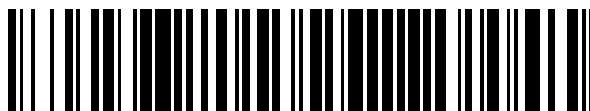


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 558 167**

51 Int. Cl.:

H04R 3/00 (2006.01)

G10L 19/018 (2013.01)

G01S 3/80 (2006.01)

G01S 5/20 (2006.01)

H04R 1/40 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.08.2012 E 12756427 (6)**

97 Fecha y número de publicación de la concesión europea: **21.10.2015 EP 2751587**

54 Título: **Estimación de dirección de llegada usando señales de audio con marca de agua y disposiciones ordenadas de micrófono**

30 Prioridad:

31.08.2011 EP 11179591

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

02.02.2016

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**THIERGART, OLIVER;
DEL GALDO, GIOVANNI;
KOLBECK, FLORIAN;
CRACIUN, ALEXANDRA;
KRÄGELOH, STEFAN;
BORSUM, JULIANE y
BLIEM, TOBIAS**

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 558 167 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

Estimación de dirección de llegada usando señales de audio con marca de agua y disposiciones ordenadas de micrófono

La presente invención se refiere a determinar la dirección de llegada de una fuente acústica tal como un altavoz, y un aparato y método para estimación de posición espacial.

Recuperar información geométrica de una escena de sonido, tal como estimación de localización y estimación de dirección de llegada de señales entrantes, se está haciendo cada vez más importante, ya que esta información es valiosa para una variedad de aplicaciones, tales como control de eco acústico, análisis de escena, formación de haz, procesamiento de señales para sensores acústicos distribuidos.

Por ejemplo, hay métodos que analizan una señal de audio registrada, en donde no se incluye señal con marca de agua alguna, si un único altavoz está reproduciendo un tema de audio y una disposición ordenada de micrófono está disponible para hacer la grabación. Si no hay otra fuente de sonido presente, por caso, un humano que habla u otra interferencia, es posible estimar la dirección de llegada, por ejemplo, vía Codificación de Audio Direccional (*Directional Audio Coding*). Se hace referencia al documento de Jukka Ahonen, Giovanni Del Galdo, Markus Kallinger, Fabian Küch, Ville Pulkki, y Richard Schultz-Amling, "Planar microphone array processing for the analysis and reproduction of spatial audio using directional audio coding", en Audio Engineering Society Convention 124, 5, 2008. Tan pronto como otra fuente está emitiendo sonido al mismo tiempo, cuya dirección de llegada no se desea, se dificulta la determinación del ángulo de incidencia correcto. Uno obtendrá dos estimaciones diferentes de la dirección de llegada si las señales presentes se originan en diferentes posiciones espaciales. En este caso, se necesita información adicional acerca de la señal de sonido favorecida.

En un escenario de un sistema de reproducción con más que solo un altavoz, si las señales reproducidas son diferentes, es posible estimar la dirección de llegada usando los enfoques explicados previamente. Sin embargo, si se reproducen temas de audio similares, esto es, en una organización de estéreo, aparece un fenómeno conocido, llamado fuente de imagen fantasma. Esto significa que el sonido es percibido como si fuera reproducido por una fuente virtual ubicada entre los altavoces. En este caso, falla la estimación de los ángulos de incidencia.

Para la aplicación especial que involucra la calibración de un sistema de reproducción, existen otros métodos conocidos, por ejemplo, reproducir señales MLS o de barrer, las cuales se usan para determinar el ángulo de incidencia del sonido reproducido. Se hace referencia a Giovanni Del Galdo, Matthias Lang, Jose Angel Pineda Pardo, Andreas Silzle, y Oliver Thiergart, "Acoustic measurement system for 3-D loudspeaker set-ups", en la Conferencia de la Sociedad de Ingeniería de Audio: Captar el Sonido del Espacio, 10 2010. Sin embargo, estas señales son audibles y con frecuencia fastidiosas. Asimismo, la calibración necesita ser hecha antes de que se pueda usar el sistema de reproducción. Por eso, el uso de señales de audio con marca de agua es preferible teniendo en cuenta que la calibración puede ser realizada durante la operación del sistema.

Una técnica frecuentemente usada para determinar la dirección o la posición de un sensor emplea señales de radio frecuencia y varios sensores. Los métodos basados en esta tecnología proveen buenas estimaciones, pero no son aplicables en áreas donde están prohibidas las señales de radio frecuencia o es difícil usarlas. Asimismo, en tales escenarios, habría que instalar los sensores de radio frecuencia.

Otro método que también hace uso de señales con marca de agua está presentado por Ryuki Tachibana, Shuichi Shimizu, Seiji Kobayashi, y Taiga Nakamura, en "An audio watermarking method using a two-dimensional pseudo-random array", Signal Process, vol. 82, pp. 1455-1469, octubre de 2002. Aquí se estima una posición de grabación usando sólo un micrófono y midiendo el retardo de tiempo de llegada de varias señales de audio con marca de agua emitidas. En este escenario no hay información de dirección disponible y el sistema está restringido al uso de un mínimo número de altavoces.

El estado de la técnica también está representado adicionalmente por los documentos de patente US2008/247274 y US2010/0286949 o por el documento "Real-Time User Position Estimation in Indoor Environments Using Digital Watermarking for Audio Signals" de Ryosuke Kaneto et al. (ICPR 2010)

Como se describió arriba, la técnica previa exhibe varios problemas. De acuerdo con algunos métodos de la técnica previa, se puede localizar una fuente fantasma que no corresponde al altavoz deseado. Otros métodos de la técnica previa exhiben el problema de que localizarían otras fuentes en el ambiente o que necesitan señales de medición especiales las cuales no pueden ser reproducidas en una situación real donde los se usan realmente los altavoces. El objetivo de la presente invención es proveer conceptos mejorados para determinar una dirección de llegada de una señal de transmisión. De acuerdo con otro aspecto de la presente invención, se proveen conceptos mejorados para estimación de posición espacial. El objetivo de la presente invención se resuelve mediante un aparato para proveer información de dirección de acuerdo con la reivindicación 1, un método para proveer información de

dirección de acuerdo con la reivindicación 14 y un programa de ordenador para proveer información de dirección de acuerdo con la reivindicación 16.

Se provee un aparato para proveer información de dirección en base a una señal de audio reproducida con una marca de agua incrustada. El aparato comprende un procesador de señal, el cual está adaptado para procesar por lo menos dos señales de audio con marca de agua registradas por al menos dos receptores de audio en diferentes posiciones espaciales. Los receptores de audio están adaptados para registrar ondas de sonido de la señal de audio reproducida para obtener las señales de audio con marca de agua registradas. Cada señal de audio con marca de agua registrada comprende la marca de agua incrustada. El procesador de señal está adaptado para procesar las señales de audio con marca de agua registradas para obtener información de fase como información específica de receptor para cada señal de audio con marca de agua registrada. Asimismo, la información específica de receptor depende de las marcas de agua incrustadas, incrustadas en las señales de audio con marca de agua registradas. Asimismo, el aparato comprende un proveedor de información de dirección para proveer información de dirección en base a la información específica de receptor para cada señal de audio con marca de agua registrada.

Una realización describe un método usado para determinar la información de dirección, por ejemplo, dirección de llegada (DoA) de una señal de audio reproducida por un altavoz. La señal actúa como una portadora para una señal de marca de agua inaudible que fue generada en el dominio de tiempo-frecuencia. Es registrada por una disposición ordenada de micrófonos que comprende por lo menos dos micrófonos. Luego de esto, los registros son procesados adicionalmente con el objetivo de detectar la señal con marca de agua incrustada. Una vez que se recupera la marca de agua, se la puede usar para estimar la dirección del origen del sonido relativo a una orientación de la disposición ordenada. Dependiendo de la aplicación, este concepto puede ser adaptado adicionalmente y aplicado a una organización que comprende un número arbitrario de altavoces.

En una realización, el procesador de señal comprende un módulo de análisis para transformar las señales de audio con marca de agua recibidas desde un dominio del tiempo a un dominio de tiempo-frecuencia para obtener una señal de transmisión del dominio de tiempo-frecuencia que comprende una pluralidad de señales de subbanda. En una realización, el procesador de señal puede estar adaptado para determinar un número complejo como información específica de receptor. Por ejemplo, el número complejo puede indicar el valor de un coeficiente complejo o un coeficiente complejo promediado de una señal de audio con marca de agua registrada por uno de los receptores de audio. En otra realización, el procesador de señal puede estar adaptado para determinar información de fase como información específica de receptor. Por ejemplo, la información de fase puede ser un valor de fase o un valor de fase promediado de una señal de audio con marca de agua registrada por uno de los receptores de audio. En una realización adicional, el procesador de señal puede estar adaptado para determinar un valor de amplitud o un valor de amplitud promediado de una señal de audio con marca de agua registrada por uno de los receptores de audio como información específica de receptor.

De acuerdo con una realización, el procesador de señal está adaptado para determinar un valor de fase promediado como información de fase en base a un coeficiente complejo promediado determinado. El procesador de señal puede ser adaptado para generar el coeficiente complejo promediado determinando un bloque periódico promediado que comprende coeficientes complejos procesados de una señal subbanda.

En una realización, el procesador de señal está adaptado para determinar coeficientes complejos promediados combinando por lo menos dos coeficientes complejos procesados de una señal subbanda, en donde los coeficientes complejos procesados tienen iguales valores de fase o valores de fase que difieren en no más que 0,5 medido en radianes. El procesador de señal puede ser adaptado asimismo para generar coeficientes complejos procesados cambiando el valor de fase de un coeficiente complejo por un número predeterminado, por ejemplo, por 180° cuando se usa BPSK antípoda, o por un valor de fase apropiados dependiendo del esquema de modulación usado y de la secuencia incrustada conocida.

De acuerdo con una realización, el procesador de señal está adaptado para determinar una posición de impacto de sincronización que indica la posición de un coeficiente complejo que codifica un primer bit de una firma de marca de agua.

El proveedor de información de dirección puede ser adaptado para emplear valores de fase para determinar y proveer información de dirección. Asimismo, el proveedor de información de dirección puede ser adaptado para proveer una dirección de vector de llegada como información de dirección.

Asimismo, se provee un método para proveer información de dirección. El método comprende los pasos de registrar ondas de sonido de la señal de audio reproducida mediante por lo menos dos receptores de audio en diferentes posiciones espaciales para obtener al menos dos señales de audio con marca de agua registradas, donde cada una de las por lo menos dos señales de audio con marca de agua registradas comprende la marca de agua incrustada, procesar las por lo menos dos señales de audio con marca de agua registradas para obtener información de fase como información específica de receptor para cada señal de audio con marca de agua registrada, en donde la información específica de receptor depende de las marcas de agua incrustadas, incrustadas en las señales de audio

con marca de agua registradas, y proveer información de dirección en base a la información específica de receptor para cada una de las por lo menos dos señales de audio con marca de agua registrada.

5 Asimismo, se provee un aparato para estimación de posición espacial. El aparato comprende un aparato para proveer información de dirección de acuerdo con una realización, y un estimador de posición para estimar una posición del aparato para estimar posición espacial, en donde el estimador de posición está adaptado para estimar la posición del aparato para estimación de posición espacial en base a la información de dirección provista por el aparato para proveer información de dirección.

10 El concepto para estimación de posición espacial se basa en determinar la posición de registro espacial de la disposición ordenada de micrófonos provisto que un número suficiente de altavoces está disponible y se conocen sus posiciones espaciales. Por ejemplo, se pueden emplear tres altavoces para determinar la posición de una disposición ordenada de micrófonos en un plano común. Se debe mencionar que el concepto propuesto no está restringido a la determinación del azimut. Además de esto, se puede estimar la elevación dependiendo del método de estimación de dirección usado.

15 Asimismo, se provee un método para estimación de posición espacial. El método comprende recibir señales de audio con marca de agua empleando una disposición ordenada de por lo menos dos receptores de audio, cada señal de audio con marca de agua recibida que comprende una marca de agua incrustada; procesar por lo menos dos señales de audio con marca de agua recibidas registradas por al menos dos receptores de audio en diferentes posiciones espaciales para determinar una información específica de receptor para cada señal de audio con marca de agua recibida, en donde la información específica de receptor depende de las marcas de agua incrustadas, incrustadas en las señales de audio con marca de agua recibidas; proveer información de dirección en base a la información específica de receptor para cada señal de audio con marca de agua recibida.; y estimar una posición de la disposición ordenada de por lo menos dos receptores de audio, en donde la posición es determinada en base a la información de dirección.

20 Ejemplos de posibles aplicaciones son, por ejemplo, juguetes que puedan reaccionar a la reproducción de una cierta señal de audio, por ejemplo, una muñeca con acción podría mover su cabeza hacia una TV cuando se reproduce una cierta serie de TV.

25 Un ejemplo de aplicación adicional es calibración de sistema de reproducción: La presente invención podría ser usado para calibrar automáticamente un sistema de reproducción, por ejemplo, una organización para teleconferencia o un sistema de "home theater", mientras la conferencia ya está en proceso o se ejecuta un tema de audio. En este caso, no es necesaria calibración previa.

