los datos de audio de información de difusión de televisión y la comunicación de las firmas monitorizadas a la instalación 106 de recopilación de datos central a través de una red 108. Las firmas de referencia pueden generarse en el sitio 104 de referencia y también podrán comunicarse a la instalación 106 de recopilación de datos central a través de la red 108. El contenido de audio representado por una firma monitorizada que se genera en el sitio 102 de monitorización puede identificarse en la instalación 106 de recopilación de datos central comparando la firma monitorizada con una o más firmas de referencia hasta que se encuentra un emparejamiento. Como alternativa, las firmas monitorizadas pueden comunicarse desde el sitio 102 de monitorización al sitio 104 de referencia y compararse con una o más firmas de referencia en el sitio 104 de referencia. En otro ejemplo, las firmas de referencia podrán comunicarse al sitio 102 de monitorización y compararse con las firmas monitorizadas en el sitio 102 de monitorización.

El sitio 102 de monitorización puede ser, por ejemplo, un hogar para el que se monitoriza el consumo de medios de una audiencia. En general, el sitio 102 de monitorización puede incluir una pluralidad de dispositivos 110 de suministro de medios, una pluralidad de dispositivos 112 de presentación de medios y un generador 114 de firma que se usa para generar firmas monitorizadas asociadas con los medios presentados en el sitio 102 de monitorización.

La pluralidad de dispositivos 110 de suministro de medios puede incluir, por ejemplo, sintonizadores decodificadores (por ejemplo, sintonizadores de cable, sintonizadores de satélite, etc.), dispositivos PVR, reproductores de DVD, reproductores de CD, radios, etc. Algunos o todos los dispositivos 110 de suministro de medios, tales como, por ejemplo, los sintonizadores decodificadores pueden acoplarse de forma comunicativa a uno o más dispositivos 116 de recepción de información de difusión, que pueden incluir un cable, una antena parabólica, una antena, y/o cualquier otro dispositivo adecuado para recibir información de difusión. Los dispositivos 110 de suministro de medios pueden configurarse para reproducir la información de los medios (por ejemplo, información de audio, información de vídeo, páginas web, imágenes fijas, etc.) en base a, por ejemplo, la información de difusión y/o la información almacenada. La información de difusión pueden obtenerse a partir de los dispositivos 116 de recepción de información de difusión y la información almacenada puede obtenerse a partir de cualquier medio de almacenamiento de información (por ejemplo, un DVD, un CD, una cinta, etc.). Los dispositivos 110 de suministro de medios se acoplan de forma comunicativa a los dispositivos 112 de presentación de medios y configurables para comunicar información de los medios a los dispositivos 112 de presentación de medios para su presentación. Los dispositivos 112 de presentación de medios pueden incluir televisores que tengan un dispositivo de pantalla y/o un conjunto de altavoces por los que consumen los miembros de la audiencia, por ejemplo, información de televisión de difusión, música, películas, etc.

El generador 114 de firmas puede usarse para generar firmas digitales monitorizadas en base a la información de audio, como se describe en mayor detalle a continuación. En particular, en el sitio 102 de monitorización, el generador 114 de firmas puede configurarse para generar firmas monitorizadas en base a flujos de audio monitorizados que se reproducen por los dispositivos 110 de suministro de medios y/o presentados por los dispositivos 112 de presentación de medios. El generador 114 de firmas puede acoplarse de forma comunicativa a los dispositivos 110 de suministro de medios y/o a los dispositivos 112 de presentación de medios a través de una interfaz 118 de monitorización de audio. De esta manera, el generador 114 de firma puede obtener los flujos de audio asociados con la información de los medios que se reproduce por los dispositivos 110 de suministro de medios y/o presentada por los dispositivos 112 de presentación de medios. Adicional o alternativamente, el generador 114 de firma puede acoplarse de forma comunicativa a los micrófonos (no mostrados) que se colocan en la proximidad de los dispositivos 112 de presentación de medios para detectar los flujos de audio. El generador 114 de firmas puede acoplarse también de forma comunicativa a la instalación 106 de recopilación de datos central a través de la red 108.

La red 108 puede usarse para comunicar firmas (por ejemplo, firmas espectrales digitales), información de control, y/o información de configuración entre el sitio 102 de monitorización, el sitio 104 de referencia, y la instalación 106 de recopilación de datos central. Cualquier sistema de comunicación por cable o inalámbrico tal como, por ejemplo, una red de cable de banda ancha, una red DSL, una red telefónica móvil, una red de satélite, y/o cualquier otra red de comunicaciones pueden usarse para implementar la red 108.

Como se muestra en la figura 1A, el sitio 104 de referencia puede incluir una pluralidad de sintonizadores 120 de información de difusión, un generador 122 de firmas de referencia, un transmisor 124, una base de datos o la memoria 126, y unos dispositivos 128 de recepción de información de difusión. El generador 122 de firmas de referencia y el transmisor 124 pueden acoplarse de forma comunicativa a la memoria 126 para almacenar las firmas de referencia en la misma y/o para recuperar las firmas de referencia almacenadas de la misma.

Los sintonizadores 120 de información de difusión pueden acoplarse de forma comunicativa a los dispositivos 128 de recepción de información de difusión, que pueden incluir un cable, una antena, una antena parabólica, y/o cualquier otro dispositivo adecuado para recibir información de difusión. Cada uno de los sintonizadores 120 de información de difusión puede configurarse para sintonizar un canal de difusión específico. En general, el número de sintonizadores en el sitio 104 de referencia es igual al número de canales disponibles en una región de difusión específica. De esta manera, pueden generarse las firmas de referencia para toda la información de los medios transmitida a través de todos los canales en una región de difusión. La parte de audio de la información de medios sintonizada podría comunicarse desde los sintonizadores 120 de información de difusión al generador 122 de firmas de referencia.

El generador 122 de firmas de referencia puede configurarse para obtener la parte de audio de toda la información de los medios que está disponible en una región de difusión específica. A continuación, el generador 122 de firmas de referencia puede generar una pluralidad de firmas de referencia (usando, por ejemplo, el procesamiento descrito en mayor detalle a continuación) en base a la información de audio y almacenar las firmas de referencia en la memoria 126. Aunque el generador de firmas de referencia se muestra en la figura 1, puede usarse una pluralidad de generadores de firma de referencia en el sitio 104 de referencia. Por ejemplo, cada uno de la pluralidad de generadores de firma puede acoplarse de forma comunicativa a uno de los sintonizadores 120 de información de difusión respectivo.

El transmisor 124 puede acoplarse de forma comunicativa a la memoria 126 y configurarse para recuperar las firmas de la misma y comunicar las firmas de referencia a la instalación 106 de recopilación de datos central a través de la red 108.

La instalación 106 de recopilación de datos central puede configurarse para comparar las firmas monitorizadas recibidas desde el sitio 102 de monitorización con las firmas de referencia recibidas desde el sitio 104 de referencia. Además, la instalación 106 de recopilación de datos central puede configurarse para identificar los flujos de audio monitorizados, emparejando las firmas monitorizadas con las firmas de referencia y usando la información de emparejamiento para recuperar la información de identificación del programa de televisión (por ejemplo, el título del programa, la hora de difusión, el canal de difusión, etc.) a partir de una base de datos. La instalación 106 de recopilación de datos central incluye un receptor 130, un analizador 132 de firma y una memoria 134, todos los cuales están acoplados de forma comunicativa como se muestra.

El receptor 130 puede configurarse para recibir las firmas monitorizadas y las firmas de referencia a través de la red 108. El receptor 130 está acoplado de forma comunicativa a la memoria 134 y se configura para almacenar las firmas monitorizadas y las firmas de referencia en la misma.

El analizador 132 de firma puede usarse para comparar las firmas de referencia con las firmas monitorizadas. El analizador 132 de firma está acoplado de forma comunicativa con la memoria 134 y configurado para recuperar las firmas monitorizadas y las firmas de referencia de la misma. El analizador 132 de firma puede configurarse para recuperar las firmas de referencia y las firmas monitorizadas desde la memoria 134 y comparar las firmas monitorizadas con las firmas de referencia hasta que se encuentre un emparejamiento. La memoria 134 puede implementarse usando cualquier medio de almacenamiento de información accesible por máquina tal como, por ejemplo, uno o más discos duros, uno o más dispositivos de almacenamiento óptico, etc.

Aunque el analizador 132 de firma se localiza en la instalación 106 de recopilación de datos central en la figura 1A, el analizador 132 de firma puede localizarse en el sitio 104 de referencia. En tal configuración, las firmas monitorizadas pueden comunicarse desde el sitio 102 de monitorización al sitio 104 de referencia a través de la red 108. Como alternativa, la memoria 134 puede localizarse en el sitio 102 de monitorización y las firmas de referencia pueden añadirse periódicamente a la memoria 134 a través de la red 108 mediante el transmisor 124. Además, aunque el analizador 132 de firma se muestra como un dispositivo separado de los generadores 114 y 122 de firma, el analizador 132 de firma puede integrarse con el generador 122 de firmas de referencia y/o el generador 114 de firmas. Aún más, aunque la figura 1 representa un único sitio de monitorización (es decir, el sitio 102 de monitorización) y un único sitio de referencia (es decir, el sitio 104 de referencia), pueden acoplarse múltiples sitios a través de la red 108 a la instalación 106 de recopilación de datos central.

El sistema 150 de identificación de flujo de audio de la figura 1B puede configurarse para monitorizar e identificar los flujos de audio asociados con la información de difusión de radio, o cualquier otro medio. En general, el sistema 150 de identificación de flujo de audio se usa para monitorizar el contenido que se difunde mediante una pluralidad de estaciones de radio en una región de difusión específica. A diferencia del sistema 100 de identificación de flujo de audio usado para monitorizar el contenido de televisión consumido por una audiencia, el sistema 150 de identificación de flujo de audio puede usarse para monitorizar música, canciones, etc. que se difunden dentro de una región de difusión y el número de veces que se difunden. Este tipo monitorizado de medios puede usarse para determinar el pago de derechos, el uso adecuado de los derechos de autor, etc. asociados con cada composición de audio. El sistema 150 de identificación de flujo de audio incluye un sitio 152 de monitorización, una instalación 154 de recopilación de datos central y la red 108.

El sitio 152 de monitorización está configurado para recibir toda la información de difusión de radio que está disponible en una región de difusión específica, y para generar firmas monitorizadas en base a la información de difusión de radio. El sitio 152 de monitorización incluye la pluralidad de sintonizadores 120 de información de difusión, el transmisor 124, la memoria 126, y los dispositivos 128 de recepción de información de difusión, todos los cuales se han descrito anteriormente en relación con la figura 1A. Además, el sitio 152 de monitorización incluye un generador 156 de firma. Cuando se usan en el sistema 150 de identificación de flujo de audio, los dispositivos 128 de recepción de información de difusión están configurados para recibir información de difusión de radio y los sintonizadores 120 de información de difusión están configurados para sintonizar las estaciones de difusión de radio. El número de sintonizadores 120 de información de difusión en el sitio 152 de monitorización puede ser igual al número de estaciones de difusión de radio en una región de difusión específica.

El generador 156 de firma está configurado para recibir la sintonización de la información de audio de cada uno de los sintonizadores 120 de información de difusión y generar firmas monitorizadas para el mismo. Aunque se muestra un generador de firma (es decir, el generador 156 de firma), el sitio 152 de monitorización puede incluir múltiples generadores de firma, cada uno de los cuales puede acoplarse de forma comunicativa a uno de los sintonizadores 120 de información de difusión. El generador 156 de firma puede almacenar las firmas monitorizadas en la memoria 126. El transmisor 124 puede recuperar las firmas monitorizadas de la memoria 126 y comunicarlás a la instalación 154 de recopilación de datos central a través de la red 108.

La instalación 154 de recopilación de datos central está configurada para recibir firmas monitorizadas desde el sitio 152 de monitorización, generar firmas de referencia en base a los flujos de audio de referencia y comparar las firmas monitorizadas con las firmas de referencia. La instalación 154 de recopilación de datos central incluye el receptor 130, el analizador 132 de firmas y la memoria 134, todos los cuales se han descrito con mayor detalle anteriormente en relación con la figura 1A. Además, la instalación 154 de recopilación de datos central incluye un generador 158 de firmas de referencia.