30 Además los canales de una señal de audio podrían ser manejados para subir o bajar el sonido de ciertos altavoces en ciertas posiciones dependiendo de la información de dirección provista. Asimismo, un usuario podría recibir instrucciones del sistema sobre cómo ubicar de modo óptimo los altavoces de un sistema de sonido dependiendo de la información de dirección provista.

Un ejemplo de aplicación adicional es el aparato descrito arriba para estimación de posición espacial.

35 A continuación se explicarán realizaciones preferidas con referencia a los dibujos, en los cuales:

la Figura 1 muestra un diagrama de bloques esquemático de un aparato para proveer información de dirección de acuerdo con una realización;

40 la Figura 2 ilustra un aparato para proveer información de dirección en un escenario de aplicación de acuerdo con una realización;

las Figuras 3a-3c ilustran una marca de agua que se dispersa en frecuencia y se dispersa en tiempo;

45 la Figura 4 muestra un modulador para generar una señal con marca de agua;

la Figura 5 representa un aparato para generar una señal con marca de agua;

50 las Figuras 6a-6b ilustran un aparato para proveer información de dirección en otros escenarios de aplicación de acuerdo con realizaciones;

60 la Figura 7 muestra los componentes de un aparato para proveer información de dirección de acuerdo con una realización;

65 la Figura 8a provee un panorama ejemplar de la localización de los coeficientes sobre el plano de tiempo-frecuencia cuando se emplea sobremuestreo de acuerdo con una realización;

- la Figura 8b ilustra una señal subbanda del dominio de tiempo-frecuencia que comprende marcas de agua incrustadas;
- 5 las Figuras 9a-9d representan firmas de marca de agua, bloques de sincronización, bloques periódicos y un bloque de sincronización promediado de acuerdo con una realización;
- la Figura 10 ilustra marcas de agua multiplexadas de acuerdo con una realización;
- 10 las Figuras 11a-11b muestran diagramas en bloques esquemáticos de alternativas de implementación de una unidad de soporte de sincronización;
- la Figura 11c muestra un diagrama de bloques esquemáticos de un correlacionador de firma de sincronización;
- 15 la Figura 12a muestra una representación gráfica del problema de hallar la alineación temporal de una marca de agua;
- la Figura 12b muestra una representación gráfica del problema de identificar el inicio del mensaje;
- 20 la Figura 12c muestra una representación gráfica de datos usados para sincronización;
- la Figura 12d muestra una representación gráfica de un concepto de identificar un impacto de sincronización;
- la Figura 13a muestra una representación gráfica de un ejemplo de una desdispersión temporal;
- 25 la Figura 13b muestra una representación gráfica de un ejemplo para una multiplicación elemento por elemento entre bits y secuencias de dispersión;
- la Figura 13c muestra una representación gráfica de una salida del correlacionador de firma de sincronización después del promediado temporal;
- 30 la Figura 13d muestra una representación gráfica de una salida del correlacionador de firma de sincronización filtrada con la función autocorrelación de la firma de sincronización;
- 35 la Figura 14 ilustra la relación de bloques periódicos que comprenden bits decodificados diferencialmente y bloques periódicos que comprenden coeficientes complejos;
- la Figura 15a representa vectores complejos de una porción de señal de audio y una porción de señal de marca de agua;
- 40 la Figura 15b representa un vector complejo promediado de porciones de señal de audio y un vector complejo promediado de porciones de señal de marca de agua;
- la Figura 15c ilustra integración de muestra en diferentes etapas;
- 45 la Figura 15d ilustra dos vectores de coeficientes complejos de diferentes subbandas;
- la Figura 16 ilustra seis micrófonos acomodados a lo largo de un sistema de coordenadas cartesiano;
- 50 la Figura 17 muestra determinación de azimut en un plano xy de acuerdo con una realización;
- la Figura 18a-d representa dirección de vectores de llegada determinados de acuerdo con una realización;
- la Figura 19 ilustra estimación de posición de acuerdo con una realización;
- 55 la Figura 20a-b representa desafíos y soluciones de estimación de posición;
- la Figura 21 ilustra un aparato para estimación de posición de acuerdo con una realización; y
- 60 la Figura 22 ilustra un aparato para estimación de posición espacial de acuerdo con una realización.

1. Aparato para proveer Información de Dirección de acuerdo con la Figura 1

La Figura 1 ilustra un aparato 100 para proveer información de dirección en base a una señal de audio reproducida con una marca de agua incrustada de acuerdo con una realización. El aparato 100 comprende un procesador de señal 110. El procesador de señal 110 está adaptado para procesar por lo menos dos señales de audio con marca de agua recibidas rx1, rx2 registradas por al menos dos receptores de audio en diferentes posiciones espaciales. Los por lo menos dos receptores de audio están adaptados para registrar ondas de sonido de la señal de audio reproducida para obtener las por lo menos dos señales de audio con marca de agua recibidas. Cada señal de audio con marca de agua recibida rx1, rx2 comprende la marca de agua incrustada. Asimismo, el procesador de señal 110 está adaptado para procesar las señales de audio con marca de agua para obtener una información específica de receptor rx1, rx2 para cada señal de audio recibida rx1, rx2. La información específica de receptor rsi1, rsi2 depende de las marcas de agua incrustadas en las señales de audio con marca de agua recibidas rx1, rx2. Asimismo, el aparato 100 para proveer información de dirección comprende un proveedor de información de dirección 120. El proveedor de información de dirección está adaptado para proveer información de dirección di en base a la información específica de receptor rsi1, rsi2 para cada señal de audio recibida rx1, rx2.

2. Escenario de Aplicación de acuerdo con la Figura 2:

La Figura 2 ilustra un escenario, en donde un aparato 100 para proveer información de dirección recibe dos señales de audio con marca de agua rx1, rx2 desde dos receptores de audio 221, 222. En la Figura 2, una señal de audio awn que comprende una marca de agua es reproducida por un altavoz 210 como ondas de sonido sw. Las ondas de sonido sw son registradas por al menos dos receptores de audio 221, 222, por ejemplo, dos micrófonos para obtener por lo menos dos señales de audio recibidas rx1, rx2. Cada señal de audio recibida rx1, rx2 comprende la marca de agua incrustada.

Las señales de audio con marca de agua recibidas rx1, rx2 entonces son ingresadas al procesador de señal 110 del aparato 100 para proveer información de dirección. El procesador de señal 110 está adaptado para procesar las señales de audio con marca de agua rx1, rx2 para obtener una información específica de receptor rx1, rx2 para cada señal de audio con marca de agua recibida rx1, rx2. La información específica de receptor puede ser información de fase que depende de la marca de agua que está comprendida en las señales de audio con marca de agua recibidas. Asimismo, la información de fase depende de la posición espacial de los receptores de audio. Por ejemplo, una onda de sonido de una señal de audio reproducida que comprende la marca de agua puede llegar un poco antes a un primer receptor de audio 221 que a un segundo receptor de audio 222. Por lo tanto, la marca de agua en la primera señal recibida rx1 es recibida un poco antes que en la segunda señal recibida rx2. En general, ambas señales recibidas rx1, rx2 son muy similares. Sin embargo, como ambas señales están levemente desplazadas en el tiempo una respecto a la otra, transformar ambas señales recibidas rx1, rx2 desde un dominio del tiempo a un dominio de tiempo-frecuencia, en generar da por resultado diferentes valores de fase de ambas señales del dominio de la frecuencia con respecto a igual subbanda de frecuencia e índice de tiempo.

El procesador de señal 110 alimenta la información específica de receptor rsi1, rsi2 en el proveedor de información de dirección 120. El proveedor de información 120 de dirección está adaptado para proveer información de dirección di en base a la información específica de receptor rsi1, rsi2 para cada señal de audio recibida rx1, rx2. Por ejemplo, la información de dirección di puede ser provista en base a información de fase. Un primer valor de fase de la primera señal recibida rx1 y un segundo valores de fase de la segunda señal recibida rx2 pueden ser procesados, por ejemplo, calculando un diferencia de fase. En base a la diferencia de fase calculada, el proveedor de información de dirección 120 puede determinar la información de dirección di. Por ejemplo, se pueden determinar posibles direcciones desde donde se originan las ondas de sonido de la señal de audio reproducida.

3. Generación de marca de agua:

A continuación, para proveer información de fondo acerca de marcas de agua, se explica generación de marca de agua en un codificador con referencia a los dibujos representados en las Figuras 3a - 5. En las Solicitudes de Patente Europea, Solicitud Número 10154953.3, 10154960.8, 10154964.0, 10154948.3, 10154956.6 y 10154951.7 las cuales son incorporada a la presente como referencia, se presentan técnicas de marcado de agua que generan la marca de agua en el dominio de tiempo-frecuencia.

Básicamente, una marca de agua es una firma de bit codificada que está escondida en otra señal, por ejemplo una señal de audio. La firma bit puede ser codificada empleando una pluralidad de bandas de frecuencia (subbandas).

Las Figuras 3a - 3c ilustran los valores binarios de una firma de marca de agua que han de ser codificados en diferentes bandas de frecuencias. A pesar de que una marca de agua puede ser usada para transmitir información, en lo que sigue, no enfocaremos en el aspecto de usar una marca de agua para transmitir una firma de marca de agua.

En general, una firma de marca de agua puede ser una secuencia arbitraria de bits, por ejemplo una secuencia de bits: [1, 1, -1, 1, -1, -1]. La Figura 3a ilustra una correspondiente firma de marca de agua 310.

La secuencia de bit de marca de agua es dispersada en frecuencia usando una secuencia de dispersión. Por ejemplo, para dispersar la firma de marca de agua a m bandas de frecuencia, se puede emplear un vector de dispersión que comprende m valores. En el caso más simple, un vector de dispersión solo contiene 1s, por ejemplo, se puede emplear [1, 1, 1] que da por resultado que la firma de marca de agua es copiada para cada respectiva banda de frecuencia. En una realización más sofisticada, el vector de dispersión también puede comprender -1s, por ejemplo [1, -1, 1], lo cual da por resultado que la firma de marca de agua es invertida antes de ser copiada para ser codificada en la respectiva banda de frecuencia para los respectivos bits -1. La Figura 3b ilustra la firma de marca de agua de arriba que es dispersada en frecuencia empleando un vector de dispersión [1, -1, 1, -1] para obtener una marca de agua dispersada en frecuencia 320.

La marca de agua dispersada en frecuencia también puede ser dispersada en un dominio del tiempo empleando un vector de dispersión. Por ejemplo, empleando un vector [1, -1] para dispersión temporal, cada una de las firmas de marca de agua dispersadas en frecuencia puede al principio ser codificada en su forma original, y luego puede ser invertida para ser codificada. La Figura 3c ilustra la firma de marca de agua dispersada en frecuencia y dispersada en tiempo 330 en donde se ha realizado la dispersión en tiempo empleando un vector de dispersión [1, -1].

Después de ser dispersada en frecuencia y posiblemente en tiempo, la firma de marca de agua puede ser codificada repetidamente.

Las diferentes secuencias de bits pueden ser codificadas diferencialmente. Un codificador diferencial puede realizar la codificación diferencial de los bits. Este paso da al sistema robustez adicional contra posibles corrimientos de fase debido a movimiento o faltas de coincidencia de oscilador local. Si $b(i; j)$ es el bit para la i-ésima banda de frecuencia y j-ésimo bloque de tiempo, el bit de salida $b_{\text{diff}}(i; j)$ es

$$b_{\text{diff}}(i, j) = b_{\text{diff}}(i, j - 1) \cdot b(i, j).$$

En el comienzo de la secuencia, que es para $j = 0$, $b_{\text{diff}}(i, j - 1)$ se fija en 1.

Un modulador 307 lleva a cabo la modulación actual, esto es, la generación de la forma de onda de señal de marca de agua dependiendo de la información binaria dada en su entrada. Un esquema más detallado de un correspondiente modulador está dado en la Figura 4. N_f entradas en paralelo, 401 a 40Nf contienen las secuencias para las diferentes subbandas. Cada bit de cada secuencia de subbanda es procesado por un bloque de conformación de bit (411 a 41Nf). La salida de los bloques de conformación de bit son formas de onda en el dominio del tiempo. La forma de onda generada para el j-ésimo bloque de tiempo y la i-ésima subbanda, denotada por $s_{i,j}(t)$, sobre la base del bit de entrada $b_{\text{diff}}(i, j)$ es computada como sigue:

$$s_{i,j}(t) = b_{\text{diff}}(i, j) \gamma(i, j) \cdot g_i(t - j \cdot T_b),$$

donde $\gamma(i; j)$ es un factor de ponderación provisto por una unidad de procesamiento psicoacústico, T_b es el intervalo de tiempo de bit, y $g_i(y)$ es la función formadora de bit para la i-ésima subbanda. La función formadora de bit se obtiene de una función de banda base $g_i^T(t)$ modulada en frecuencia con un coseno

$$g_i(t) = g_i^T(t) \cdot \cos(2\pi f_i t)$$

donde f_i es la frecuencia central de la i-ésima subbanda y el superíndice T significa transmisor. Las funciones de banda base pueden ser diferentes para cada subbanda. Si se eligen idénticas, es posible una implementación más eficiente en el decodificador.

El proceso de incrustación descrito arriba se refiere a un esquema de modulación BPSK, esto es, donde el bit de información codificado simplemente cambia el signo de la función formadora de bit. Alternativamente, dos o más bits pueden ser agrupados y se puede usar una constelación de modulación de orden superior (tal como QAM o M-PSK). Todavía se puede aplicar codificación diferencial, como se hace comúnmente, por ejemplo, en QPSK codificado diferencialmente.