El generador 158 de firmas de referencia está configurado para generar firmas de referencia en base a los flujos de audio de referencia. Los flujos de audio de referencia pueden almacenarse en cualquier tipo de medio accesible por máquina tal como, por ejemplo, un CD, un DVD, una cinta de audio digital (DAT), etc. En general, los artistas y/o las empresas productoras discográficas envían sus trabajos de audio (es decir, la música, las canciones, etc.) a la instalación 154 de recopilación de datos central para añadirse a una biblioteca de referencia. El generador 158 de firmas de referencia puede leer datos de audio del medio accesible por máquina y generar una pluralidad de firmas de referencia en base a cada trabajo de audio (por ejemplo, el audio 300 capturado de la figura 3). A continuación, el generador 158 de firmas de referencia puede almacenar las firmas de referencia en la memoria 134 para su posterior recuperación mediante el analizador 132 de firma. La información de identificación (por ejemplo, el título de la canción, el nombre del artista, el número de pista, etc.) asociada con cada flujo de audio de referencia puede almacenarse en una base de datos y puede indexarse en base a las firmas de referencia. De esta manera, la instalación 154 de recopilación de datos central incluye una base de datos de firmas de referencia y la información de identificación correspondiente a todos los títulos de las canciones conocidas y disponibles.

El receptor 130 está configurado para recibir firmas monitorizadas desde la red 108 y almacenar las firmas monitorizadas en la memoria 134. Las firmas monitorizadas y las firmas de referencia se recuperan de la memoria 134 mediante el analizador 132 de firma para su uso en la identificación de los flujos de audio monitorizados difundidos dentro de una región de difusión. El analizador 132 de firma puede identificar los flujos de audio monitorizados mediante un primer emparejamiento de una firma monitorizada con una firma de referencia. La información de emparejamiento y/o la firma de referencia de emparejamiento se usan a continuación para recuperar la información de identificación (por ejemplo, un título de una canción, una pista de una canción, un artista, etc.) a partir de una base de datos almacenada en la memoria 134.

Aunque se muestra un sitio de monitorización (por ejemplo, el sitio 152 de monitorización) en la figura 1B, pueden acoplarse múltiples sitios de monitorización de forma comunicativa a la red 108 y configurarse para generar firmas monitorizadas. En particular, cada sitio de monitorización puede localizarse en una región de difusión respectiva y configurarse para monitorizar el contenido de las estaciones de difusión dentro de una región de difusión respectiva.

La figura 2 es un diagrama de flujo que representa un proceso 200 de generación de firmas de ejemplo. Como se muestra en la figura 2, el proceso 200 de generación de firmas captura primero un bloque de audio que se caracteriza por una firma (bloque 202). Una representación gráfica de un dominio de tiempo de audio de ejemplo que puede capturarse se muestra en la figura 3 con el número de referencia 300. El audio puede capturarse desde una fuente de audio a través de, por ejemplo, una conexión cableada a una fuente de audio o a través de una conexión inalámbrica, tal como un sensor de audio, a una fuente de audio. Si la fuente de audio es analógica, la captura incluye un muestreo de la fuente de audio analógica usando, por ejemplo, un convertidor analógico a digital. En un ejemplo, la fuente de audio puede muestrearse a una velocidad de 8 kilohercios (kHz), que se denomina como una velocidad (F_s) de muestreo. Esto significa que el audio analógico está representado por las muestras digitales de la misma que se toman a la velocidad de ocho mil muestras por segundo, o cada 125 microsegundos (μs). Cada una de las muestras de audio puede representarse mediante monoaural, 16 bits de resolución.

En un ejemplo, se captura un bloque de audio correspondiente a 8192 muestras para su procesamiento. En la frecuencia de muestreo anterior de 8 kHz, esto corresponde a 1,024 segundos de audio. Sin embargo, esto es simplemente un ejemplo, y el número de muestras que se recopilan puede corresponder a segmentos de audio que van desde aproximadamente 1 segundo a 2 segundos de duración. Genéricamente, en el presente documento el número de muestras capturadas en un bloque de audio se denomina con la variable N . Por lo tanto, en el ejemplo anterior, $N = 8192$ y el intervalo de tiempo de audio capturado corresponde a $t \dots t + N/F_s$. Se muestra una representación de un bloque de audio en la figura 4 con el número de referencia 402, en el que, para fines de ejemplo, el bloque de audio corresponde a una senoide.

Después de que se haya capturado el bloque de audio (bloque 202), el proceso 200 aplica la primera función de ventana, denominada como W_1 (bloque 204A), al bloque de audio para producir un primer bloque de audio de ventana. Además, el proceso 200 visualiza el bloque de audio usando una segunda función de ventana, denominada como W_2 (bloque 204B) para producir un segundo bloque de audio de ventana. Por ejemplo, la ventana puede ser una función en forma de campana o de Gauss tal como la mostrada con el número de referencia 502 en la figura 5, en la que los extremos alto y bajo de W_1 502 tienen un valor cero y el centro de la ventana 502 tiene un valor de uno. En un ejemplo, la visualización es una multiplicación a manera de muestra entre los valores de la función de ventana y las muestras respectivas del bloque de audio. Por ejemplo, visualizando el bloque 402 de audio con la ventana 502, resulta en un bloque 602 de audio de ventana, como se muestra en la figura 6, en la que la amplitud del bloque 602 de audio de ventana es cero en los extremos de la ventana 502 y es la misma amplitud que el bloque 402 de audio en el centro del bloque 602 de audio de ventana.

Como alternativa, en lugar de aplicar las funciones de ventana en el dominio de tiempo usando la multiplicación a manera de muestra de las funciones de ventana al bloque de audio, la visualización podría hacerse en el dominio de frecuencia, en el que una respuesta de frecuencia correspondiente a una ventana del dominio de tiempo puede convolucionarse con el espectro de frecuencia de un bloque de audio. Como se ha señalado anteriormente, si se usa un procesamiento de dominio de frecuencia que incluya una convolución, puede realizarse una conversión del bloque de audio al dominio de frecuencia usando una transformada de Fourier, en la que los ajustes se hacen entre los bloques de audio para tener en cuenta la discontinuidad. Además, si el procesamiento y la aplicación de las ventanas se hacen en el dominio de frecuencia, puede seleccionarse una ventana del dominio de tiempo que tenga una característica de frecuencia con un número de elementos distinto de cero (por ejemplo, 3-5 elementos distintos de cero).

Las ventanas seleccionadas para W_1 y W_2 pueden ser complementarias por naturaleza. Por ejemplo, si se selecciona la ventana 502 mostrada en la figura 5 para W_1 , puede seleccionarse la ventana 702 de la figura 7 para W_2 . Como se muestra en la figura 7, la ventana 702 es una versión invertida de la ventana W_1 , concretamente $W_2(k) = 1 - W_1(k)$, en la que k es un índice de muestra en el dominio de la ventana. La ventana W_2 se aproxima a un valor unitario en los extremos alto y bajo de la ventana 702 y tiene un valor cero en el centro de la ventana 702. Por lo tanto, cuando la ventana 702 se aplica al bloque 402 de audio, resulta un bloque 802 de audio de ventana, como se muestra en la figura 8. Como se muestra en la figura 8, el bloque 802 de audio de ventana tiene una amplitud cero en el centro de la misma, pero tiene una amplitud que se empareja sustancialmente con la amplitud del bloque 402 de audio en los extremos del bloque 802 de audio de ventana.

Como se muestra en las figuras 9 y 10, las ventanas 502 y 702 tienen unas respuestas 902 y 1002 de frecuencia respectivas. Por lo tanto, la aplicación de las ventanas 502 y 702 al bloque de audio (por ejemplo, el bloque 402 de audio), afecta al espectro del bloque de audio. Como se explica a continuación, estos son los diferentes efectos de diferentes ventanas en el bloque de audio que se examinan para determinar las firmas representativas del bloque de audio.

Mientras que las ventanas 502, 702 seleccionadas para la descripción anterior se asemejan a una ventana de Hann y a una ventana de Hann invertida, respectivamente, pueden usarse otras formas de ventana. Por ejemplo, como se muestra en las figuras 11 y 12, pueden seleccionarse dos ventanas asimétricas 1102, 1202, en la que una primera ventana 1102 ocupa una mitad superior del espacio de ventana y en la que una segunda ventana de 1202 ocupa una mitad inferior del espacio de ventana. Las respuestas de frecuencia de las ventanas 1102, 1202 asimétricas son idénticas, como se muestra en las figuras 11 y 12 con los números de referencia 1104 y 1204, pero debido a que las ventanas funcionan en su mayoría en partes distintas de un bloque de audio, los resultados de la visualización tienen diferentes características espectrales para las señales de audio que no son sinusoidales.

Mientras que ciertos ejemplos de formas de ventana se describen en el presente documento, pueden usarse otras formas de ventana. Por ejemplo, pueden seleccionarse formas de ventana arbitrariamente, tanto para la primera ventana como para la segunda ventana (por ejemplo, W_1 y W_2), en la que la selección se realiza a partir de un conjunto de funciones de ventana. Por supuesto, pueden usarse diferentes ventanas en diferentes momentos, proporcionados los sitios de monitor y de referencia que usan los mismos tiempos. Además, pueden usarse más de dos ventanas.

Volviendo a la figura 2, después de que se complete el windowing (bloques 204A y 204B), los bloques de audio de ventana se transforman respectivamente (bloques 206A y 206B). En un ejemplo, la transformación puede ser una

transformación del dominio de tiempo al dominio de frecuencia. Por ejemplo, las N muestras de audio capturado que se han visualizado se pueden convertir en un espectro de audio que se representa por $N/2$ coeficientes de la DFT complejos. Como alternativa, pueden usarse cualquier transformación adecuada, tal como las transformadas wavelet, DCT, TCMD, las transformadas Haar, las transformadas Walsh, etc.

Después de que las transformaciones se hayan completado (bloques 206A y 206B), el proceso 200 caracteriza los resultados de cada transformación (208A y 208B del bloque). Por ejemplo, el proceso puede determinar la energía en cada una de las $K + 1$ bandas diferentes de cada uno de los resultados de la transformación. Es decir, los resultados de la transformación en el bloque de audio de ventana resultante a partir del uso de la ventana W_1 (bloque 206A) pueden dividirse en, por ejemplo, 16 bandas diferentes y puede determinarse la energía en cada una de las 16 bandas diferentes. Esto puede representarse por $E_j(w_1)$, en la que j varía de 0 a 15, y w_1 indica que la energía está asociada con el espectro resultante de la aplicación de la ventana W_1 al audio de muestreo (es decir, al bloque de audio). Del mismo modo, los resultados de la transformación en el bloque de ventana resultante a partir del uso de la ventana W_2 (bloque 206B) pueden dividirse en, por ejemplo, 16 bandas diferentes, la energía de las cuales puede determinarse y se representa como $E_j(w_2)$, en la que j varía de 0 a 15, y w_2 indica que la energía está asociada con el espectro resultante de la aplicación de la ventana W_2 . Como alternativa, pueden usarse diferentes características espectrales distintas de la energía para caracterizar los resultados. Por ejemplo, puede usarse la planitud espectral de la distribución de energía.

Después de que se haya caracterizado cada conjunto de resultados de transformación (bloques 208A y 208B), el proceso 210 compara los resultados de las caracterizaciones. Por ejemplo, los resultados de las caracterizaciones de cada banda pueden restarse entre sí. En un ejemplo, puede calcularse un valor intermedio como $d_j = E_j(w_2) - E_j(w_1)$, en la que j varía de 0 a K . Siguiendo con el ejemplo específico anterior en el que $K = 15$, puede calcularse un valor d_j intermedio, en el que $d_j = E_j(w_2) - E_j(w_1)$, y j varía de 0 a 15. Por lo tanto, tal comparación resulta en 16 valores intermedios diferentes (por ejemplo, $d_0, d_1, d_2, \dots, d_{15}$), en el que cada valor intermedio es la diferencia en las características en, por ejemplo, las bandas de frecuencia similares del espectro resultante de las transformaciones de los bloques de audio de ventana.

Después de que se haya calculado el valor intermedio para representar la comparación de las caracterizaciones (bloque 210), el proceso 200 determina los bits de la firma en base a las comparaciones (bloque 212). Por ejemplo, puede asignarse a un bit S_j de firma un valor de 1 si el valor intermedio $d_j > 0$, y puede asignarse un valor de 0 en caso contrario, en el que j varía de 0 a K . Más específicamente, como se indica en el ejemplo anterior $K = 15$ y, por lo tanto, habrá 16 comparaciones de los valores intermedios al valor de 0 y, en base a estas comparaciones, se generará una firma de 16 bits para representar el bloque de audio, que se ha capturado en el bloque 202 de la figura 2. Después de que se haya determinado la firma (bloque 212), el proceso 200 se repite (bloque 214) y se captura audio adicional (bloque 202) para desarrollar firmas adicionales.