Se repita la conformación de bit para cada bit en un proceso iterativo controlado por un módulo de procesamiento psicoacústico 102. Las iteraciones pueden ser necesarias para el ajuste fino de las ponderaciones $\gamma(i, j)$ para asignar toda la energía posible a la marca de agua mientras se la mantiene inaudible.

La forma de onda completa en la salida del i-ésimo filtro de conformación de bit 41i es

$$s_i(t) = \sum_j s_{i,j}(t).$$

La función de banda base de conformación de bit $g_i^T(t)$ es normalmente no cero durante un intervalo de tiempo mucho mayor que T_b , a pesar de que la energía principal está concentrada dentro del intervalo de bit. Se puede ver un ejemplo en la Figura 9a donde la misma función de banda base formadora de bit está graficada para dos bits adyacentes. En la Figura 9a, tenemos $T_b = 40$ ms. La elección de T_b así como también de la forma de la función afecta considerablemente al sistema. De hecho, símbolos más largos proveen respuestas de frecuencia más angostas. Esto es particularmente beneficioso en ambientes reverberantes. De hecho, en tales escenarios, la señal de marca de agua alcanza al micrófono vía varias trayectorias de propagación, cada una caracterizada por un diferente tiempo de propagación. El canal resultante exhibe fuerte selectividad de frecuencia. Interpretado en el dominio del tiempo, símbolos más largos son beneficiosos como ecos con una demora comparable a la interferencia constructiva que produce el intervalo de bit, que significa que aumentan la energía de la señal recibida. No obstante, los símbolos más largos también traen algunas pocas desventajas; superposiciones más grandes podrían conducir a interferencia intersímbolo (ISI) y seguramente son más difíciles de esconder en la señal de audio, de modo que el módulo de procesamiento psicoacústico permitiría menos energía que para símbolos más cortos.

La señal de marca de agua se obtiene sumando todas las salidas de los filtros de conformación de bit

$$\sum_i s_i(t).$$

La Figura 5 muestra un diagrama de bloques esquemático de un insertador de marca de agua 500. En el lado del codificador, se genera la señal de marca de agua 501b en el bloque de procesamiento 501 (también designado como generador de marca de agua) a partir de datos binarios 501a y sobre la base de la información 504, 505 intercambiada con la módulo de procesamiento psicoacústico 502. La información provista desde el bloque 502 típicamente garantiza que la marca de agua sea inaudible. La marca de agua generada por el generador de marca de agua 501 es agregada a la señal de audio 506.

Como ya se mencionó arriba, se necesitan factores de ponderación $\gamma(i, j)$ para calcular las formas de onda $s_{i,j}(t)$ para cada subbanda. El propósito de los factores de ponderación $\gamma(i, j)$ es esconder la marca de agua inaudible en la señal de marca de agua resultante 507. Los factores de ponderación $\gamma(i, j)$ son provistos por el módulo de procesamiento psicoacústico 502 que comprende 3 partes. La primera parte es un módulo de análisis que transforma la señal de audio del tiempo en el dominio de tiempo/frecuencia. Este módulo de análisis puede llevar a cabo análisis en paralelo en diferentes resoluciones de tiempo/frecuencia. Después del módulo de análisis, los datos de tiempo/frecuencia son transferidos al modelo psicoacústico (PAM), en el cual se calculan umbrales de enmascaramiento para la señal de marca de agua de acuerdo con consideraciones psicoacústicas. Los umbrales de enmascaramiento indican la magnitud de energía que puede ser escondida en la señal de audio para cada subbanda y bloque de tiempo. El último bloque del módulo de procesamiento psicoacústico 502 es el módulo de cálculo de amplitud. Este módulo determina las ganancias de amplitud a ser usadas en la generación de la señal de marca de agua para que se cumplan los umbrales de enmascaramiento, esto es, que la energía incrustada sea menor o igual que la energía definida por los umbrales de enmascaramiento.

En la Figura 6a se usa un único altavoz 610 para la reproducción de una señal de audio con marca de agua awm. Como un primer paso, la señal de audio con marca de agua wm es combinada con una señal de audio a mediante una unidad de combinación 605 para obtener una señal de audio con marca de agua awm. La señal con marca de agua wm puede haber sido creada como se describe con referencia a las Figuras 3a - 4. La señal con marca de agua awm puede haber sido generada mediante un generador de marca de agua como se descrito con referencia a la Figura 5. Por ejemplo, la señal con marca de agua wm puede haber sido agregada a la señal de audio a.

La señal de audio con marca de agua awm que resulta de la combinación realizada mediante la unidad de combinación 605 luego es reproducida por el altavoz 610 y registrada por los sensores 621, 622, 62N, por ejemplo, N micrófonos de una disposición ordenada de micrófonos tal que se obtienen N señales $s_1, s_2, \dots, s_N$. Las señales registradas $s_1, s_2, \dots, s_N$ no comprenden las señales de altavoz con marca de agua, sino que comprenden ruido que interfiere, por ejemplo, gente que habla en la habitación y efectos causados por reverberación. Las señales registradas $s_1, s_2, \dots, s_N$ entonces son ingresadas en el aparato 100 para proveer información de dirección de acuerdo con una realización. El aparato 100 determina y entrega información de dirección, por ejemplo, puede determinar y entregar una dirección de llegada (DoA).

Sin embargo, en otras configuraciones, se pueden adaptar múltiples altavoces 611, 612, ..., 61N para entregar múltiples señales con marca de agua $awm_1, awm_2, \dots, awm_M$. Esto se ilustra en la Figura 6b. Se agregan diferentes señales con marca de agua $wm_1, wm_2, \dots, wm_M$ a las señales de audio $a_1, a_2, \dots, a_M$, por ejemplo canales cribados

de un tema de audio usado para reproducción en *home theater*. El procesamiento para el presente escenario es similar a una situación donde sólo un altavoz está reproduciendo una señal con marca de agua.

4. Aparato para proveer Información de Dirección de acuerdo con la Figura 7:

La Figura 7 ilustra un aparato 100 para proveer información de dirección de acuerdo con una realización. Se ilustra una pluralidad de señales de audio con marca de agua $s_1, s_2, \dots, s_N$, por ejemplo, señales que pueden haber sido registradas por los micrófonos 621, 622, ..., 62N representadas en la Figura 6. Las señales con marca de agua recibidas $s_1, s_2, \dots, s_N$ que están representadas en un dominio del tiempo, son alimentadas a los módulos de análisis 711, 712, ..., 71N para ser transformadas a un dominio de tiempo-frecuencia y para ser normalizadas.

4.1 Módulos de análisis 711, 712, ..., 71N:

Para explicar los módulos de análisis 711, 712, ..., 71N, se considera una de las señales con marca de agua recibidas. Uno de los módulos de análisis transforma la señal de audio con marca de agua desde un dominio del tiempo a un dominio de tiempo-frecuencia empleando un banco de filtros de análisis. La entrada del banco de filtros es la respectiva señal de audio con marca de agua. Su salida son coeficientes complejos $b_i^{AFB}(j)$ para la i -ésima rama o subbanda en el instante j . Estos valores contienen información acerca de la amplitud y la fase de la señal en la frecuencia central f_i y el instante $j T_b$.

El banco de filtros comprende N_f ramas, una por cada subbanda espectral i . Cada rama se divide en una subbanda superior para el componente de fase de entrada y una subbanda inferior para el componente de cuadratura de la subbanda i . A pesar de la modulación en el generador de marca de agua y por ende la señal de audio con marca de agua son valores puramente reales, el análisis de valores complejos de la señal en el receptor es necesario porque se sabe que hay rotaciones de la constelación de modulación introducidas por el canal y por mala alineación de sincronización en el receptor. A continuación consideramos la i -ésima rama del banco de filtros. Combinando la subbanda de fase de entrada y cuadratura, podemos definir la señal de banda base con valores complejos $b_i^{AFB}(t)$ como

$$b_i^{AFB}(t) = r(t) \cdot e^{-j2\pi f_i t} * g_i^R(t)$$

donde $*$ indica convolución y $g_i^R(t)$ es la respuesta de impulso del filtro pasa bajos del receptor de la subbanda i . Usualmente $g_i^R(t)$ es igual a la función formadora de bit de banda base $g_i^T(t)$ de la subbanda i en el modulador para cumplir la condición de filtro apareado, pero otras respuestas de impulso son posibles también.

Para obtener los coeficientes $b_i^{AFB}(j)$ con índice $1=T_b$, se debe muestrear $b_i^{AFB}(t)$ la salida continua. Si la temporización correcta de los bits fuera conocida por el receptor, el muestreo con índice $1=T_b$ sería suficiente. Sin embargo, como la sincronización de bits no es conocida todavía, se lleva a cabo el muestreo con índice N_{os}/T_b donde N_{os} es el factor de sobremuestreo del banco de filtros de análisis. Eligiendo N_{os} suficientemente grande (por ejemplo $N_{os} = 4$), podemos asegurar que por lo menos un ciclo de muestreo es suficientemente cercano a la sincronización de bit ideal. La decisión sobre la mejor capa de sobremuestreo se toma durante el proceso de sincronización, de modo que los datos sobremuestrados se mantengan hasta entonces.

En la salida de la i -ésima rama tenemos los coeficientes $b_i^{AFB}(j, k)$, donde j indica el número de bit o instante de tiempo y k indica la posición de sobremuestreo dentro de este bit singular, donde $k = 1; 2; \dots, N_{os}$.

La Figura 8a da un panorama ejemplar de la ubicación de los coeficientes en el plano de tiempo-frecuencia. En la Figura 8a, el factor de sobremuestreo es $N_{os} = 2$. La altura y el ancho de los rectángulos indican, respectivamente, el ancho de banda y el intervalo de tiempo de la parte de la señal que está representada por el correspondiente coeficiente $b_i^{AFB}(j, k)$.

Si las frecuencias subbanda f_i son elegidas como múltiples de un cierto intervalo Δf , el banco de filtros de análisis puede ser implementado eficientemente usando la Transformada de Fourier rápida (FFT).

La normalización de los coeficientes complejos obtenidos para cada banda de frecuencia se puede realizar como sigue: Sin perder generalidad y para simplificar la descripción, suponemos en lo que sigue, que la sincronización de bit es conocida y que $N_{os} = 1$. Antes de la normalización, tenemos los coeficientes complejos $b_i^{AFB}(j)$. Como no hay disponible información de estado de canal (esto es, el canal de propagación es desconocido), se usa un esquema de combinación de ganancia igual (EGC). Debido al canal dispersivo de tiempo y frecuencia, la energía del

bit enviado $b_i(j)$ no sólo se halla alrededor de la frecuencia central f_i y el instante de tiempo j , sino también en frecuencias e instantes de tiempo adyacentes. Por lo tanto, para una ponderación más precisa, se calculan coeficientes adicionales en las frecuencias $f_i \pm n \Delta f$ y se usan para la normalización del coeficiente $b_i^{\text{AFB}}(j)$. Si $n = 1$ tenemos, por ejemplo,

$$b_i^{\text{norm}}(j) = \frac{b_i^{\text{AFB}}(j)}{\sqrt{1/3 \cdot (|b_i^{\text{AFB}}(j)|^2 + |b_{i-\Delta f}^{\text{AFB}}(j)|^2 + |b_{i+\Delta f}^{\text{AFB}}(j)|^2)}}$$

La normalización para $n > 1$ es una extensión directa de la fórmula de arriba. La normalización es llevada a cabo para cada subbanda i y cada instante de tiempo j .

Asimismo, el factor de normalización puede ser computado para un micrófono solo y luego aplicado a todos los micrófonos, en caso de que la dirección de estimación de llegada usa también la amplitud de la información específica de receptor.

La Figura 8b ilustra una señal subbanda, en donde dos porciones de señal con marca de agua WM1 y WM2 son multiplexadas en tiempo. La señal subbanda comprende la porción de sonido de una primera fuente de audio LS1 proveniente de un primer altavoz y la porción de sonido de una segunda fuente de sonido LS2 proveniente de un segundo altavoz. Asimismo, la señal subbanda también comprende, o bien la porción de señal de marca de agua WM1 proveniente de un primer altavoz, o la porción de señal de marca de agua WM2 proveniente de un segundo altavoz.

4.2 Módulo de Selección 720:

Volviendo a la Figura 7, después de la transformación en un dominio de tiempo-frecuencia y la normalización, las señales del dominio de tiempo-frecuencia $S_1, S_2, \dots, S_N$ son alimentadas a un módulo de selección 720. El módulo de selección 720 determina cuáles de las señales disponibles $S_1, S_2, \dots, S_N$ deben ser usadas para realizar la subsiguiente sincronización. Existen varias posibilidades para el módulo de selección 720 para elegir entre las señales. En una realización, el módulo de selección 720 está adaptado para elegir una de las señales $S_1, S_2, \dots, S_N$ aleatoriamente. Por ejemplo, el módulo de selección podría seleccionar la señal del dominio de tiempo-frecuencia S_1 . En una realización alternativa, el módulo de selección 720 está adaptado para promediar las señales $S_1, S_2, \dots, S_N$ y usar la señal promediada.

4.3 Decodificador Diferencial 730:

El módulo de selección 720 entonces alimenta la señal resultante al decodificador diferencial 730. En la entrada del decodificador diferencial 730 tenemos coeficientes complejos normalizados de amplitud $b_i^{\text{norm}}(j)$ los cuales contienen información acerca de la fase de los componentes de la señal en la frecuencia f_i y el instante de tiempo j . Como los bits son codificados diferencialmente en el transmisor, aquí se debe realizar la operación inversa. Los bits blandos $\hat{b}_i(j)$ se obtienen calculando primero la diferencia de fase de dos coeficientes consecutivos y luego tomando la parte real:

$$\begin{aligned} \hat{b}_i(j) &= \text{Re}\{b_i^{\text{norm}}(j) \cdot b_i^{\text{norm}*}(j-1)\} \\ &= \text{Re}\{|b_i^{\text{norm}}(j)| \cdot |b_i^{\text{norm}}(j-1)| \cdot e^{j(\varphi_j - \varphi_{j-1})}\} \end{aligned}$$

Esto tiene que ser llevado a cabo separadamente para cada subbanda porque el canal normalmente introduce diferentes rotaciones de fase en cada subbanda.

El decodificador diferencial 730 decodifica diferencialmente la señal ingresada para obtener los denominados bits blandos calculando primero la diferencia de fase de dos coeficientes consecutivos y luego tomando la parte real de la fase. Se debe notar que los bits blandos resultantes pueden tomar cualquier valor real y todavía no se toma ninguna decisión definitiva sobre el bit. Como los bits blandos pueden tomar cualquier valor, no necesariamente tienen que ser -1 o 1, pero en cambio pueden tener un valor real diferente de -1 o 1, tal como, por ejemplo, 0,92; -0,88; 0,7, y valores similares.

Asimismo, cuando se ingresa en el decodificador diferencial, la señal subbanda podría ser sobremuestreada. El decodificador diferencial toma en cuenta sobremuestreo cuando decodifica diferencialmente la señal subbanda. Por ejemplo, considerar el caso donde una señal subbanda ha sido sobremuestreada por un factor de sobremuestreo q .

Por ejemplo, tenemos para la i -ésima subbanda los coeficientes $b_i^{norm}(j, q)$, donde j indica el número de bit o instante de tiempo y q indica la posición de sobremuestreo dentro de este bit singular, donde $q = 1; 2; \dots$, Nos. Entonces, con un factor de sobremuestreo q , calculamos q bits blandos aplicando:

$$\hat{b}_i(j, q) = \text{Re} \{ b_i^{norm}(j, q) \cdot b_i^{norm*}(j-1, q) \}$$

El decodificador diferencial 730 luego alimenta a un módulo de integración 740 con la secuencia de bits blandos, el cual integra la secuencia de bits blandos en el tiempo a través de varias firmas de sincronización para mejorar adicionalmente la precisión de la sincronización.

4.4 Módulo de Integración 740:

Ahora se explicará en más detalle el módulo de integración 740. Cuando se emplea sólo una firma de marca de agua wm1 producida por un solo altavoz, la firma de marca de agua wm1 es transmitida repetidamente tal que cuando el último bit de la firma de marca de agua wm1 es transmitido, sigue a continuación el primer bit de la firma de marca de agua wm1, y así de seguido. La Figura 9a representa una correspondiente secuencia. Si la única firma de marca de agua empleada (por ejemplo dispersada en tiempo) wm1 comprende N bits, entonces la secuencia de bits se repite después de N bits, y la firma de marca de agua única forma el bloque de sincronización sync de longitud N bits.

Sin embargo, como ya se explicó arriba, un sistema puede comprender más de un altavoz. Para obtener la estimación de la dirección de llegada de cada altavoz, es necesario determinar cuál parte de la señal registrada fue emitida por un cierto altavoz. Por lo tanto, las señales de marca de agua tienen que ser multiplexadas. Hay diferentes métodos posibles, los cuales son muy conocidos en comunicaciones, a saber, acceso múltiple por división de código (CDMA), acceso múltiple por división de frecuencia (FDMA) y acceso múltiple por división de tiempo (TDMA). En una realización se emplea acceso múltiple por división de tiempo. Para determinar cuáles partes de la señal registrada pertenecen a un cierto altavoz, se incrustan diferentes firmas de marca de agua en diferentes señales de altavoz. La Figura 10 ilustra la multiplexación de tiempo para el uso de dos altavoces y firmas de sincronización que comprenden seis secuencias.

Así, diferentes altavoces pueden transmitir diferentes marcas de agua, por ejemplo, las marcas de agua wm1 wm2, las cuales pueden ser multiplexadas en tiempo. Supongamos que el número de bits de ambas marcas de agua (por ejemplo, dispersadas en tiempo) es N . Al principio, se transmite la primera firma de marca de agua wm1. Luego, se transmite la segunda firma de marca de agua. Entonces, se transmite nuevamente la primera firma de marca de agua wm1, seguida por la segunda firma de marca de agua wm2 y así de seguido. El bloque de sincronización sync entonces comprende las marcas de agua wm1, wm2 y tiene $2 \cdot N$ bits. La Figura 9b representa el escenario descrito con dos firmas de marca de agua multiplexadas.

Cada bloque de sincronización es transmitido repetidamente. Suponiendo que la secuencia de sincronización tiene N_{sync} bits, entonces bit 1 es igual a bit $N_{sync} + 1$, e igual a bit $2 \cdot N_{sync} + 1$, y así de seguido. Bit 2 es igual a bit $N_{sync} + 2$, e igual a bit $2 \cdot N_{sync} + 2$, y así de seguido. La fórmula resultante para una subbanda i considerada es:

$$\hat{b}_i(j) = \hat{b}_i(j + N_{sync}) = \hat{b}_i(j + 2 \cdot N_{sync}) = \dots = \hat{b}_i(j + K \cdot N_{sync}), \quad \text{para } j = 0, \dots, N_{sync} - 1.$$

En el caso, se ha realizado sobremuestreo mediante los módulos de análisis 711, 712, ..., 71N, por ejemplo con factor de sobremuestreo Nos , luego se han creado valores de bit decodificados diferencialmente Nos candidatos

$\hat{b}_i(j, k)$ para cada bit singular $\hat{b}_i(j)$ de una subbanda i considerada. Resulta la siguiente ecuación:

$$\hat{b}_i(j, q) = \hat{b}_i(j + N_{sync}, q) = \hat{b}_i(j + 2 \cdot N_{sync}, q) = \dots = \hat{b}_i(j + K \cdot N_{sync}, q), \quad \text{para } j = 0, \dots, N_{sync}-1,$$

en donde $q = 1, 2, \dots, Nos$ indica la posición de sobremuestreo dentro del bit considerado.

El módulo de integración está adaptado para integrar los bits blandos tal que una pluralidad de bits blandos de diferentes bloques de sincronización sean sumados para mejorar la robustez del proceso de decodificación de marca de agua.

Al principio, la secuencia de bits que comprende los bits decodificados diferencialmente es dividida en bloques periódicos subsiguientes $p1, p2, p3$ que tienen N_{sync} bits. En este punto, el módulo de integración 740 no se da cuenta dónde empieza un bloque de sincronización. Sin embargo, esto no es necesario, ya que la periodicidad de la

secuencia de bits decodificada diferencialmente no depende del punto de inicio del bloque de sincronización, sino en cambio, depende de la longitud del bloque de sincronización. Para eso, el primer bit de un bloque de sincronización puede ser ubicado en algún lugar dentro del bloque periódico. La Figura 9c ilustra este escenario.

Para formar un bloque de firma sincronización promediado p_0 , se agregan correspondientes bits de diferentes bloques periódicos. Por ejemplo, se agrega el primer bit del primero, segundo y tercer periodo de marca de agua, se agrega el segundo bit del primero, segundo y tercer periodo de marca de agua, y así de seguido hasta el bit N_{sync} del primero, segundo y tercer periodo de marca de agua. Cada bit resultante puede ser dividido por el número de bloques periódicos que son promediados. Esto da por resultado la siguiente fórmula:

$$b_i^{\text{avg}}(j) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{b}(j + k \cdot N_{\text{sync}}), \text{ donde } j = 0, 1, 2, \dots, N_{\text{sync}}-1.$$

En esta fórmula, i es la subbanda considerada, K representa el número de bloques de sincronización que son promediados, k representa el bloque de sincronización actual que se considera, mientras j es la posición de bit

considerada dentro de los bloques periódicos. El bloque de sincronización promediado tiene N_{sync} bits: $b_i^{\text{avg}}(0)$, $b_i^{\text{avg}}(1)$, $b_i^{\text{avg}}(2)$, ..., $b_i^{\text{avg}}(N_{\text{sync}} - 1)$.

El primer bit de la primera firma de marca de agua puede ser ubicado en algún lugar en el bloque de sincronización promediado p_0 . La Figura 9d ilustra el escenario mostrando un bloque de sincronización promediado p_0 .

Si el análisis en los módulos de análisis 711, 712, ... 71N ha sido realizado con un factor de sobremuestreo de N_{os} , entonces también se toma en cuenta el sobremuestreo. En este caso, se generan valores promediados para todo valor de bit decodificado diferencialmente N_{os} candidato $\hat{b}(i, q)$ para cada bit de los bloques periódicos. Esto da por resultado la fórmula:

$$b_i^{\text{avg}}(j, q) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{b}(j + k \cdot N_s, q), \text{ donde } j = 0, 1, 2, \dots, N_{\text{sync}}-1$$

en donde $q = 1, 2, \dots, N_{\text{os}}$ indica la posición de sobremuestreo dentro del bit considerado.

La integración realizada por el módulo de integración 740 produce bits blandos promediados o en caso de sobremuestreo, una pluralidad de N_{os} valores de bits blandos candidatos para cada uno de los bits del bloque de sincronización promediado p_0 .

En otras realizaciones donde señales de voz actúan como portadoras de marca de agua, la integración es particularmente importante, ya que con frecuencia contienen huelgos sin señal con marca de agua. Por eso, normalmente se reducirán errores existentes en la rehíla de bits blandos cuando se realiza integración en tiempo.

4.5 Módulo de Sincronización 750:

Después que se ha integrado la secuencia de bits en el tiempo, el módulo de integración 740 alimenta la secuencia de bits blandos integrada al módulo de sincronización 750 el cual realiza una sincronización robusta. En más detalle, el módulo de sincronización 750 realiza una correlación con la firma de sincronización incrustada. Lo que resulta de esto es una posición de impacto de sincronización. Ahora se explicará en más detalle la sincronización en el módulo de sincronización 750.

La tarea del módulo de sincronización 750 es hallar la alineación temporal de la una o más marcas de agua dentro del bloque de sincronización promediado.

El problema de sincronizar el decodificador a los datos codificados es doble. En un primer paso, el banco de filtros de análisis debe ser alineado con los datos codificados, a saber, las funciones de conformación de bits $g_i^T(t)$ usadas en la síntesis en el modulador deben ser alineadas con los filtros $g_i^R(t)$ usados para el análisis. Este problema se ilustra en la Figura 12a, donde los filtros de análisis son idénticos a los de síntesis. En la parte superior hay tres bits visibles. Por razones de simplicidad, las formas de onda para los tres bits no están a escala. El desplazamiento temporal entre diferentes bits es T_b . La parte inferior ilustra el problema de sincronización en el decodificador: el filtro puede ser aplicado a diferentes instantes de tiempo, sin embargo, sólo la curva 1299a es correcta y permite extraer el primer bit con la mejor relación señal a ruido SNR y relación señal a interferencia SIR.

De hecho, una alineación incorrecta conduciría a una degradación tanto de la SNR como de la SIR. Nos referimos a este primer problema de alineación como "sincronización de bit". Una vez que se ha logrado la sincronización de bit, se puede extraer bits en forma óptima. Sin embargo, también puede ser necesario conocer dónde empieza una firma de marca de agua. El problema está ilustrado en la Figura 12b y se lo refiere como sincronización de mensaje. En la

secuencia de bits decodificada sólo la posición inicial marcada por una flecha gruesa 1299b es correcta y permite decodificar el k-ésimo mensaje.

Para explicar el concepto de sincronización básico primero nos referimos a un sistema en donde sólo un altavoz entrega una sola firma de marca de agua. Por lo tanto, el bloque de sincronización promediado sólo comprende los bits de la única firma de marca de agua. La firma de marca de agua también puede ser denominada como firma de sincronización.

Asimismo, nos referiremos al término secuencia de sincronización. Una firma de sincronización puede comprender una pluralidad de secuencias de sincronización. En el caso más simple, la secuencia de sincronización puede comprender un solo bit. Sin embargo, también es posible que una firma de sincronización comprenda una pluralidad de bits.

Primero apuntamos a sincronización de mensaje. Se supone que la firma de sincronización está compuesta por N_s secuencias de sincronización en un orden predeterminado las cuales están incrustadas continuamente y periódicamente en la señal de audio con marca de agua. El módulo de sincronización 750 es capaz de recuperar la alineación temporal de las secuencias de sincronización.

El módulo de sincronización puede identificar el comienzo de cada firma de sincronización hallando la alineación temporal de las secuencias de sincronización. Nos referimos a las posiciones temporales en las cuales empieza una nueva firma de sincronización como impactos de sincronización.