Aunque lo anterior describe la selección de una primera ventana (W_1) y una segunda ventana (W_2) y que todos los bits de firma de un bloque de audio capturado se determinan usando las ventanas seleccionadas, son posibles otras configuraciones. Por ejemplo, algunos bits de una firma que representan el bloque de audio capturado pueden determinarse usando un primer par de ventanas (por ejemplo, W_1 y W_2), mientras que los otros bits de la firma pueden determinarse usando un par diferente de ventanas (por ejemplo, W_3 y W_4). Además, puede usarse un tercer par de ventanas (por ejemplo, W_1 y W_3) para determinar los bits de firma adicionales. En algunos casos, podría seleccionarse un único par de ventanas de una manera predeterminada o arbitraria para determinar el valor de cada bit de firma, siempre y cuando se hayan seleccionado los mismos pares de ventanas para operar en los mismos bloques de ventana en el sitio de referencia.

Lo anterior ha descrito las técnicas de firmas que se pueden realizar para determinar las firmas representativas de una parte del audio capturado. La figura 13 muestra un ejemplo de un proceso 1300 de emparejamiento de firmas que puede realizarse para comparar las firmas de referencia (es decir, las firmas determinadas en un sitio(s) de referencia) con las firmas monitorizadas (es decir, las firmas determinadas en un sitio de monitorización). El objetivo último del emparejamiento de firmas es encontrar el emparejamiento más cercano entre una firma de audio de consulta (por ejemplo, audio monitorizado) y las firmas en una base de datos (por ejemplo, las firmas tomadas en base a un audio de referencia). La comparación puede realizarse en un sitio de referencia, un sitio de monitorización, o cualquier otro sitio de procesamiento de datos que tenga acceso a las firmas monitorizadas y a una base de datos que contenga las firmas de referencia.

Volviendo ahora en detalle al método del ejemplo de la figura 13, el proceso 1300 de ejemplo, consiste en obtener una firma monitorizada y su sincronización (bloque 1302). Como se muestra en la figura 14, una recopilación de firmas puede incluir un número de firmas monitorizadas, tres de las cuales se muestran en la figura 14 en los números de referencia 802, 804 y 806. Cada una de las firmas está representada por una sigma (σ). Cada una de las firmas 1402, 1404, 1406 monitorizadas pueden incluir información 1408, 1410, 1412, de sincronismo ya que la información de sincronismo está implícita o explícita.

A continuación, se hace una consulta a una base de datos que contiene firmas de referencia (bloque 1304) para identificar la firma en la base de datos que tiene el emparejamiento más cercano. En una implementación, se toma la

medida de similitud (proximidad) entre las firmas para que sea una distancia de Hamming, concretamente, el número de posición en el que difieren los valores de consulta y las cadenas de bits de referencia. En la figura 14, se muestra una base de datos de firmas e información de sincronismo con el número de referencia 1416. Por supuesto, la base de datos 1406 puede incluir cualquier número de diferentes firmas de diferentes presentaciones de medios. A continuación, se hace una asociación entre el programa asociado con la firma de referencia emparejada y la firma desconocida (bloque 1306).

Opcionalmente, a continuación, el procedimiento 1300 puede establecer una compensación entre la firma monitorizada y la firma de referencia (bloque 1308). Se requiere el valor de compensación con el fin de hacer una mejor determinación, más segura, que si un bloque de firmas de consulta se empareja bien con la firma de referencia. Normalmente los valores de compensación para todas las firmas en un bloque de consulta corto se mantendrán casi constantes en relación con las firmas de referencia respectivas debido a la continuidad de la monitorización (visualización).

En los casos en los que todos los descriptores de más de una firma de referencia se asocian con una distancia de Hamming por debajo del umbral de la distancia de Hamming predeterminada, más de una firma monitorizada puede necesitar emparejarse con las firmas de referencia respectivas de los posibles flujos de audio de referencia emparejados. Será relativamente poco probable que todas las firmas monitorizadas generadas en base al flujo de audio monitorizado se emparejen con todas las firmas de referencia de más de un flujo de audio de referencia, y, por lo tanto, puede evitarse emparejar erróneamente más de un flujo de audio de referencia a un flujo de audio monitorizado.

Los métodos de ejemplo, los procesos y/o las técnicas descritas anteriormente pueden implementarse mediante hardware, software, y/o cualquiera de sus combinaciones. Más específicamente, los métodos de ejemplo pueden ejecutarse en un hardware definido por los diagramas de bloques de las figuras 15 y 16. Los métodos de ejemplo, los procesos y/o las técnicas pueden implementarse también mediante software ejecutado en un sistema de procesador tal como, por ejemplo, el sistema 1610 de procesador de la figura 16.

La figura 15 es un diagrama de bloques de un sistema 1500 de generación de firma de ejemplo para generar firmas espectrales digitales. En particular, el sistema 1500 de generación de firma de ejemplo puede usarse para generar firmas monitorizadas y/o firmas de referencia en base a una visualización, transformación, caracterización, y comparación, de un bloque de audio, como se ha descrito anteriormente. Por ejemplo, el sistema 1500 de generación de firma de ejemplo puede usarse para implementar los generadores 114 y 122 de firma de la figura 1A o los generadores 156 y 158 de firma de la figura 1B. Además, el sistema 1500 de generación de firma de ejemplo puede usarse para implementar los métodos de ejemplo de la figura 2.

Como se muestra en la figura 15, el sistema 1500 de generación de firma de ejemplo incluye un generador 1502 de muestras, un dispositivo 1503 de sincronismo, un generador 1504 de tiempo de referencia, un windower 1506, un transformador 1508, un determinante 1510 de características, un comparador 1512, un determinador 1514 de firma, un almacenamiento 1516 y una interfaz 1518 de comunicación de datos, todos los cuales puede estar acoplados de forma comunicativa tal como se muestra. El sistema 1500 de generación de firma de ejemplo puede configurarse para obtener un flujo de audio de ejemplo, para adquirir una pluralidad de muestras de audio a partir del flujo de audio de ejemplo para formar un bloque de audio y a partir de ese único bloque de audio, generar una firma representativa del mismo.

El generador 1502 de muestras puede configurarse para obtener el flujo de audio de ejemplo, tal como un flujo resultante en el audio 300 capturado de la figura 3. El flujo 300 puede ser cualquier flujo de audio analógico o digital. Si el flujo de audio de ejemplo es un flujo de audio analógico, el generador 1502 de muestras puede implementarse usando un convertidor analógico a digital. Si el flujo de audio de ejemplo es un flujo de audio digital, el generador 1502 de muestras puede implementarse usando un procesador de señal digital. Además, el generador 1502 de muestras puede configurarse para adquirir y/o extraer muestras de audio en cualquier frecuencia F_s de muestreo deseada. Por ejemplo, como se ha descrito anteriormente, el generador de muestras puede configurarse para adquirir N muestras a 8 kHz y puede usar 16 bits para representar cada muestra. En tal disposición, N puede ser cualquier número de muestras tales como, por ejemplo, $N = 8192$. El generador 1502 de muestras también podría notificar al generador 1504 de tiempo de referencia cuando comienza un proceso de adquisición de muestras de audio. El generador 1502 de muestras comunica las muestras al windower 1506.

El dispositivo 1503 de sincronización puede configurarse para generar datos de tiempo y/o información de marca de tiempo y puede implementarse mediante un reloj, un temporizador, un contador, y/o cualquier otro dispositivo adecuado. El dispositivo 1503 de sincronización puede acoplarse de forma comunicativa al generador 1504 de tiempo de referencia y puede configurarse para comunicar datos y/o marcas de tiempo al generador 1504 de tiempo de referencia. El dispositivo 1503 de sincronización también puede acoplarse de forma comunicativa al generador 1502 de muestras y puede declarar una señal de arranque o una interrupción para indicar al generador 1502 de muestras que empiece a recopilar o a adquirir los datos de muestra de audio. En un ejemplo, el dispositivo 1503 de sincronización puede implementarse mediante un reloj en tiempo real que tenga un período de 24 horas que controle el tiempo con una resolución de milisegundos. En este caso, el dispositivo 1503 de sincronización puede

configurarse para reiniciarse a cero a medianoche y controlar el tiempo en milisegundos con respecto a la medianoche. Sin embargo, en general, las marcas de tiempo pueden representar un año completo, un mes, un día, una hora, un minuto, un segundo de información como un número de segundos transcurridos desde un momento predeterminado en el pasado, tal como las 00:00 AM, del 1 de enero del 2005. Puede añadirse una resolución por debajo del segundo derivando a partir de la velocidad de adquisición determinística de las firmas de audio recopiladas.

El generador 1504 de tiempo de referencia puede inicializar un tiempo t_0 de referencia cuando se recibe una notificación desde el generador 1502 de muestras. El tiempo t_0 de referencia puede usarse para indicar el tiempo en un flujo de audio en el que se genera una firma. En particular, el generador 1504 de tiempo de referencia puede configurarse para leer los datos de tiempo y/o un valor de marca de tiempo desde el dispositivo 1503 de sincronización cuando se notifique el inicio de un proceso de adquisición de muestras mediante el generador 1502 de muestras. A continuación, el generador 1504 de tiempo de referencia puede almacenar el valor de la marca de tiempo como el tiempo t_0 de referencia.

El windower 1506 aplica, por ejemplo, dos ventanas a la salida del bloque de audio desde el generador 1502 de muestras. Por lo tanto, los resultados del windower 1506 son dos bloques de audio de ventana. Como se ha descrito anteriormente, las ventanas pueden ser cualquier conjunto de ventanas. Sin embargo, pueden preferirse las ventanas incluidas porque garantizarían fácilmente que, en promedio, la energía de ambos valores es la misma, lo que conduce a la distribución de bits equi-probable.

El transformador 1508 puede configurarse para realizar un punto N de la DFT en cada uno de los bloques de audio de ventana, en el que N es el número de muestras obtenidas por el generador 1502 de muestras. Por ejemplo, si el generador de muestras obtiene 8192 muestras, el transformador producirá un espectro a partir de las muestras en el que el espectro se representa por 4096 coeficientes de Fourier de valores complejos.

El determinante 1510 de características puede configurarse para identificar varias bandas de frecuencia (por ejemplo, 16 bandas) dentro de la caracterización del espectro de la DFT generada por el transformador 1508. Las bandas seleccionadas pueden, pero preferentemente no, solaparse entre sí. Las bandas pueden seleccionarse de acuerdo con cualquier técnica. Por supuesto, pueden seleccionarse cualquier número de bandas adecuadas (por ejemplo, 48). A continuación, el determinante 1510 de características determina una característica en cada una de las bandas. Por ejemplo, el determinante 1510 de características puede determinar la energía en cada banda. Por lo tanto, los resultados del determinante 1510 de características son dos conjuntos de características por cada una de, por ejemplo, las 16 bandas. Por ejemplo, si se seleccionan 16 bandas, la salida del determinante 1510 de características sería de 32 medidas de energía, una por cada una de las bandas en cada una de las DFT. Las características pueden representarse por $E_j(w1)$ y $E_j(w2)$, en la que j varía de 0 a K (por ejemplo, de 0 a 15), y w1 y w2 representan la ventana 1 y la ventana 2, respectivamente.

El comparador 1512 compara las características de las bandas respectivas para determinar los valores intermedios. Por ejemplo, el comparador 1512 puede generar valores intermedios de acuerdo con $d_j = E_j(w2) - E_j(w1)$, de manera que las energías en las bandas respectivas de la DFT se restan entre sí.

El determinador 1514 de firma opera en los valores resultantes del comparador 1512 para producir un bit de firma para cada uno de los valores intermedios. Esta operación puede ser muy similar o idéntica al proceso 212 descrito anteriormente junto con la figura 2. Es decir, los valores de bit de firma pueden basarse en una comparación del valor intermedio a cero. Los bits de firma son la salida a la memoria 1516.

El almacenamiento puede ser cualquier medio adecuado para alojar el almacenamiento de la firma. Por ejemplo, el almacenamiento 1516 puede ser una memoria tal como una memoria de acceso aleatorio (RAM), una memoria flash, o similares. Adicional o alternativamente, el almacenamiento 1516 puede ser una memoria masiva, tal como un disco duro, un medio de almacenamiento óptico, una unidad de cinta, o similares.

El almacenamiento 1516 está acoplado a la interfaz 1518 de comunicación de datos. Por ejemplo, si el sistema 1500 está en un sitio de monitorización (por ejemplo, el hogar de una persona), la información de firma en el almacenamiento 1516 podría comunicarse a una instalación de recopilación, un sitio de referencia, o similares, usando la interfaz 1518 de comunicación de datos.