Los bloques de procesamiento del módulo de sincronización están representados en las Figuras 11a y 11b. La unidad de soporte de sincronización lleva a cabo la sincronización de bits y la sincronización de mensaje a la vez analizando la salida del correlacionador de firma de sincronización 1201. Los datos en el dominio de tiempo/frecuencia 204 son provistos por el módulo de análisis. Como todavía la sincronización de bits no está disponible, el módulo de análisis sobremuestrea los datos con factor N_{os} .

Una ilustración de los datos de entrada se da en la Figura 12c. Para este ejemplo hemos tomado $N_{os} = 4$, $N_t = 2$, y $N_s = 3$. Nuevamente, N_{os} indica el factor de sobremuestreo, N_t indica el factor de dispersión en el tiempo y N_s indica el número de secuencias de sincronización. Así, en otras palabras, la firma de sincronización consiste en 3 secuencias (denotadas con a, b y c). La dispersión de tiempo, en este caso con secuencia de dispersión $c_t = [1 \ 1]^T$, simplemente repite cada bit dos veces en el dominio del tiempo.

Los impactos de sincronización exactos son denotados con flechas y corresponden al comienzo de cada firma de sincronización. El periodo de la firma de sincronización es $N_t \cdot N_{os} \cdot N_s = N_{sbl}$ que es $2 \cdot 4 \cdot 3 = 24$.

Debido a la periodicidad del bloque de sincronización promediado, es suficiente considerar sólo el bloque de sincronización promediado como es entregado por el módulo de integración 740. El bloque de sincronización debe contener (o típicamente contiene) un impacto de sincronización como se representa en la Figura 12d. Cada uno de los bits del bloque de sincronización promediado es un impacto de sincronización candidato. En la Figura 12d, el bloque de sincronización promediado está ilustrado en forma duplicada en las figuras sólo con propósitos ilustrativos. La tarea del bloque 1201 es computar una medida de probabilidad para cada bit candidato del bloque de sincronización. Esta información luego es pasada al bloque 1204 el cual computa los impactos de sincronización.

Para cada una de las posiciones de sincronización candidato N_{sbl} , el correlacionador de firma de sincronización computa una medida de probabilidad, cuanto más grandes es esta última, más probable es que se haya hallado la alineación temporal (tanto sincronización de bit como mensaje parcial o completo). Los pasos de procesamiento están representados en la Figura 11c.

Consiguientemente, se puede obtener una secuencia 1204a de valores de probabilidad, asociada con diferentes elecciones posicionales. El bloque 1301 lleva a cabo dispersión temporal, esto multiplica todos los bits N_t con la secuencia de dispersión temporal c_t y luego los suma. Esto es llevado a cabo para cada una de las N_t subbandas de frecuencia.

La Figura 13a muestra un ejemplo. Tomamos los mismos parámetros que se describieron en la sección previa, a saber $N_{os} = 4$, $N_t = 2$, y $N_s = 3$. La posición de sincronización candidato está marcada. Desde ese bit, con desplazamiento N_{os} , $N_t \cdot N_s$ son tomados por el bloque 1301 y son dispersados en tiempo con la secuencia c_t , de modo que se dejan los bits N_s .

En el bloque 1302 los bits son multiplicados elemento por elemento con las secuencias de dispersión N_s (ver Figura 13b).

En el bloque 1303 se lleva a cabo la desdispersión de frecuencia, a saber, cada bit es multiplicado con la secuencia de dispersión c_r y luego sumados a lo largo de la frecuencia.

En este punto, si la posición de sincronización fuera correcta, tendríamos N_s bits decodificados. Como los bits no son conocidos para el receptor, el bloque 1304 computa la medida de probabilidad tomando los valores absolutos de los N_s valores y suma.

La salida del bloque 1304 es en principio un correlacionador no coherente que busca la firma de sincronización. De hecho, cuando se elige un N_s pequeño, a saber el modo de sincronización de mensaje parcial, no es posible usar secuencias de sincronización (por ejemplo, a, b, c) que sean mutuamente ortogonales. Al hacerlo, cuando el correlacionador no está correctamente alineado con la firma, su salida será muy pequeña, idealmente cero. Cuando se usa el modo de sincronización de mensaje completo se aconseja usar tantas secuencias de sincronización ortogonales como sea posible, y luego crear una firma eligiendo cuidadosamente el orden en el cual se las usa. En este caso, se puede aplicar la misma teoría que cuando se buscan secuencias de dispersión con buenas funciones de autocorrelación. Cuando el correlacionador está sólo levemente mal alineado, entonces la salida del correlacionador no será cero aun en el caso ideal, pero de todos modos será más pequeña comparada con la alineación perfecta, ya que los filtros de análisis no pueden capturar la energía de la señal en forma óptima.

Este bloque de computación de impactos de sincronización analiza la salida del correlacionador de firma de sincronización para decidir dónde están las posiciones de sincronización. Como el sistema es bastante robusto contra errores de alineación de hasta $T_b/4$ y el T_b normalmente se toma alrededor de 40 ms, es posible integrar la salida de 1201 en el tiempo para lograr una sincronización más estable. Una implementación posible de esto está dada por un filtro IIR aplicado a lo largo del tiempo con una respuesta de impulso que decae exponencialmente. Alternativamente, se puede aplicar un filtro de promedio móvil FIR tradicional. Una vez que se llevó a cabo el promediado, se lleva a cabo una segunda correlación a lo largo de diferentes $N_t \cdot N_s$ ("diferente elección posicional"). De hecho, queremos aprovechar la información que la función de autocorrelación de la función de sincronización es conocida. Esto corresponde a un estimador de máxima probabilidad. La idea se muestra en la Figura 13c. La curva muestra la salida del bloque 1201 después de integración temporal. Una posibilidad para determinar el impacto de sincronización es simplemente hallar el máximo de esta función. En la Figura 13d vemos la misma función (en negro) filtrada con la función de autocorrelación de la firma de sincronización. La función resultante está indicada por la línea más ancha. En este caso el máximo es más pronunciado y nos da la posición del impacto de sincronización. Los dos métodos son bastante similares para SNR alta pero el segundo método se desempeña mucho mejor en regímenes de SNR más bajo. Una vez que se hallaron los impactos de sincronización, son pasados al extractor de marca de agua 202 el cual decodifica los datos.

Más de una firma de marca de agua:

Un sistema también puede emplear dos o más firmas de marca de agua entregadas por diferentes altavoces, en donde las firmas de marca de agua son multiplexadas en el tiempo. La Figura 10 ilustra tales firmas de marca de agua multiplexadas en el tiempo. En un caso así, el módulo de sincronización 750 revisa separadamente cada una de las firmas de marca de agua incrustadas. Por lo tanto, detecta tantas posiciones de impacto de sincronización como altavoces y diferentes firmas de marca de agua se usan y suministra estas posiciones de impacto de sincronización al estimador de dirección 760.

De acuerdo con esto, el concepto descrito arriba para hallar una posición de impacto de sincronización de una firma de sincronización, es aplicado repetidamente para cada marca de agua. Así, para cada marca de agua, se determina una posición de impacto de sincronización dentro del bloque de sincronización promediado. La determinación de la posición de impacto de sincronización de cada marca de agua se realiza repetidamente como se explicó con referencia a las Figuras 11a - 13d para cada una de las marcas de agua, en donde la firma de marca de agua de la marca de agua considerada es empleada como firma de sincronización.

La señal registrada luego es dividida en M segmentos, perteneciendo cada segmento M a una cierta marca de agua de un cierto altavoz. La estimación de la dirección de llegada luego es llevada a cabo para segmento. Como resultado, el estimador de dirección 760 tendrá M salidas, representando cada una la dirección de llegada del sonido emitido por un cierto altavoz.

El módulo de sincronización 750 provee la una o más posiciones de impacto de sincronización dentro del bloque de sincronización promediado y suministra la posición de impacto de sincronización a un estimador de dirección 760.

En otras realizaciones, puede no ser necesario que los impactos de sincronización sean calculados si la marca de agua está incrustada en forma continua y sólo se usa un altavoz. Si este no es el caso, se ha de realizar la sincronización para hallar el punto de inicio para la estimación de dirección.

4.6 Módulos de Integración 741, 742, ..., 74N:

Volviendo a la Figura 7, los módulos de análisis 711, 712, ..., 71N no sólo proveen las señales del dominio de tiempo-frecuencia $S_1, S_2, \dots, S_N$ al módulo de selección 720, sino también proveen las señales $S_1, S_2, \dots, S_N$ a los módulos de integración 741, 742, ..., 74N. Lo que se suministra a los módulos de integración 741, 742, ..., 74N de

hecho son los coeficientes complejos normalizados (sobremuestreados) $b_i^{norm}(j, q)$ los cuales son los mismos coeficientes que los provistos al módulo de selección 720.

Sin embargo, no se realiza decodificación diferencial antes de que los coeficientes complejos normalizados sobremuestreados sean alimentados a los módulos de integración 741, 742, ..., 74N, ya que es necesaria la información de fase comprendida en los coeficientes complejos, la cual se perdería después de una decodificación diferencial. En cambio, la información de fase de la señal tiene que ser mantenida. Aplicando decodificación diferencial, esta información sería descartada.

En cada uno de los módulos de integración 741, 742, ..., 74N y para cada subbanda i , al principio, la secuencia que comprende los coeficientes complejos normalizados sobremuestreados es dividida en bloques periódicos subsiguientes $p11, p12, p13$. La secuencia de coeficientes complejos es dividida exactamente de la misma manera en bloques como la secuencia de bits decodificados diferencialmente ha sido separada en los bloques periódicos $p1, p2, p3$. Bits y muestras con correspondientes índices de tiempo y (posiciones sobremuestreadas) están comprendidos en correspondientes bloques periódicos.

Esto se explica con referencia a las Figuras 7 y 14. En la Figura 7, una señal $s1$ puede ser provista al módulo de análisis 711. El módulo de análisis 711 genera una señal del dominio de tiempo-frecuencia $S1$ que comprende una pluralidad de señal subbanda, estando representada cada una de las señales subbanda como una secuencia de coeficientes complejos. En una realización, el módulo de selección 720 puede seleccionar $S1$ tal que la subsiguiente sincronización sea realizada en base a la señal del dominio de tiempo-frecuencia $S1$.

En la Figura 14 se considera una particular subbanda de la señal del dominio de tiempo-frecuencia $S1$. Esta subbanda puede comprender coeficientes complejos $c(i, j)$, en donde i denota la subbanda y j denota el índice de tiempo. Por ejemplo, considere los coeficientes complejos $c(i, 1), c(i, 2), c(i, 3), c(i, 4), c(i, 5)$ y $c(i, 6)$ de esta subbanda. La decodificación diferencial en base a estos valores mediante el decodificador diferencial provee seis bits blandos, a saber $d(i, 1)$ en base a $c(i, 1)$ y la muestra de subbanda precedente; $d(i, 2)$ en base a $c(i, 2)$ y la precedente $c(i, 1)$; $d(i, 3)$ en base a $c(i, 3)$ y la precedente $c(i, 2)$; $d(i, 4)$ en base a $c(i, 4)$ y la precedente $c(i, 3)$; $d(i, 5)$, en base a $c(i, 5)$ y la precedente $c(i, 4)$; y $d(i, 6)$ en base a $c(i, 6)$ y la precedente $c(i, 5)$. En $d(i, j)$, i representa la subbanda y j representa el índice de tiempo.

Los coeficientes complejos de la banda de frecuencia i también son alimentados al módulo de integración 741 de la Figura 7. Si el módulo de integración 740 ahora divide la secuencia de bits decodificados diferencialmente que comprende los bits blandos $d(i, 1), d(i, 2), d(i, 3), d(i, 4), d(i, 5), d(i, 6)$ en un primer bloque periódico $p1$ que comprende $d(i, 1), d(i, 2)$ y $d(i, 3)$ y en un segundo bloque periódico $p2$ que comprende $d(i, 4), d(i, 5)$ y $d(i, 6)$ entonces los coeficientes complejos $c(i, 1), c(i, 2), c(i, 3), c(i, 4), c(i, 5)$ y $c(i, 6)$ también son divididos en dos bloques, esto es, un primer bloque que comprende los coeficientes complejos $c(i, 1), c(i, 2)$ y $c(i, 3)$ y en un segundo bloque que comprende los coeficientes complejos $c(i, 4), c(i, 5)$ y $c(i, 6)$ (ver Figura 14).

Lo mismo se aplica para otras señales del dominio de tiempo-frecuencia S_N generadas por otros módulos de análisis 71N incluso si estas señales no han sido usadas para la decodificación diferencial realizada por el decodificador diferencial.

El concepto descrito arriba es igualmente aplicable si se toma en cuenta sobremuestreo. Si los bits blandos sobremuestreados $b(i, j, q)$ son divididos en un primer y un segundo bloque periódico $p1$ y $p2$, entonces los coeficientes complejos $c(i, j, q)$ también son divididos en dos bloques $p11$ y $p12$. (Aquí, i indica la subbanda, j indica el índice de tiempo y q indica la posición de sobremuestreo.) Si el primer bloque periódico $p1$ comprende el bit blando $b(i, j, q)$, entonces el primer bloque $p11$ comprende el coeficiente complejo $c(i, j, q)$ con los mismos índices j, q . Si el segundo bloque periódico $p2$ comprende el bit blando $b(i, j, q)$, entonces el segundo bloque $p12$ también comprende el coeficiente complejo $c(i, j, q)$ con los mismos índices j, q .