La figura 16 es un diagrama de bloques de un sistema 1600 de comparación de firma de ejemplo para comparar firmas espectrales digitales. En particular, el sistema 1600 de comparación de firma de ejemplo puede usarse para comparar firmas monitorizadas con firmas de referencia. Por ejemplo, el sistema 1600 de comparación de firma de ejemplo puede usarse para implementar el analizador 132 de firma de la figura 1A para comparar firmas monitorizadas con firmas de referencia. Además, el sistema 1600 de comparación de firma de ejemplo puede usarse para implementar el proceso de ejemplo de la figura 13.

El sistema 1600 de comparación de firma de ejemplo incluye un receptor 1602 de firma monitorizada, un receptor 1604 de firma de referencia, un comparador 1606, un filtro 1608 de distancia de Hamming, un identificador 1610 de

medios, y una interfaz 1612 de tabla de consulta de identificación de medios, todos los cuales pueden acoplarse de forma comunicativa como se muestra.

El receptor 1602 de firma monitorizada puede configurarse para obtener las firmas monitorizadas a través de la red 108 (figura 1) y comunicar las firmas monitorizadas al comparador de 1606. El receptor 1604 de firma de referencia puede configurarse para obtener las firmas de referencia de la memoria 134 (figuras 1A y 1B) y comunicar las firmas de referencia al comparador 1606.

El comparador de 1606 y el filtro 1608 de distancia de Hamming pueden configurarse para comparar firmas de referencia con las firmas monitorizadas usando las distancias de Hamming. En particular, el comparador 1606 puede configurarse para comparar los descriptores de firmas monitorizadas con descriptores de una pluralidad de firmas de referencia y para generar los valores de distancia de Hamming para cada comparación. A continuación, el filtro 1608 de distancia de Hamming puede obtener los valores de la distancia de Hamming del comparador 1606 y filtrar las firmas de referencia no emparejadas en base a los valores de la distancia de Hamming.

Después de que se encuentre una firma de referencia emparejada, el identificador 1610 de medios puede obtener la firma de referencia emparejada y en cooperación con la interfaz 1612 de tabla de consulta de identificación de medios puede identificar la información de los medios asociados con un flujo de audio no identificado (por ejemplo, el flujo 300 de audio monitorizado de ejemplo de la figura 3). Por ejemplo, la interfaz 1612 de tabla de consulta de identificación de medios puede acoplarse de forma comunicativa a una tabla de consulta de identificación de medios o a una base de datos que se usa para una referencia cruzada de la información de identificación de los medios (por ejemplo, el título de una película, el título de un programa, el título de una canción, el nombre del artista, el número del episodio, etc.) en base a las firmas de referencia. De esta manera, el identificador 1610 de medios puede recuperar la información de identificación de los medios a partir de la base de datos de identificación de los medios en base a las firmas de referencia emparejadas.

La figura 17 es un diagrama de bloques de un sistema 1710 de procesador de ejemplo que puede usarse para implementar el aparato y los métodos descritos en el presente documento. Como se muestra en la figura 17, el sistema 1710 de procesador incluye un procesador 1712 que está acoplado a un bus de interconexión o a la red 1714. El procesador 1712 incluye un conjunto de registros o un espacio 116 de registros, que se representa en la figura 17 como que está enteramente en el chip, pero que como alternativa podría estar localizado total o parcialmente fuera del chip y directamente acoplado al procesador 1712 a través de conexiones eléctricas dedicadas y/o a través de la red de interconexión o un bus 1714. El procesador 1712 puede ser cualquier procesador adecuado, una unidad de procesamiento o un microprocesador. Aunque no se muestra en la figura 17, el sistema 1710 puede ser un sistema multiprocesador y, por lo tanto, puede incluir uno o más procesadores adicionales que son idénticos o similares al procesador 1712 y que están acoplados de forma comunicativa al bus de interconexión o a la red 1714.

El procesador 1712 de la figura 17 está acoplado a un chipset 1718, que incluye un controlador 1720 de memoria y un controlador 1722 de entrada/salida (E/S). Como es bien sabido, un chipset proporciona normalmente funciones de E/S y de gestión de memoria, así como una pluralidad de registros de propósito general y/o de propósito especial, temporizadores, etc. que son accesibles o usados por uno o más procesadores acoplados al chipset. El controlador 1720 de memoria realiza funciones que permiten que el procesador 1712 (o procesadores si hay múltiples procesadores) acceda a una memoria 1724 del sistema y a una memoria 1725 de almacenamiento masivo.

La memoria 1724 del sistema puede incluir cualquier tipo deseado de memoria volátil y/o no volátil, tal como, por ejemplo, una memoria de acceso aleatorio estática (SRAM), una memoria de acceso aleatorio dinámica (DRAM), una memoria flash, una memoria de solo lectura (ROM), etc. La memoria 1725 de almacenamiento masivo puede incluir cualquier tipo deseado de dispositivo de almacenamiento masivo que incluya unidades de disco duro, unidades ópticas, dispositivos de almacenamiento de cinta, etc.

El controlador 1722 de E/S realiza funciones que permiten al procesador 1712 comunicarse con los dispositivos 1726 y 1728 de entrada/salida (E/S) periféricos a través de un bus 1730 de E/S. Los dispositivos 1726 y 1728 de E/S pueden ser de cualquier tipo deseado de dispositivo de E/S, tal como, por ejemplo, un teclado, una pantalla o monitor de vídeo, un ratón, etc. Mientras que el controlador 1720 de memoria y el controlador 1722 de E/S se representan en la figura 17 como bloques funcionales independientes dentro del chipset 1718, las funciones realizadas por estos bloques pueden integrarse dentro de un único circuito semiconductor o pueden implementarse usando dos o más circuitos integrados separados.

Los métodos descritos en el presente documento pueden implementarse usando instrucciones almacenadas en un medio legible por ordenador que se ejecutan por el procesador 1712. El medio legible por ordenador puede incluir cualquier combinación deseada de un medio de estado sólido, magnético y/o óptico implementada usando cualquier combinación deseada de dispositivos de almacenamiento masivo (por ejemplo, una unidad de disco), dispositivos de almacenamiento extraíbles (por ejemplo, disquetes, tarjetas o cartuchos de memoria, etc.) y/o dispositivos de memoria integrada (por ejemplo, memoria de acceso aleatorio, memoria flash, etc.).

Aunque se han descrito ciertos métodos, aparatos y artículos de fabricación en el presente documento, el alcance de la cobertura de esta patente no se limita a los mismos.

REIVINDICACIONES

1. Un método de generación de una firma que representa una parte de una señal de audio, comprendiendo el método:

capturar una señal de audio;

multiplicar una primera función de ventana por una parte de la señal de audio capturada que representa un mismo número de muestras que la primera función de ventana para producir un primer bloque de audio de ventana que representa el mismo número de muestras que la primera función de ventana;

multiplicar una segunda función de ventana diferente de la primera función de ventana y que representa el mismo número de muestras que la primera función de ventana por una parte idéntica de la señal de audio capturada que representa el mismo número de muestras que la primera función de ventana para producir un segundo bloque de audio de ventana que representa el mismo número de muestras que la primera función de ventana;

transformar el primer bloque de audio de ventana en una primera representación del dominio de frecuencia;

transformar el segundo bloque de audio de ventana en una segunda representación del dominio de frecuencia;

determinar una primera característica para cada una de una pluralidad de bandas de las frecuencias a partir de la primera representación del dominio de frecuencia;

determinar una segunda característica para cada una de las bandas de las frecuencias a partir de la segunda representación del dominio de frecuencia, siendo la segunda característica de un mismo tipo que la primera característica;

restar la primera característica de la segunda característica para determinar un valor intermedio para cada una de las bandas de las frecuencias; y

asignar un valor de bit de firma en base al valor intermedio determinado para cada una de las bandas de las frecuencias.

2. Un método como se define en la reivindicación 1, en el que capturar la señal de audio comprende una captura de audio inalámbrica.

3. Un método como se define en la reivindicación 2, en el que capturar la señal de audio comprende un muestreo digital.

4. Un método como se define en la reivindicación 1, en el que las funciones de ventana primera y segunda comprenden unas funciones complementarias.

5. Un método como se define en la reivindicación 4, en el que la primera función de ventana comprende una gran magnitud en los extremos superior e inferior de la primera función de ventana.

6. Un método como se define en la reivindicación 4, en el que la primera función de ventana comprende una gran magnitud en un extremo superior de la primera función de ventana y una pequeña magnitud en un extremo inferior de la primera función de ventana.

7. Un método como se define en la reivindicación 1, en el que la primera función de ventana y la parte de la señal de audio capturada se multiplican en una operación del dominio de tiempo.

8. Un método como se define en la reivindicación 1, en el que las características primera y segunda comprenden unas energías primera y segunda.

9. Un método como se define en la reivindicación 1, que comprende además:

aplicar una tercera función de ventana y una cuarta función de ventana a la parte del audio capturado para producir un tercer bloque de audio de ventana y un cuarto bloque de audio de ventana;

transformar el tercer bloque de audio de ventana en una tercera representación del dominio de frecuencia;

transformar el cuarto bloque de audio de ventana en una cuarta representación del dominio de frecuencia;

determinar las terceras características de las bandas de las frecuencias solo a partir de la tercera representación del dominio de frecuencia;

determinar las cuartas características de las bandas de las frecuencias solo a partir de la cuarta representación del dominio de frecuencia, siendo la cuarta característica de un mismo tipo que la segunda característica;

comparar la tercera característica con la cuarta característica de una o más de las bandas de las frecuencias; y

asignar uno o más bits de firma adicionales representativos de las partes respectivas de la señal de audio capturada en base a las comparaciones de las terceras características y las cuartas características.

10. Un aparato para generar una firma que representa una parte de una señal de audio, comprendiendo el aparato:

un generador de muestras para capturar una señal de audio;

un windower para multiplicar una primera función de ventana por una parte de la señal de audio capturada que representa un mismo número de muestras que la primera función de ventana para producir un primer bloque de

- audio de ventana que representa el mismo número de muestras que la primera función de ventana y para multiplicar una segunda función de ventana diferente a partir de la primera función de ventana y que representa el mismo número de muestras que la primera función de ventana por una parte idéntica de la señal de audio capturada que representa el mismo número de muestras que la primera función de ventana para producir un segundo bloque de audio de ventana que representa el mismo número de muestras que la primera función de ventana;
- 5 un transformador para transformar el primer bloque de audio de ventana en una primera representación del dominio de frecuencia y para transformar el segundo bloque de audio de ventana en una segunda representación del dominio de frecuencia;
- 10 un determinador de características para determinar una primera característica para cada una de una pluralidad de bandas de las frecuencias a partir de la primera representación del dominio de frecuencia y para determinar una segunda característica para cada una de las bandas de las frecuencias a partir de la segunda representación del dominio de frecuencia, siendo la segunda característica de un mismo tipo que la primera característica;
- un comparador para restar la primera característica de la segunda característica para determinar un valor intermedio para cada una de las bandas de las frecuencias; y
- 15 un determinador de firmas para asignar un valor de bit de firma en base al valor intermedio determinado para cada una de las bandas de las frecuencias.
11. Un aparato como se define en la reivindicación 10, en el que el windower multiplica la primera función de ventana y la parte de la señal de audio capturada en una operación del dominio de tiempo, y las funciones de ventana primera y segunda comprenden unas funciones complementarias.
- 20