Así, la longitud de cada uno de los bloques periódicos $p11, p12, p13$ generados por los módulos de integración 741, 742, ..., 74N corresponde a la longitud de los bloques periódicos $p1, p2, p3$ generados por el módulo de integración 740.

Para formar un bloque periódico promediado para cada subbanda i , se promedia correspondientes coeficientes complejos. Por ejemplo, se suman correspondientes coeficientes complejos de diferentes bloques periódicos. Por ejemplo, el primer coeficiente complejo del primer, segundo y tercer período de sincronización son sumados, el segundo coeficiente complejo de los primer, segundo y tercer período de marca de agua son sumados, y así de

seguido. Cada coeficiente complejo resultante puede ser dividido por el número de bloques periódicos que son promediados. Esto da por resultado la siguiente fórmula:

$$s_i^{avg}(j) = \frac{1}{K} \sum_{k=0}^{K-1} b_i^{norm}(j + k \cdot N_s), \text{ donde } j = 0, 1, 2, \dots, N_{sync}-1.$$

En esta fórmula, K representa el número de bloques periódicos de una subbanda i considerada que son promediados, k representa el bloque periódico actual que se considera, mientras j es la posición del coeficiente complejo dentro de los bloques periódicos.

Tomando sobremuestreo con un factor de sobremuestreo N_{os} en cuenta, esto da por resultado la fórmula:

$$s_i^{avg}(j, q) = \frac{1}{K} \sum_{k=0}^{K-1} b_i^{norm}(j + k \cdot N_s, q), \text{ donde } j = 0, 1, 2, \dots, N_{sync}-1$$

en donde $q = 1, 2, \dots, N_{os}$ indica la posición de sobremuestreo dentro del coeficiente complejo considerado.

Se determina un bloque periódico promediado para cada banda de frecuencia y cada señal.

Este efecto de formar bloques periódicos promediados se explica con referencia a la Figura 15a y a la Figura 15b. Las Figuras 15a y 15b ilustran cómo ayuda la integración a adquirir la información de fase deseada de la señal de marca de agua.

Cuando se usan varios altavoces, aparece un fenómeno conocido, llamado fuente de imagen fantasma. En este escenario, el sonido emitido es percibido como si fuera reproducido por una fuente ubicada entre los altavoces. Este también es el caso para la estimación de la dirección de llegada, ya que la parte de audio de la señal de marca de agua registrada se duplica proporcionalmente a la parte de marca de agua. Una solución para este problema es el uso de módulos de integración 441, 442, ..., 44N, lo que ahora se explica en más detalle.

La Figura 15a muestra el vector complejo de la parte de audio $\hat{A}_n = \sum_{m=1}^M A_{mn}$ y de la parte de señal de

marca de agua $\widehat{WM}_n$ de una señal registrada y analizada S_n para una cierta frecuencia subbanda l y diferentes ranuras de tiempo i suponiendo un canal ideal. Aquí, A_{mn} representa la parte de audio emitida por el altavoz m y recibida en el micrófono n. Se debería notar que en el ejemplo TDMA, cuando mirando las ranuras de tiempo dadas, la señal registrada comprende la superposición de todas las señales de audio emitidas juntas con sólo una señal de marca de agua. K representa el número de bloques de sincronización a ser integrados. N_{synch} representa el número de ranuras de tiempo para un bloque de sincronización completo. Comprende las diferentes firmas de sincronización para todos los altavoces:

$$N_{synch} = M \cdot N_{seq},$$

donde M es el número de altavoces y N_{seq} es el número de secuencias de una firma de sincronización. El vector complejo de la porción de audio rota desde una ranura de tiempo a otra. Por el contrario, la fase de la señal de marca de agua permanece sin embargo casi igual.

En la Figura 15b se ilustra la situación después que la integración fue realizada. El vector complejo de la señal de

audio resultante $\frac{1}{K} \sum_{k=0}^{K-1} \hat{A}_n(l, i + k \cdot N_{synch})$ casi desaparece, mientras que el vector de la marca de agua obtenido

$\frac{1}{K} \sum_{k=0}^{K-1} \widehat{WM}_n(l, i + k \cdot N_{synch})$ permanece. Por lo tanto, los métodos que logran una estimación del ángulo de incidencia usando la información de fase de las señales de entrada registradas reciben datos usables para obtener una estimación factible para la dirección de llegada de cada altavoz.

Después de la integración de bit, hay disponible un coeficiente complejo integrado para cada señal recibida $S_1, S_2, \dots, S_N$ de los diferentes micrófonos 611, 612, ..., 61N y para cada banda de frecuencia i.

Por el momento, no tomaremos en cuenta sobremuestreo y explicamos algunos principios sobre los cuales se basa La presente invención. La modulación de fase como se emplea en el contexto de esta realización se basa en algunas ideas fundamentales: Suponiendo que cada coeficiente complejo muestreado corresponde a una posición de bit codificada, el valor de fase de la muestra actual es comparado con el valor de fase de la muestra previa. Considere por ejemplo, manipulación por desplazamiento de fase binaria (BPSK). Si los valores de fase de ambas muestras son idénticos, esto corresponde a un primer valor de bit codificado, por ejemplo, el valor de bit codificado

es 1. Sin embargo, si ambos valores de fase difieren, por ejemplo por 180° ($0 \pm \pi$), esto corresponde a un segundo valor de bit codificado, por ejemplo, el valor de bit codificado es -1. Así, para emplear modulación de fase, por ejemplo, manipulación por desplazamiento de fase binaria, es esencial que las muestras de las posiciones de bit subsiguientes tengan, o bien el mismo valor de fase, o un valor de fase que difiera por 180° ($\pm \pi$). Por lo tanto, cuando se codifican valores modulados por fase, se emplea una modulación, tal que los valores de fase de las muestras correspondientes a posiciones de bit subsiguientes, o bien son iguales, o bien difieren por 180° dependiendo del valor de bit codificado.

Con respecto a una cierta firma de marca de agua y una cierta banda de frecuencia, un altavoz transmite una secuencia de bits codificada mediante modulación de fase. Todos los receptores de audio, por ejemplo, micrófonos reciben la secuencia de bits modulada por fase. La diferencia de fase de las muestras correspondientes a posiciones de bit subsiguientes de las señales de audio recibidas $S_1, S_2, \dots, S_N$ deben ser (casi) iguales para todas las señales de diferentes micrófonos. Si dos micrófonos reciben la misma porción de la señal de audio al mismo tiempo, los valores absolutos de fase también deben ser idénticos. Sin embargo, como los micrófonos están separados (a pesar de que pueden estar sólo poco separados), un primer micrófono puede recibir la misma porción de la señal de audio un poco antes que un segundo micrófono, por ejemplo, si el primer micrófono está más cerca del altavoz que el segundo micrófono. Una tal diferencia en tiempo de dos señales del dominio del tiempo recibidas da por resultado una diferencia de fase de dos señales del dominio de la frecuencia recibidas, por ejemplo, señales del dominio de tiempo-frecuencia. Por lo tanto, una diferencia de fase de dos señales de marca de agua recibidas por dos micrófonos comprende información acerca de la posición del altavoz que emitió la señal de marca de agua relativa a la ubicación de los micrófonos.

4.7 Estimador de Dirección 760:

Los módulos de integración 741, 742, ..., 74N alimentan los bloques periódicos promediados determinados para banda de frecuencia y cada señal en el estimador de dirección 760 el cual ya ha recibido la posición de impacto de sincronización. El estimador de dirección 760 entonces realiza la estimación de la dirección de llegada sobre la parte de los datos entrantes señalados por la posición de impacto de sincronización para cada bloque periódico promediado.

En lo que sigue, nos enfocaremos en la situación de que un sistema sólo comprende un altavoz que transmite una sola marca de agua. Sin embargo, la invención es igualmente aplicable en el caso que se transmita más de una marca de agua. En este caso se provee una posición de impacto de sincronización para cada marca de agua transmitida y el bloque periódico promediado luego es dividido en diferentes segmentos en base a la posición de impacto de sincronización y en base a la longitud de la marca de agua. Los siguientes conceptos son aplicados luego para cada firma de marca de agua y su segmento de bloque periódico promediado.

4.7.1 Descartado de muestras en posiciones de sobremuestreo incorrectas:

En un primer paso, el estimador de dirección realiza integración de muestra en base a los bloques periódicos promediados. Sólo se considera la posición de sobremuestreo para cada bit que está indicado como la posición de sobremuestreo correcta por la posición de impacto de sincronización. Todas las otras posiciones de sobremuestreo son descartadas y no consideradas en la integración de muestra.

4.7.2 Inversión bits -1:

Luego, la firma de marca de agua es considerada bit por bit empezando con el primer bit de la primera firma de marca de agua. La posición del correspondiente valor de muestra en el bloque periódico promediado es indicada por la posición de impacto de sincronización. Se consideran todos los bits de la firma de marca de agua y sus correspondientes valores de muestra. En una realización que usa BPSK como esquema de modulación, mediante un valor de bit de 1 se indica iguales valores de fase de la muestra actual y la previa, mientras que mediante un valor de bit de -1 se indica una diferencia de fase de 180° de la muestra actual y la previa. Esto se ilustra en la Figura 15c. La secuencia de valores de muestra de coeficientes complejos está representada en 1510. La secuencia de bits 1515 indica los bits codificados suponiendo que se emplea codificación diferencial. El módulo de sincronización ahora genera una secuencia procesada de coeficientes complejos promediados 1520, tal que todos los coeficientes complejos aproximadamente tienen el mismo valor de fase ya sea dejando sin cambiar el coeficiente complejo o cambiando su valor de fase por 180° . La decisión, si un valor de fase de un coeficiente complejo es cambiado por 180° , se determina en base a evaluar la secuencia bits codificados de la marca de agua.

Por ejemplo, en base a la firma de marca de agua, un aparato puede calcular si un valor de bit ha sido codificado usando un coeficiente complejo con un deseado primer valor de fase o si un valor de bit ha sido codificado usando un coeficiente complejo con un segundo valor de fase no deseado. Si se ha determinado que el correspondiente valor de muestra tiene un segundo valor de fase no deseado, el valor de fase de este coeficiente complejo es cambiado por 180° .

Por este medio, mientras antes del procesamiento, los coeficientes complejos de una banda de frecuencia tenían o bien un primer valor absoluto de fase o tenían un segundo valor absoluto de fase que difería del primer valor de fase por alrededor de 180° , los valores absolutos de fase de los coeficientes complejos para una cierta banda de frecuencia de la secuencia de coeficientes complejos procesada ahora son casi el mismo valor para todos los coeficientes complejos con respecto a una cierta señal del dominio de la frecuencia de un cierto micrófono.

4.7.3 Integración de muestra

Después de ajustar la fase de las muestras, todos los coeficientes complejos tienen valores de fase con aproximadamente el mismo valor. El módulo de sincronización 750 ahora conduce la integración sobre muestra por muestra. Todos los valores de muestra restantes (los valores de muestreo con la correcta posición de sobremuestreo) que se relacionan con una firma de marca de agua considerada, son integrados, esto es, promediados, por ejemplo, sumados y divididos por el número de muestras sumadas para obtener un coeficiente complejo promediado 1530.

Se obtienen dos efectos: Como primer efecto, se obtiene un valor absoluto de fase más estable. Las diferentes muestras pueden ser expuestas a fluctuaciones dentro del canal tal que los diferentes valores de fase tienen aproximadamente los mismos valores absolutos de fase pero no exactamente los mismo valores de fase. Como segundo efecto, el efecto de la señal de audio adentro de la muestra promediada es minimizado mientras se enfatiza el efecto de la firma de marca de agua.

Así, se puede determinar un valor (absoluto) de fase para cada uno de los coeficientes complejos promediados para los diferentes micrófonos 611, 612, ..., 61N y para cada banda de frecuencia i . El valor de fase es un valor de fase promediado. Como los coeficientes complejos son números complejos, el valor absoluto de fase de un coeficiente complejo puede ser determinado determinando el valor de fase (el argumento) de un número complejo.

Sin embargo, se debe notar que los coeficientes complejos de diferentes bandas de frecuencias puede tener diferentes valores de fase como se ilustra en la Figura 15d.

4.7.4 Determinación de dirección:

En base a la información de fase determinada, se pueden determinar diversos tipos de información de dirección en base a conceptos de determinación alternativos. Usando los métodos que dependen de la configuración de la disposición ordenada de micrófonos y el uso de diferencias de fase entre las señales entrantes, por ejemplo, codificación de audio direccional (DirAC) o Estimación de Parámetros de Señal vía Técnicas de Invariancia Rotacional (ESPRIT), se obtiene una estimación de la dirección de llegada.

En DirAC, se analiza un campo de sonido usando señales de micrófono de formato B, los cuales comprende una señal omnidireccional $w(t)$ y tres señales de dipolo $x(t)$, $y(t)$, $z(t)$, las cuales corresponden a los ejes x , y , y z de un sistema de coordenadas cartesiano.

La Figura 16 ilustra una disposición ordenada de micrófonos de seis micrófonos tal que forman tres pares de micrófonos. Se supone que el primer par de micrófonos $mx1$, $mx2$ está ubicado sobre el eje x de un sistema de coordenadas cartesiano. El segundo par de micrófonos $my1$, $my2$ está acomodado tal que puede suponerse que está ubicado sobre el eje y del sistema de coordenadas cartesiano. Además, el tercer par de micrófonos $mz1$, $mz2$ está acomodado tal que puede suponerse que está ubicado sobre el eje z del sistema de coordenadas cartesiano.