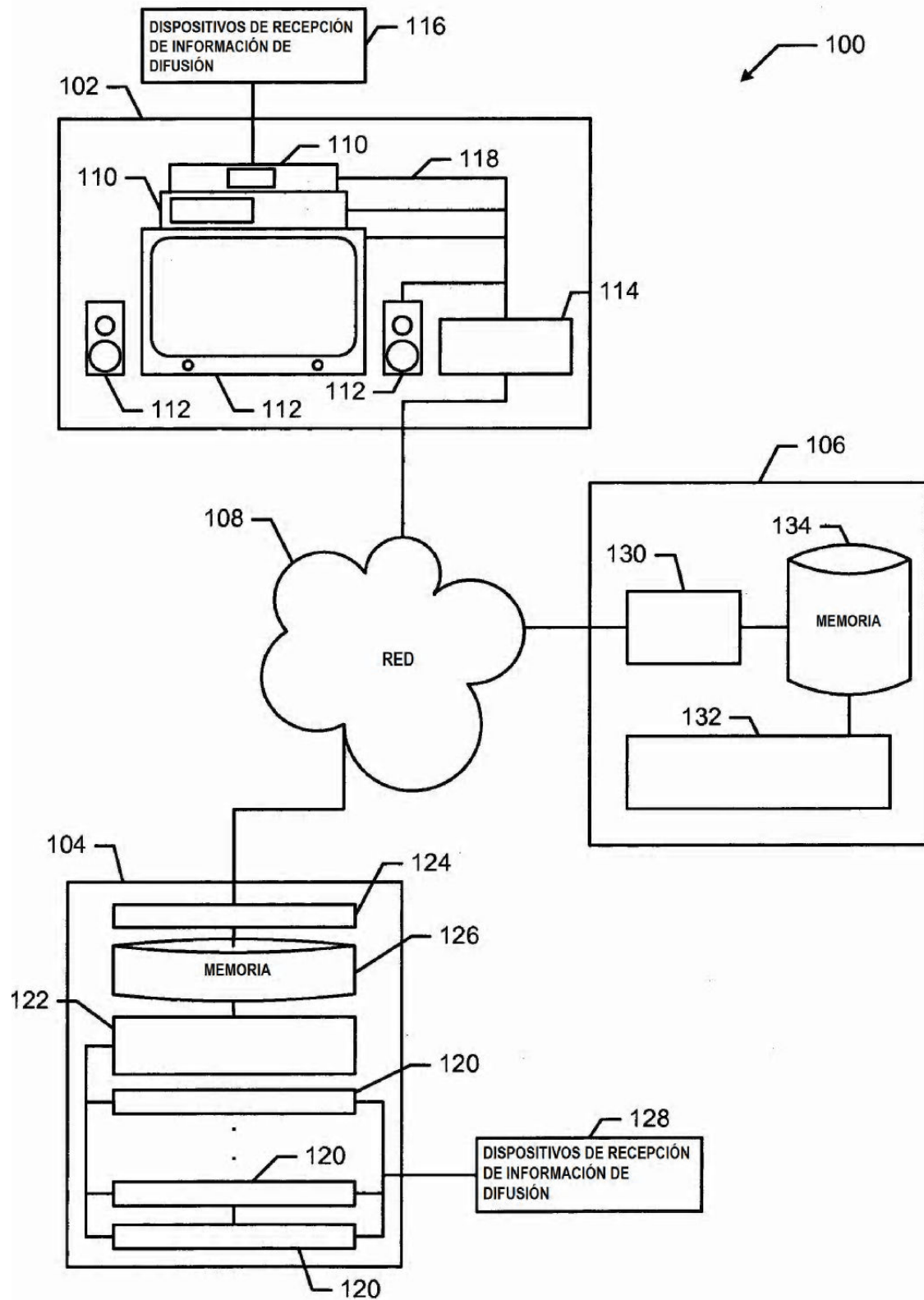


FIG. 1A

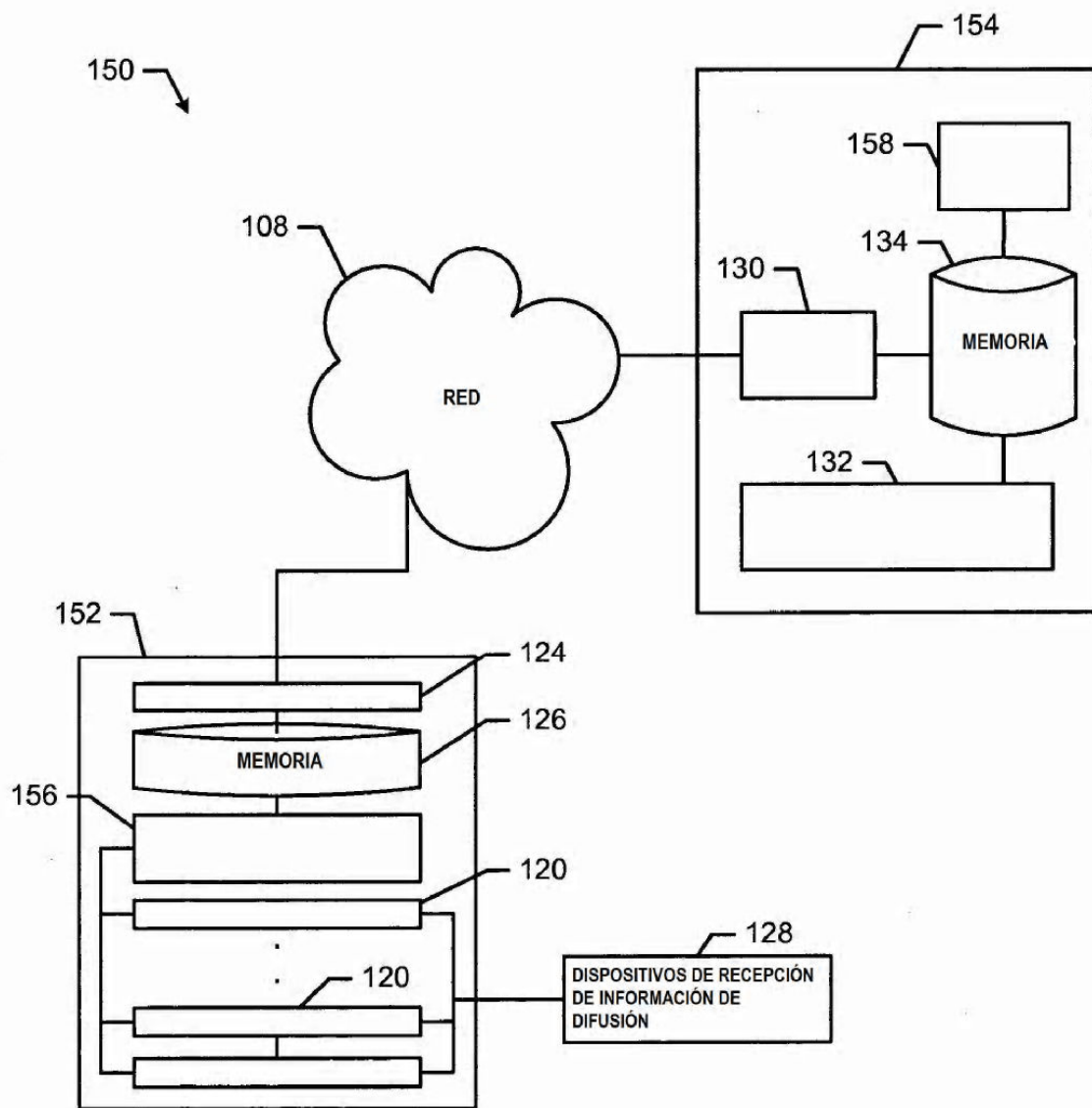


FIG. 1B

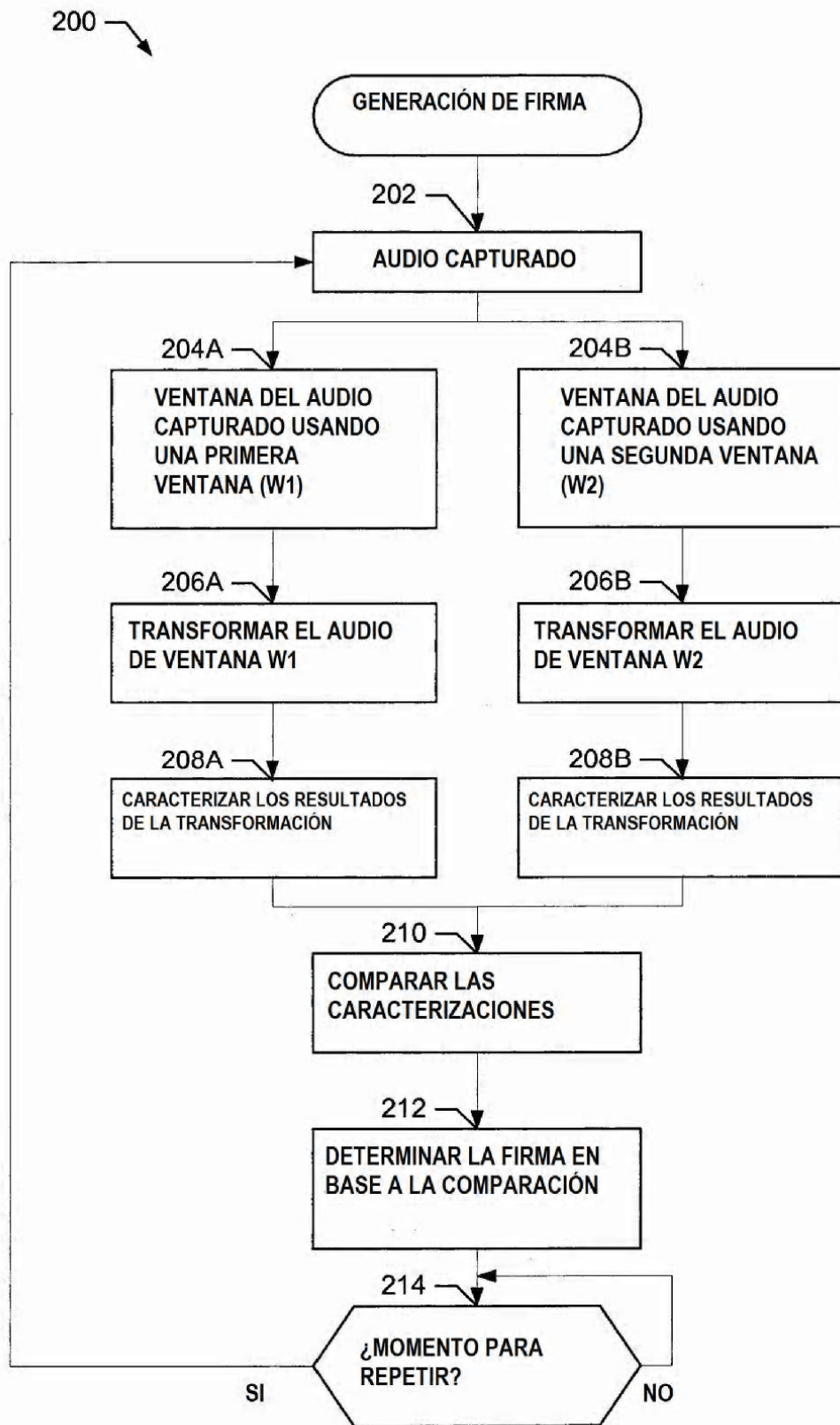


FIG. 2

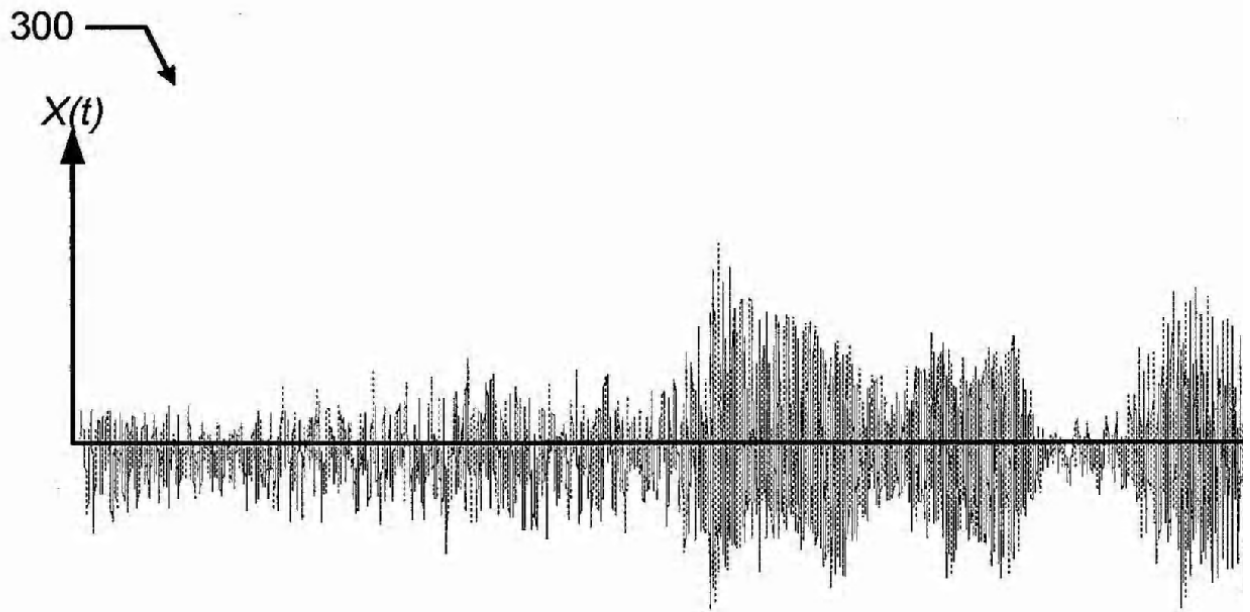


FIG. 3

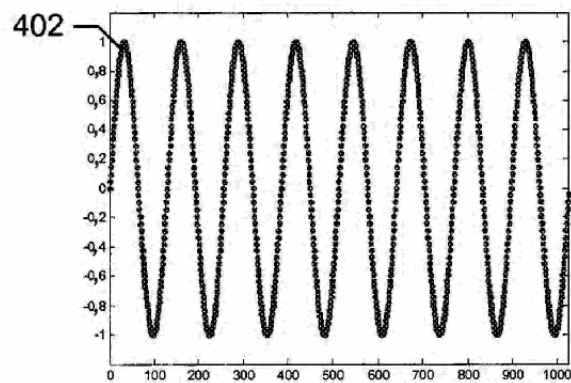


FIG. 4

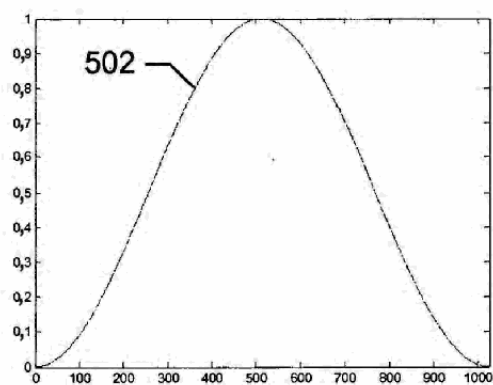


FIG. 5

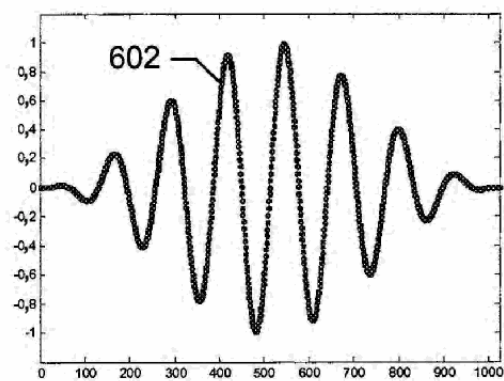


FIG. 6

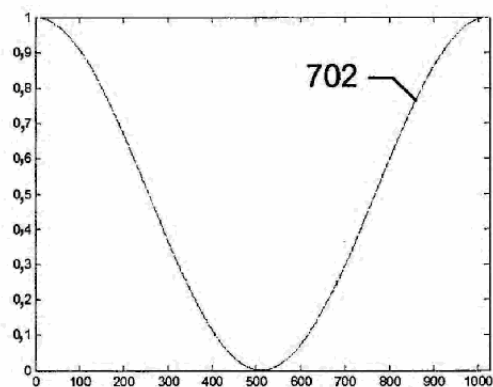


FIG. 7

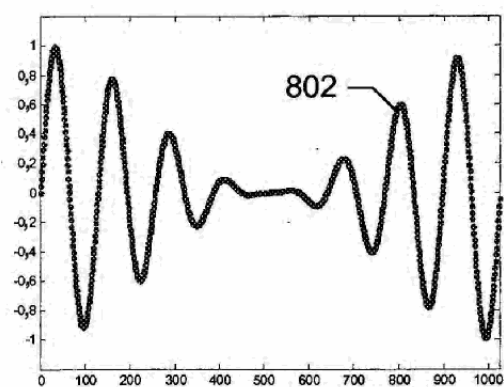


FIG. 8

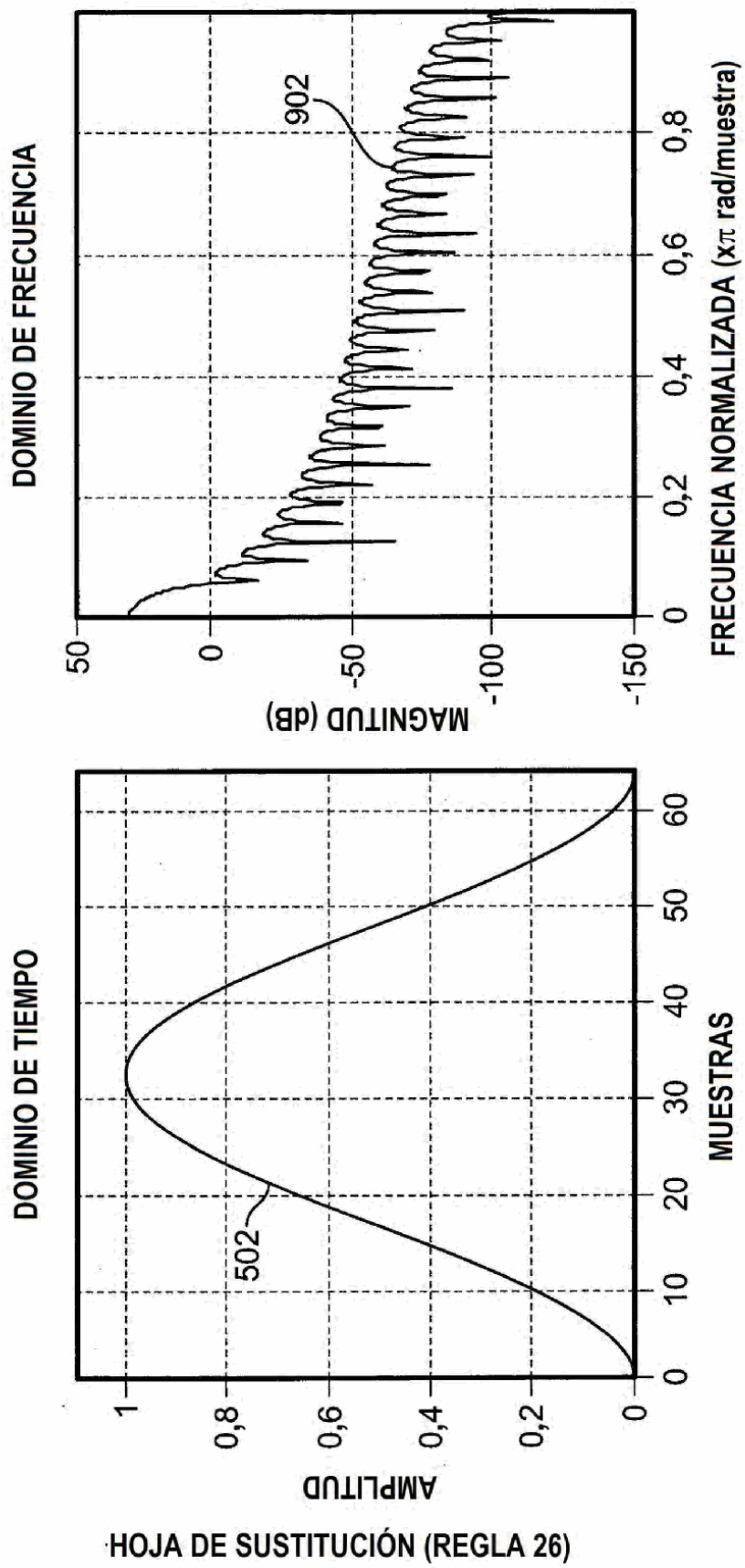


FIG. 9

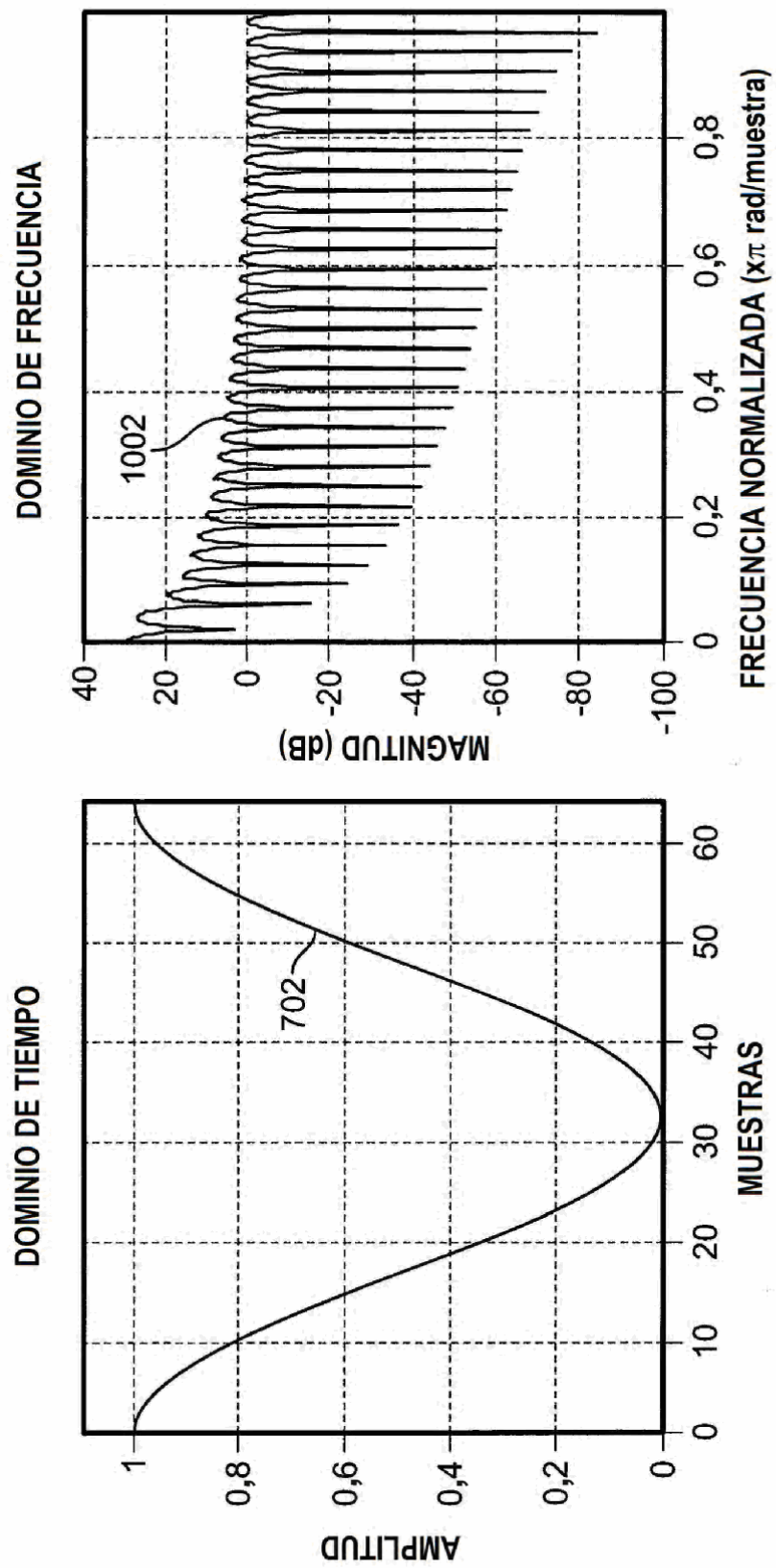


FIG. 10

HOJA DE SUSTITUCIÓN (REGLA 26)