Considere una cierta banda de frecuencia i . El valor absoluto de fase determinado con relación a una firma de marca de agua considerada para una banda de frecuencia considerada i , es llamado $vx1$ para el micrófono $mx1$, $vx2$ para el micrófono $mx2$, $vy1$ para el micrófono $my1$, $vy2$ para el micrófono $my2$, $vz1$ para el micrófono $mz1$ y $vz2$ para el micrófono $mz2$. Entonces, las diferencias de fases Δvx , Δvy y Δvz pueden ser calculadas como sigue:

$$\begin{aligned}\Delta vx &= vx1 - vx2; \\ \Delta vy &= vy1 - vy2; \\ \Delta vz &= vz1 - vz2;\end{aligned}$$

Determinación de un ángulo azimutal:

En una realización, se puede calcular un ángulo azimutal en base a los valores de fase calculados. Considere por ejemplo la ubicación de los micrófonos $mx1$ y $mx2$ en el plano xy como se ilustra en la Figura 17. Se indica la distancia entre los dos micrófonos. Hay que especificar un ángulo ϕ que especifica la dirección de una fuente de sonido que entrega la firma de marca de agua, con respecto al plano xy . $\phi=0$ corresponde a una onda que impacta los micrófonos al mismo tiempo. Esto es, se supone que la fuente de sonido está igualmente espaciada de $mx1$ que de $mx2$. Se emplea la siguiente fórmula:

$$\Delta v_x = (-2 \cdot \pi / \lambda) \cdot d \cdot \sin(\varphi)$$

En esta ecuación, λ es la longitud de onda y está definida como $\lambda = c/f_i$ en donde c es la velocidad del sonido y en donde f_i es la frecuencia central de la subbanda considerada. f_i puede corresponder a f_i en la función formadora de bit para la subbanda considerada como se describió arriba. d es la distancia entre los dos micrófonos $mx1$ y $mx2$ y φ es el ángulo que indica la dirección de llegada con respecto al plano xy a ser determinado.

Además, $\sin(\varphi)$ es:

$$\sin(\varphi) = -\Delta v_x \cdot \lambda / (2 \cdot \pi \cdot d) \quad (I)$$

Así se puede determinar el valor $\sin(\varphi)$.

Suponemos que φ está en el rango $]-\pi, \pi]$. Si $\sin(\varphi)$ es -1 o 1 , φ puede ser determinado sin ambigüedad. Sin embargo, si $\sin(\varphi)$ no es -1 o 1 , φ puede tomar uno de dos valores diferentes. Por ejemplo, si $\sin(\varphi)=0$, entonces φ es 0 o π .

Para resolver esta ambigüedad, se puede emplear información proveniente de otros micrófonos. Considere la ubicación de los micrófonos $my1$ y $my2$ en el plano xy como se ilustra en la Figura 17. Nuevamente se indica la distancia entre los dos micrófonos. La distancia puede ser idéntica a la distancia entre los micrófonos $mx1$ y $mx2$ o puede ser diferente. Hay que especificar un ángulo β que especifica la dirección de una fuente de sonido que entrega la firma de marca de agua, con respecto al plano xy . $\beta=0$ corresponde a una onda que impacta los micrófonos al mismo tiempo. Esto es, se supone que la fuente de sonido está igualmente espaciada de $mx1$ que de $mx2$. Se emplea la siguiente ecuación:

$$\Delta v_y = (-2 \cdot \pi / \lambda) \cdot d \cdot \sin(\beta)$$

Esto da por resultado la fórmula:

$$\sin(\beta) = -\Delta v_y \cdot \lambda / (2 \cdot \pi \cdot d)$$

En la realización ilustrada en la Figura 17, $\beta = \varphi - \pi/2$. (en otras realizaciones, β puede ser $\beta = \varphi + \pi/2$). Por lo tanto,

$\sin(\beta) = \sin(\varphi - \pi/2) = -\cos(\varphi)$ y por ende:

$$\cos(\varphi) = \Delta v_y \cdot \lambda / (2 \cdot \pi \cdot d) \quad (II)$$

De nuevo, si $\cos(\varphi)$ no es -1 o 1 , φ puede tomar uno de dos valores.

Calculando $\sin(\varphi)$ de acuerdo con la fórmula (I) y calculando $\cos(\varphi)$ de acuerdo con (II) y determinando φ en base al $\sin(\varphi)$ y al $\cos(\varphi)$, se puede determinar el valor de φ sin ambigüedad.

Por este medio se puede determinar este ángulo φ que especifica una dirección de llegada de una onda de sonido que comprende una firma de marca de agua con respecto al plano xy y una cierta banda de frecuencia i .

La determinación del ángulo φ podría repetirse para cada una de una pluralidad de diferentes bandas de frecuencia i . Para cada banda de frecuencia se obtiene un ángulo φ que debe ser casi igual para cada banda de frecuencia. Sin embargo, perturbaciones con respecto a ciertas bandas de frecuencia pueden influir la corrección del valor determinado del ángulo φ el cual es determinado en base a estas bandas de frecuencia. En una realización, se calcula un valor promedio de los valores determinados para el ángulo φ . Por, los valores de ángulo calculados pueden ser sumados y divididos por el número de sumandos para obtener el valor del ángulo φ como un valor promedio.

Determinación del vector dirección de llegada a partir del ángulo azimutal φ : Se puede computar el vector dirección de llegada según

$$a = (\cos(\varphi), \sin(\varphi)).$$

Determinación alternativa de un vector dirección de llegada:

En una realización, se determina un vector dirección de llegada en base a los valores de fase calculados como sigue.

Volviendo a la Figura 17, como se explicó arriba, las diferencias de fase Δv_x , Δv_y y Δv_z para los valores absolutos de fase pueden ser calculadas aplicando $\Delta v_x = v_{x1} - v_{x2}$; $\Delta v_y = v_{y1} - v_{y2}$; y $\Delta v_z = v_{z1} - v_{z2}$.

En una realización, se supone que la distancia d_x entre los micrófonos $mx1$ y $mx2$, la distancia d_y entre los micrófonos $my1$ y $my2$, y la distancia d_z entre los micrófonos $mz1$ y $mz2$, son iguales.

Entonces Δv_x , Δv_y y Δv_z forman los componentes de la dirección de un vector llegada a , cuando Δv_x , Δv_y y Δv_z se expresan en el rango $]-\pi; \pi]$:

$$a = (\Delta v_x, \Delta v_y, \Delta v_z)$$

el cual apunta a la dirección de la fuente de sonido.

En otra realización, la distancia d_x entre los micrófonos $mx1$ y $mx2$, la distancia d_y entre los micrófonos $my1$ y $my2$, y la distancia d_z entre los micrófonos $mz1$ y $mz2$, no son iguales. En una realización así, las diferencias de fases modificadas tienen que ser ajustadas para tomar en cuenta las diferentes distancias, por ejemplo aplicando un factor de corrección. Por ejemplo, los componentes a_x , a_y y a_z pueden ser determinados de acuerdo con:

$$a_x = (1/d_x) \cdot \Delta v_x; \quad a_y = (1/d_y) \cdot \Delta v_y; \quad a_z = (1/d_z) \cdot \Delta v_z$$

tal que el vector dirección de llegada resulta ser $a = (a_x, a_y, a_z)$.

En realizaciones adicionales, el vector puede ser normalizado, por ejemplo, multiplicando sus componentes por un factor k , por ejemplo,

$$k = \sqrt{\Delta v_x^2 + \Delta v_y^2 + \Delta v_z^2}$$

El vector dirección de llegada se explica con referencia a las Figuras 18a - 18d. Para mejor ilustración, la situación se explica en un plano bidimensional, el plano xy . Se supone que la distancia d_x entre los micrófonos $mx1$ y $mx2$ y la distancia d_y entre los micrófonos $my1$ y $my2$, son iguales. Sin embargo, también se aplican las explicaciones dadas al caso tridimensional y a situaciones donde las diferencias entre los micrófonos no son iguales, pero donde las diferencias de fase Δv_x y Δv_y han sido ajustadas, por ejemplo, aplicando un factor de distancia.

Considere un vector dirección de llegada $a = (\Delta v_x, \Delta v_y, \Delta v_z)$. Sus componentes Δv_x , Δv_y , Δv_z son las diferencias de fase obtenidas según las realizaciones de arriba. Por razones de claridad de ilustración, sólo consideramos las componentes x e y de un vector $a' = (\Delta v_x, \Delta v_y)$.

En la Figura 18a, consideramos una dirección de dos componentes del vector llegada $a' = (0, 1)$.

La componente x del vector Δv_x es 0, esto es no hay diferencia de fase con respecto a los valores absolutos de fase de las señales recibidas por ambos micrófonos $mx1$ y $mx2$. Así, la correspondiente onda de audio está impactando en ambos micrófonos $mx1$ y $mx2$ al mismo tiempo, y se supone que la fuente de sonido está igualmente espaciada desde ambos micrófonos $mx1$ y $mx2$. Como resultado, el vector dirección de llegada a' no es ni positiva ni negativa.

La componente y del vector Δv_y es 1. Como $\Delta v_y = v_{y1} - v_{y2}$ esto indica que el valor absoluto de fase v_{y1} es mayor que el valor absoluto de fase v_{y2} , y por ende la respectiva onda de audio arribó antes al micrófono $my1$ que al micrófono $my2$. Así, vector dirección de llegada a' apunta hacia la dirección positiva del eje y .

En la Figura 18b, se considera una dirección de dos componentes del vector llegada $a' = (-1, 0)$.

La componente y del vector Δv_y es 0, esto es no hay diferencia de fase con respecto a los valores absolutos de fase de las señales recibidas por ambos micrófonos $my1$ y $my2$. Así, la correspondiente onda de audio está impactando en ambos micrófonos $my1$ y $my2$ al mismo tiempo, y se supone que la fuente de sonido está igualmente espaciada desde ambos micrófonos $my1$ y $my2$. Como resultado, el vector dirección de llegada a' no es ni positiva ni negativa.

La componente x del vector Δv_x es -1. Como $\Delta v_x = v_{x1} - v_{x2}$ esto indica que el valor absoluto de fase v_{x2} es mayor que el valor absoluto de fase v_{x1} , y por ende la respectiva onda de audio arribó antes al micrófono $mx2$ que al micrófono $mx1$. Así, el vector dirección de llegada a' apunta hacia la dirección negativa del eje x .

En la Figura 18c, consideramos una dirección de dos componentes del vector llegada $a' = (1, 1)$.

La componente x del vector Δv_x es 1. Como $\Delta v_x = v_{x1} - v_{x2}$ esto indica que el valor absoluto de fase v_{x1} es mayor que el valor absoluto de fase v_{x2} , y por ende la respectiva onda de audio arribó antes al micrófono $mx1$ que al micrófono $mx2$. Así, el vector dirección de llegada a' apunta hacia la dirección positiva del eje x .

La componente y del vector Δv_y es 1. Como $\Delta v_y = v_{y1} - v_{y2}$ esto indica que el valor absoluto de fase v_{y1} es mayor que el valor absoluto de fase v_{y2} , y por ende la respectiva onda de audio arribó antes al micrófono my_1 que al micrófono my_2 . Así, el vector dirección de llegada a' apunta hacia la dirección positiva del eje y.

5 Como las diferencias de fase Δv_x y Δv_y son iguales, y como se puede suponer que una correspondiente onda de audio se propaga igualmente rápido en la dirección x y en la dirección y, la componente x y la componente y del vector dirección de llegada tienen que tener valores iguales.

10 En la Figura 18d, se considera una dirección de dos componentes del vector llegada $a' = (2, 1)$.

La componente x del vector Δv_x es 2. Como $\Delta v_x = v_{x1} - v_{x2}$ esto indica que el valor absoluto de fase v_{x1} es mayor que el valor absoluto de fase v_{x2} , y por ende la respectiva onda de audio arribó antes al micrófono mx_1 que al micrófono mx_2 . Así, el vector dirección de llegada a' apunta hacia la dirección positiva del eje x.

15 La componente y del vector Δv_y es 1. Como $\Delta v_y = v_{y1} - v_{y2}$ esto indica que el valor absoluto de fase v_{y1} es mayor que el valor absoluto de fase v_{y2} , y por ende la respectiva onda de audio arribó antes al micrófono my_1 que al micrófono my_2 . Así, el vector dirección de llegada a' apunta hacia la dirección positiva del eje y.

20 Como las diferencias de fase Δv_x y Δv_y son iguales y como podemos suponer que una correspondiente onda de audio se propaga más rápido en una dirección x que en una dirección y, y que su velocidad en la dirección x es el doble de la velocidad en la dirección y, la componente x es el doble de tamaño del tamaño de la componente y del vector dirección de llegada.