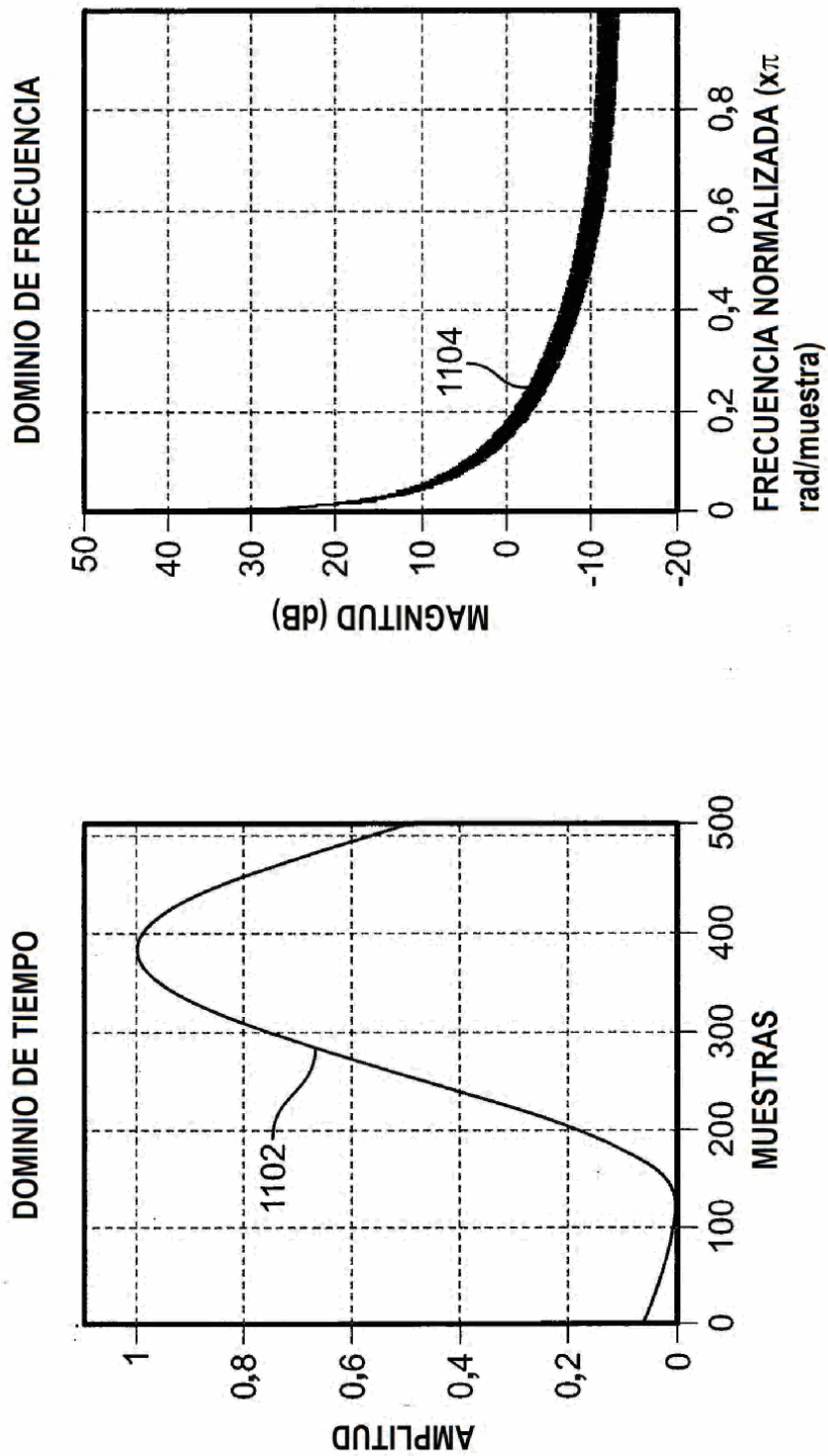


FIG. 11

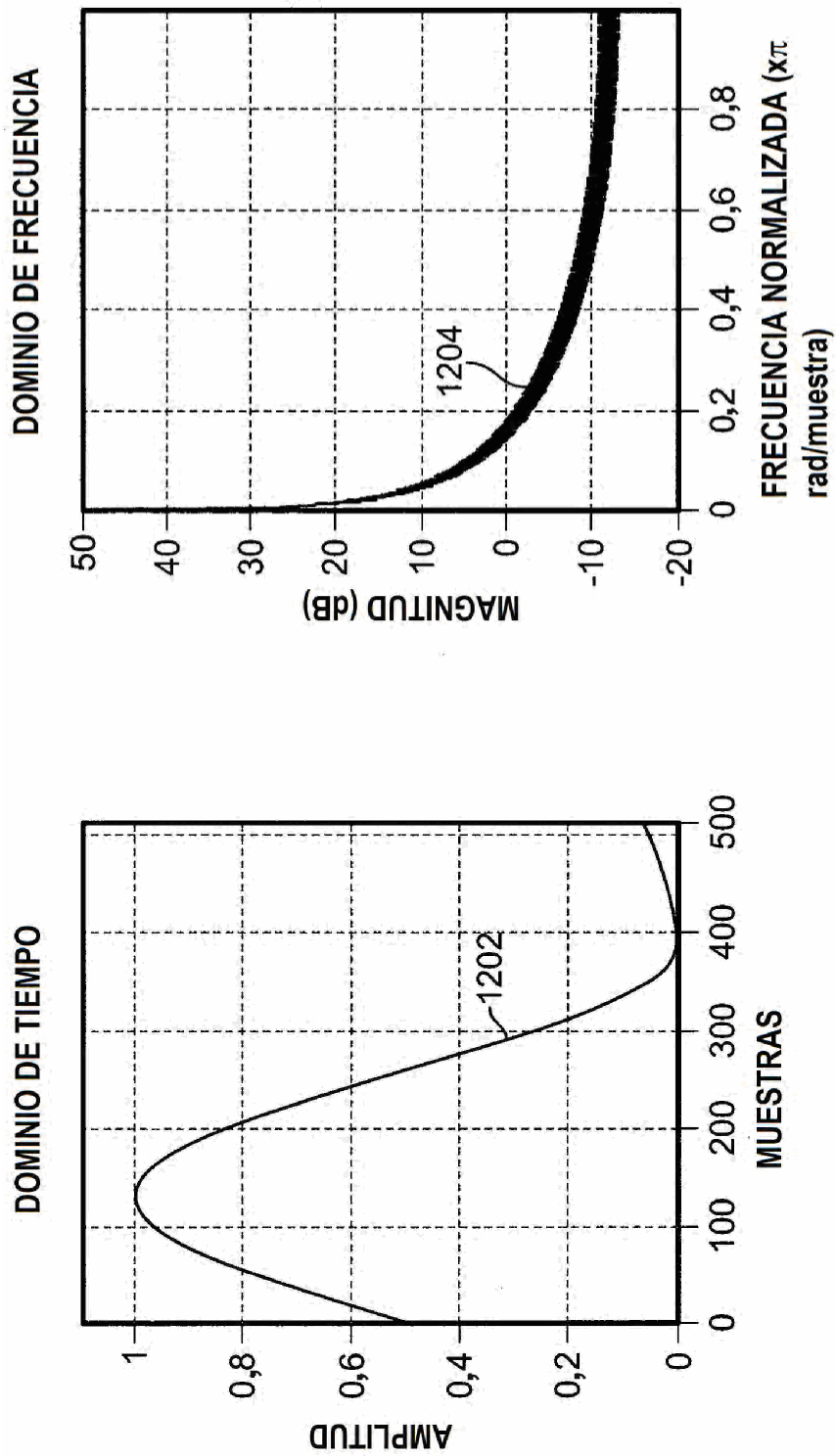


FIG. 12

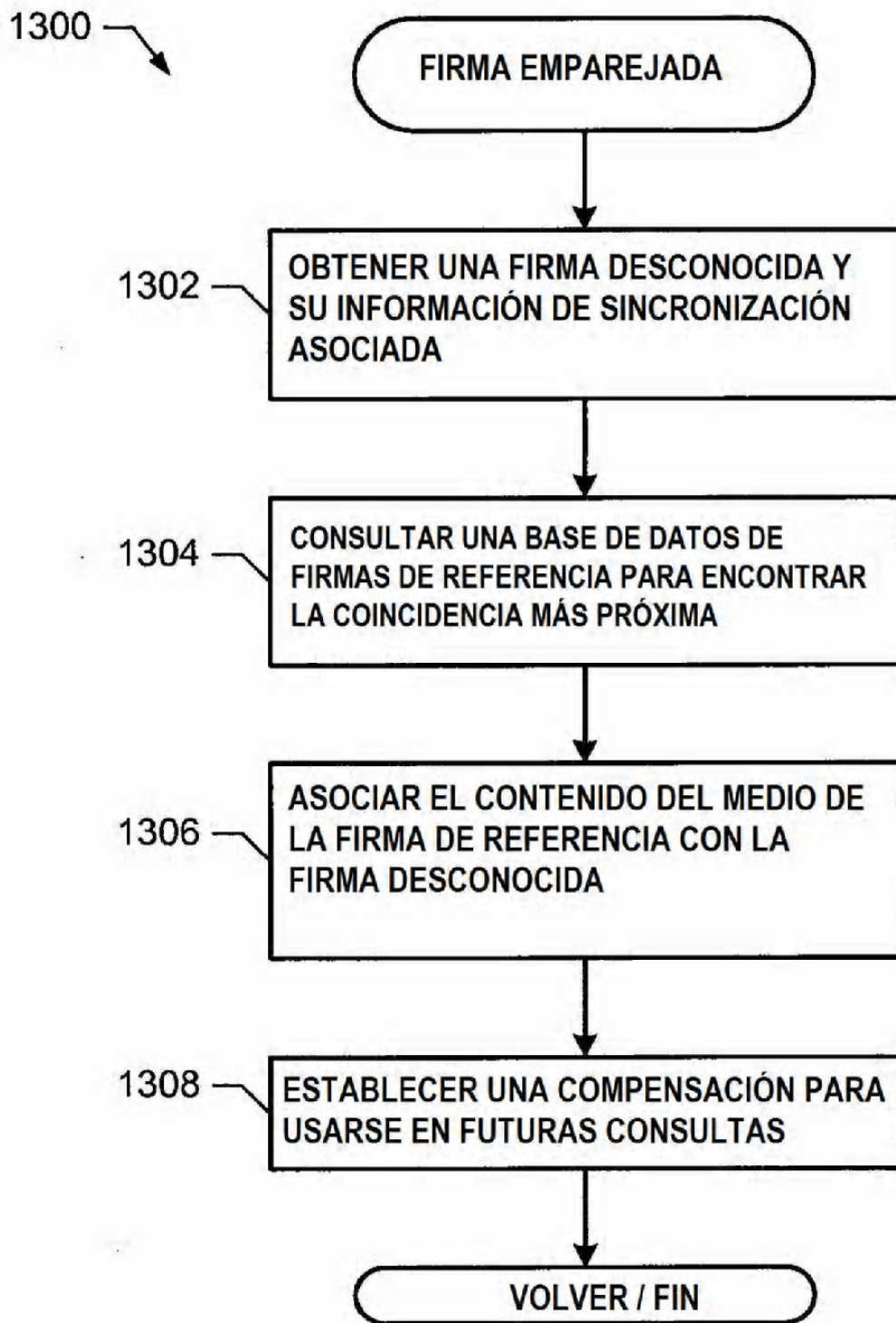


FIG. 13

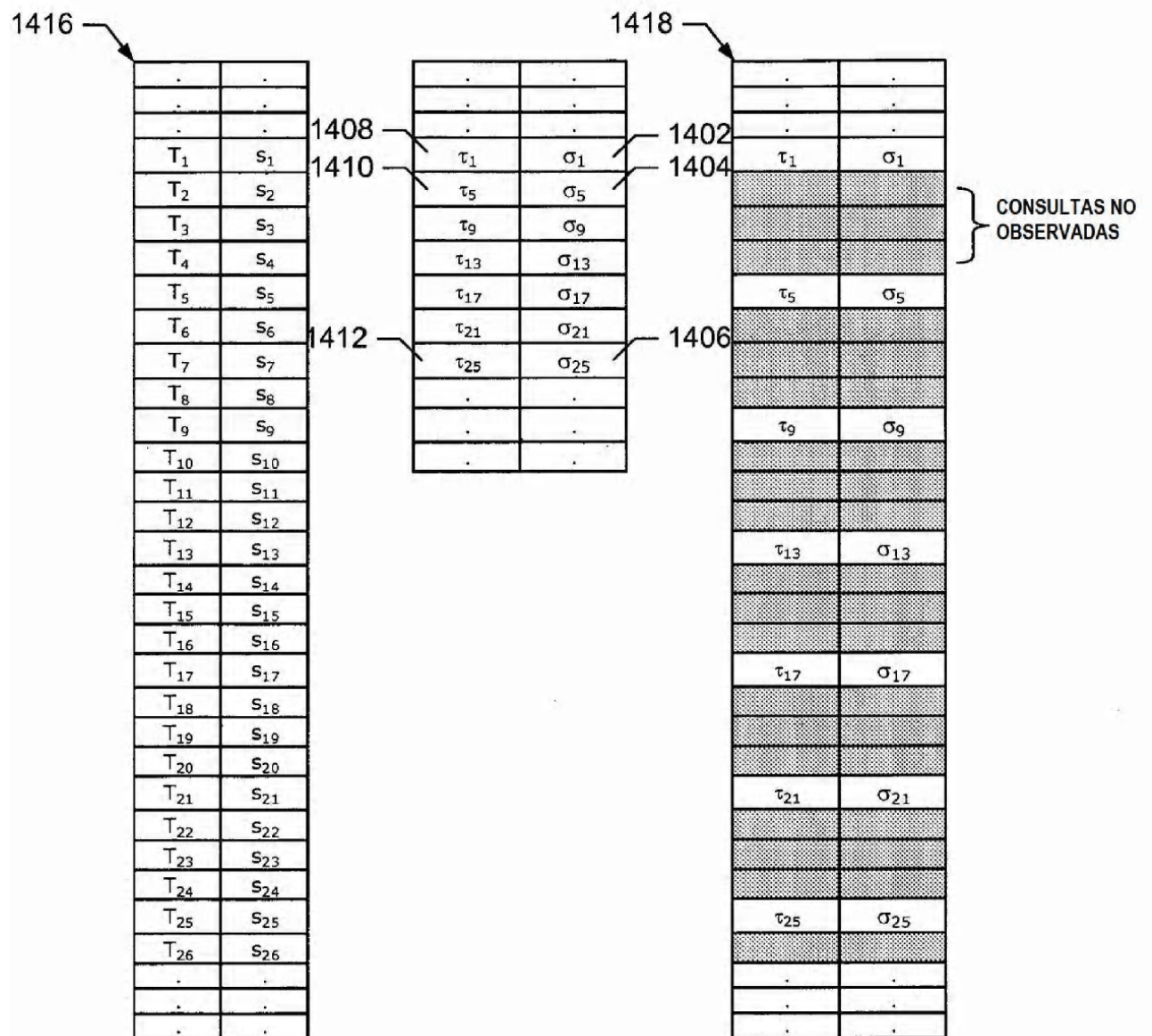


FIG. 14

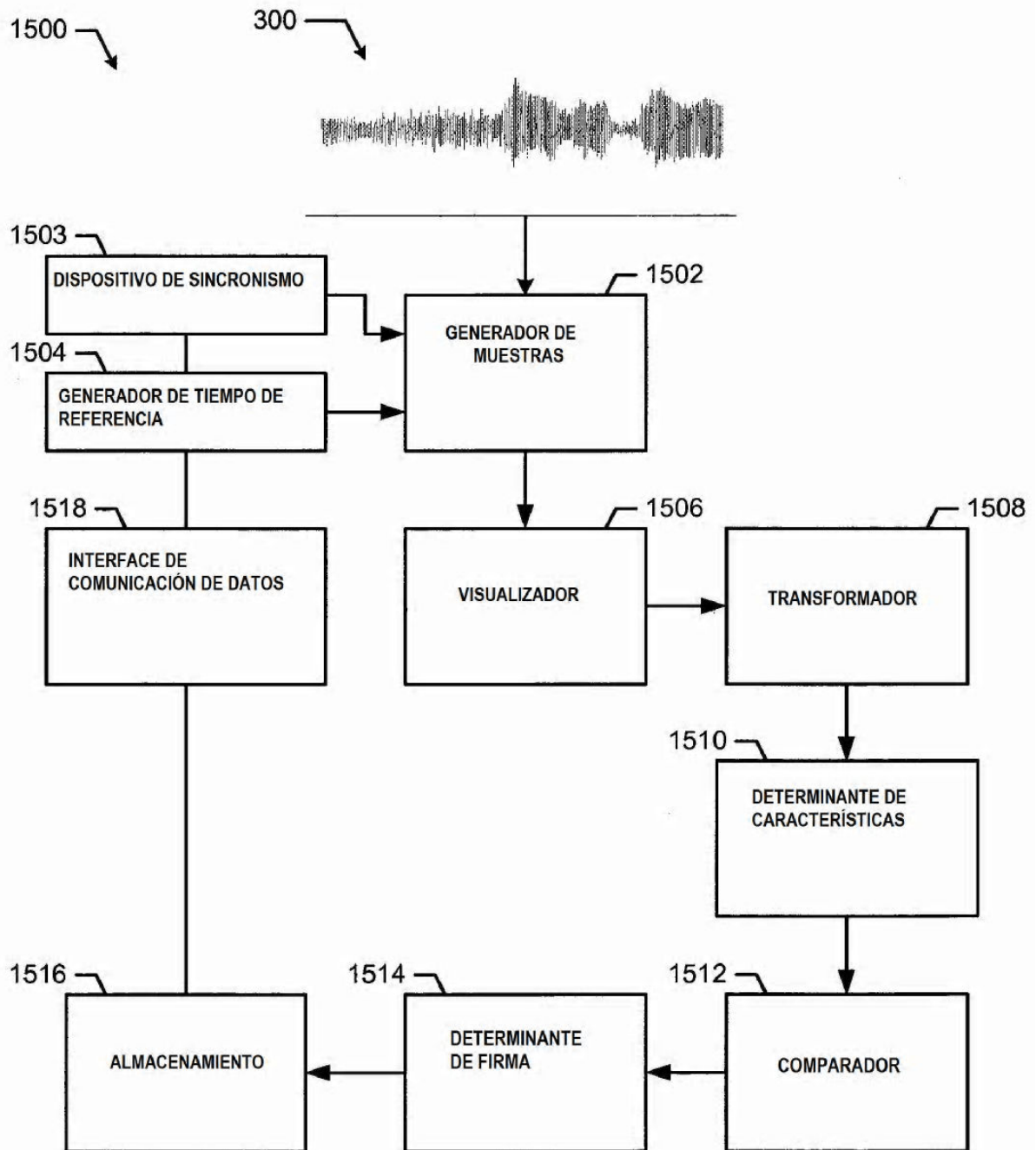


FIG. 15

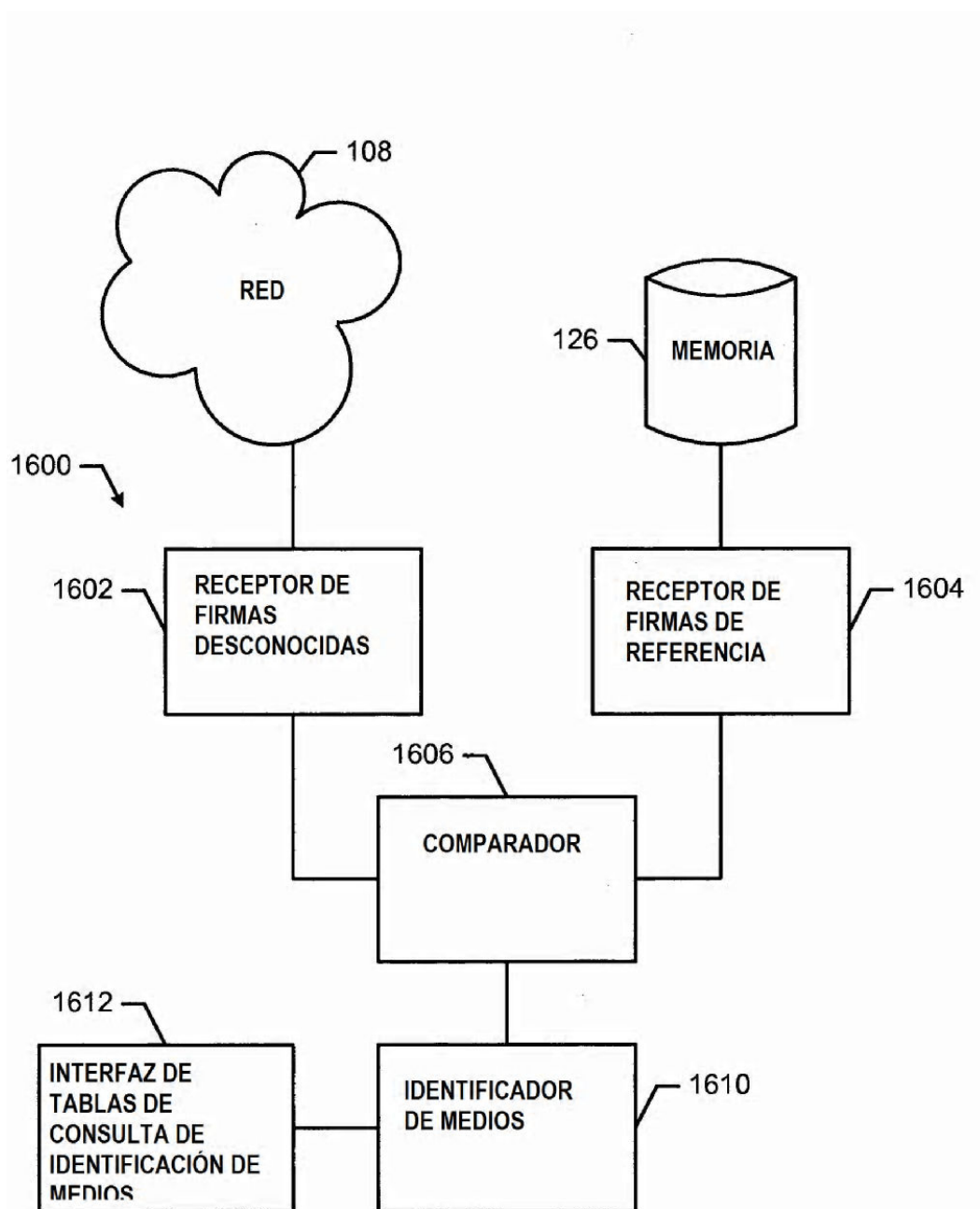


FIG. 16

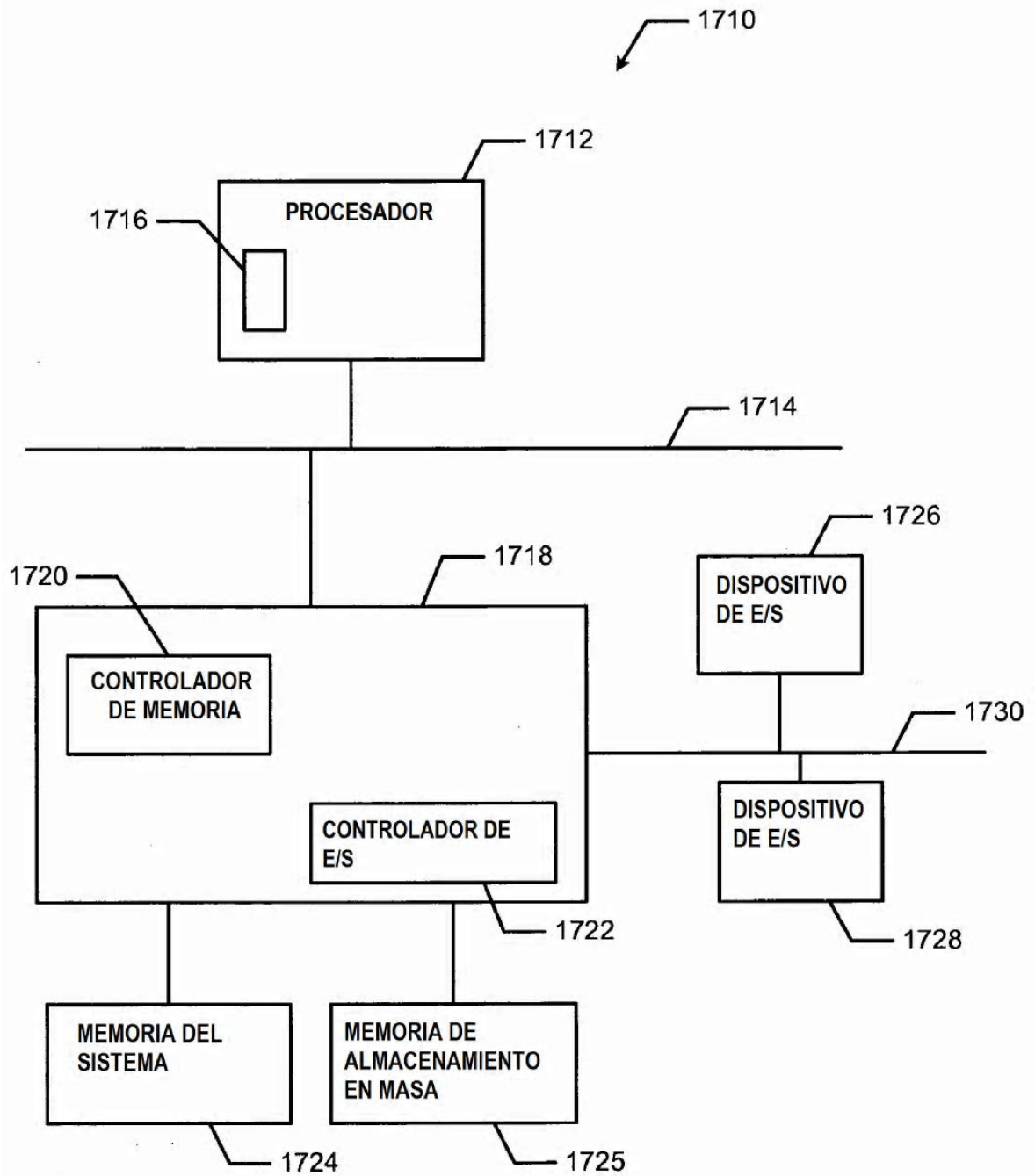


FIG. 17