25 Determinación del vector dirección de llegada en base a múltiples bandas de frecuencia:

Como se explicó arriba, el vector dirección de llegada se determina en base a valores de fase de cierta banda de frecuencia. En una realización, el vector dirección de llegada se calcula para una pluralidad de bandas de frecuencia para obtener una pluralidad de vectores dirección de llegada. La pluralidad de vectores de banda de frecuencia luego pueden ser normalizados, por ejemplo, multiplicando sus componentes por factores de normalización específicos del vector para obtener vectores dirección de llegada normalizados. Por ejemplo, para cada banda de frecuencia n se puede determinar un factor de normalización específico del vector, por ejemplo, de acuerdo con la fórmula:

35
$$n = \frac{1}{f_i}, \text{ donde } f_i \text{ es la frecuencia central de la } i\text{-ésima subbanda.}$$

Luego se puede determinar un vector dirección de llegada promedio. El promediado se puede lograr, por ejemplo, sumando la pluralidad de vectores dirección de llegada y dividiendo las componentes del vector obtenido por el número de vectores dirección de llegada normalizados que se promediaron.

40 En una realización se pueden calcular ángulos de azimut y de elevación en base al vector dirección de llegada determinado usando los métodos empleados en el contexto de codificación de audio direccional (DirAC).

Estimación alternativa del vector dirección de llegada usando información compleja específica de receptor:

45 En una realización adicional, se pueden usar valores complejos relacionados con una firma de marca de agua considerada para una banda de frecuencia i considerada para estimar la dirección de llegada usando la técnica de codificación de audio direccional (DirAC).

50 Sea Δp_x , que denota la diferencia entre los valores complejos obtenidos de los micrófonos mx_1 y mx_2 de la Figura 16. Análogamente, se obtienen Δp_y y Δp_z para los ejes y y z. Sea además p_0 el promedio de los valores complejos obtenidos en todos los micrófonos.

El vector dirección de llegada a se calcula a partir del vector intensidad activa $la = (lax, lay, laz)$ definido según

55
$$\begin{aligned} lax &= .5 \text{ Real} (p_0 \cdot \text{conj}(\Delta p_x)) \\ lay &= .5 \text{ Real} (p_0 \cdot \text{conj}(\Delta p_y)) \\ laz &= .5 \text{ Real} (p_0 \cdot \text{conj}(\Delta p_z)), \end{aligned}$$

donde $\text{Real}()$ extrae la parte real, y conj denota conjugación compleja, y donde .5 designa 0,5.

60 El vector dirección de llegada es

$$a = -la / \text{norm}(la),$$

donde $\text{norm}()$ computa la norma L2 del argumento. Los ángulos de azimut y de elevación se calculan fácilmente con

$$\begin{aligned}\text{azimut} &= \text{atan}(\text{lay}/\text{lax}) \\ \text{elevación} &= \text{atan}(\text{norm}(\text{la}), \sqrt{\text{lax}^2 + \text{lay}^2}),\end{aligned}$$

donde $\sqrt{}$ es la raíz cuadrada, donde lax^2 designa lax^2 , y donde lay^2 designa lay^2 .

Estimación de Posición de la Unidad de Registro

La Figura 19 ilustra la estimación de posición en un escenario bidimensional. Allí se supone que la posición de los altavoces es conocida. Además suponemos que los altavoces y la disposición ordenada de micrófonos de registro están ubicados en el mismo plano.

Se determina un vector dirección de llegada para cada uno de los altavoces de acuerdo con una de las realizaciones descritas arriba. El vector dirección de llegada del primer altavoz apunta al primer altavoz. El vector dirección de llegada del segundo altavoz apunta al segundo altavoz.

Básicamente, un vector dirección de llegada apunta a la dirección de llegada de una onda de sonido proveniente de un altavoz, en donde las componentes del vector pueden ser expresadas con referencia a un sistema de coordenadas cartesiano. Sin embargo, en algunas realizaciones, los ejes del sistema de coordenadas se definen mediante la ubicación de los micrófonos: La ubicación de mx1 y mx2 define el eje x , y la ubicación de my1 y my2 define el eje y del sistema de coordenadas al que se refiere el vector dirección de llegada.

Si la orientación de la disposición ordenada de micrófonos es conocida, y si las posiciones (absolutas) de los dos altavoces son conocidas, dos vectores dirección de llegada son suficientes en el caso bidimensional para definir la posición de la disposición ordenada de micrófonos. Esto se ilustra en la Figura 19.

En una realización así, un aparato para estimación de posición está adaptado para calcular un vector dirección de llegada para cada altavoz que entrega ondas de sonido que comprenden firmas de marca de agua. Se calcula una primera línea en el plano bidimensional que es paralela al vector dirección de llegada del primer altavoz y que interseca la posición del primer altavoz. Además se calcula una segunda línea en el plano bidimensional que es paralela al vector dirección de llegada del segundo altavoz y que interseca la posición del primer altavoz. Luego se calcula el punto de intersección de la primera y la segunda línea en el plano bidimensional, en donde el punto de intersección determina la posición de la disposición ordenada de micrófonos.

En otras realizaciones, el concepto descrito se aplica en ambientes en donde la disposición ordenada de micrófonos y los altavoces no está ubicados en un plano bidimensional, aplicando vectores dirección de llegada de tridimensionales (vectores dirección de llegada con tres componentes). Allí se supone que la ubicación de los altavoces es conocida. Un aparato calcula las líneas paralelas a los vectores dirección de llegada determinados, en donde las líneas calculadas intersecan la posición de altavoz del altavoz de su vector dirección de llegada. El punto en donde todas las líneas calculadas se intersecan entre sí, es la posición determinada de la disposición ordenada de micrófonos.

Con propósitos ilustrativos, volvamos al caso bidimensional. Si la orientación de la disposición ordenada de micrófonos no es conocida, entonces dos vectores dirección de llegada de dos altavoces no son suficientes para determinar la posición de la disposición ordenada de micrófonos. En un escenario así, se podría calcular un ángulo α entre el primer y el segundo vector dirección de llegada, pero como la orientación de la disposición ordenada de micrófonos y por ende la orientación del sistema de coordenadas definido por la disposición ordenada de micrófonos no sería conocida, la posición de la disposición ordenada de micrófonos no podría ser definida sin ambigüedad, incluso si se conociera la posición del altavoz (por ejemplo, expresada con referencia a un segundo sistema de coordenadas). Esto se ilustra en la Figura 20a.

Sin embargo, empleando un tercer altavoz que entrega una tercera onda de audio que comprende una firma de marca de agua, se podría resolver la falta de ambigüedad. Esto se explica con referencia a la Figura 20b. Si se estiman las direcciones de llegada de por lo menos tres altavoces ($M \geq 3$), también se puede determinar la posición espacial de la disposición ordenada de micrófonos. La Figura 21 ilustra un correspondiente estimador de posición. Cuando se usan tres altavoces, se supone que los altavoces y la unidad de registro están en un plano común. Sin embargo, las posiciones espaciales de los altavoces tienen que ser conocidas en el receptor. Apoyándose en esta información, es posible obtener la posición de registro resolviendo un sistema no lineal de ecuaciones en el estimador de posición 800.

En la realización de la Figura 20b, un aparato para estimación de posición calcula el primer vector dirección de llegada del primer altavoz, el segundo vector dirección de llegada del segundo altavoz y el tercer vector dirección de llegada del tercer altavoz. Luego se puede calcular un primer ángulo α entre el primer y el segundo vector dirección

de llegada y un segundo ángulo δ entre el segundo y el tercer vector dirección de llegada. El aparato para estimación de dirección luego determina el punto en el plano bidimensional tal que tres líneas cada una de las cuales interseca una posición de altavoz y el punto determinado tiene las relaciones angulares determinadas entre sí. Eso es, la primera línea que interseca la posición del primer altavoz y la segunda línea que interseca la posición del segundo altavoz se intersecan en el punto determinado tal que su ángulo de intersección es igual al primer ángulo. La segunda línea que interseca la posición del segundo altavoz y la tercera línea que interseca la posición del tercer altavoz se intersecan en el punto determinado tal que su ángulo de intersección es igual al segundo ángulo. El punto determinado es la posición de la disposición ordenada de micrófonos.

En realizaciones adicionales, un estimador de posición determina la posición de una disposición de micrófonos cuya orientación no es conocida para el caso tridimensional, esto es para una configuración donde los altavoces y la disposición ordenada de micrófonos no están todos ubicados en un plano bidimensional. La determinación se logra en base al concepto descrito arriba empleando vectores dirección de llegada tridimensionales, esto es, vectores dirección de llegada que tienen tres componentes.

La Figura 22 ilustra un aparato para estimación de posición espacial de acuerdo con una realización. El aparato para estimación de posición espacial comprende un aparato 100 para proveer información de dirección de acuerdo con una de las realizaciones descritas arriba, que procesa señales de audio con marca de agua recibidas rx1, rx2. Asimismo, el aparato para estimación de posición espacial comprende un estimador de posición 300 para estimar una posición del aparato para estimación de posición espacial. El estimador de posición está adaptado para estimar la posición del aparato para estimar posición espacial en base a la información de dirección provista por el aparato para proveer información de dirección.

A pesar de que se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a un paso de método o a un rasgo de un paso de método. Análogamente, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un correspondiente bloque o componente o rasgo de un correspondiente aparato.

La señal descompuesta inventiva puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida a través de un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión vía cable, tal como Internet.

Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo un diskette, un DVD, un CD, una memoria ROM, una EPROM, una EEPROM o una memoria FLASH, los cuales tienen unas señales de control electrónicamente legibles guardadas en ellos, las cuales cooperan (o son capaces de cooperar) con un sistema de computación programable de modo que se ejecuta el respectivo método.

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos no transitorio que tiene señales de control legibles electrónicamente, las cuales son capaces de cooperar con un sistema de ordenador programable, tal que uno de los métodos descrito en la presente sea ejecutado.

Generalmente, realizaciones de la presente invención pueden ser implementadas como un programa de ordenador con un código de programa, siendo código de programa operativo para ejecutar uno de los métodos cuando el producto de programa de ordenador corre en un ordenador. El código de programa puede ser almacenado, por ejemplo, sobre un portador legible por una máquina.

Otras realizaciones comprenden el programa de ordenador para ejecutar uno de los métodos descritos en la presente, almacenado en un portador legible por una máquina.

En otras palabras, una realización del método inventivo es, por lo tanto, un programa de ordenador que un código de programa para ejecutar uno de los métodos descritos en la presente, cuando el programa de ordenador corre en un ordenador.

Una realización adicional de los métodos inventivos es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa de ordenador para ejecutar uno de los métodos descritos en la presente.

Una realización adicional del método inventivo es, por lo tanto, una secuencia de datos o una secuencia de señales que representan el programa de ordenador para ejecutar uno de los métodos descritos en la presente. La secuencia de datos o la secuencia de señales pueden ser configuradas, por ejemplo, para ser transferidos vía una conexión de comunicación de datos, por ejemplo, vía Internet.

Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado para o adaptado para ejecutar uno de los métodos descritos en la presente.

5 Una realización adicional comprende un ordenador que tiene instalado en ella el programa de ordenador para ejecutar uno de los métodos descritos aquí.

10 En algunas realizaciones se puede usar un dispositivo de lógica programable (por ejemplo una disposición de puertas programables en campo) para realizar algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, la disposición de puertas programables en campo puede cooperar con un microprocesador para realizar uno de los métodos descritos en la presente. Generalmente, los métodos preferiblemente son realizados mediante algún aparato de hardware.

15 Las realizaciones que se describieron más arriba son puramente ilustrativas para los principios de la presente invención. Se entiende que las modificaciones y variaciones posibles de las disposiciones y de los detalles descritos en la presente serán evidentes para los expertos en la materia. Por lo tanto, es la intención que la invención esté limitada sólo por el alcance de las siguientes reivindicaciones de patente y no por los detalles específicos presentados por la descripción y la explicación de las realizaciones en la presente.

REIVINDICACIONES

1. Un aparato (100) para proveer información de dirección en base a una señal de audio (awm) reproducida con una marca de agua incrustada, que comprende:

un procesador de señal (110) que está adaptado para procesar por lo menos dos señales de audio con marca de agua (rx1, rx2) registradas por al menos dos receptores de audio (221, 222) en diferentes posiciones espaciales, en donde los receptores de audio (221, 222) están adaptados para registrar ondas de sonido de la señal de audio reproducida (awm) para obtener las señales de audio con marca de agua recibidas (rx1, rx2), donde cada señal de audio con marca de agua registrada (rx1, rx2) comprende la marca de agua incrustada, en donde el procesador de señal (110) está adaptado para procesar las señales de audio con marca de agua registradas (rx1, rx2) para obtener información de fase como información específica de receptor para cada señal de audio con marca de agua registrada (rx1, rx2), donde la información específica de receptor depende de las marcas de agua incrustadas en las señales de audio con marca de agua registradas (rx1, rx2), y

un proveedor de información de dirección (120) para proveer información de dirección en base a la información específica de receptor para cada señal de audio con marca de agua registrada (rx1, rx2).

2. Un aparato (100) de acuerdo con la reivindicación 1, en donde el procesador de señal (110) comprende un módulo de análisis (711, 712, 71N) para transformar las señales de audio con marca de agua registradas (rx1, rx2) desde un dominio del tiempo a un dominio de tiempo-frecuencia para obtener una señal de transmisión del dominio de tiempo-frecuencia que comprende una pluralidad de señales de subbanda.

3. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones 1 o 2, en donde el procesador de señal (110) está adaptado para determinar un número complejo como la información específica de receptor.

4. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde el procesador de señal (110) está adaptado para determinar un valor de fase promediado como la información de fase en base a un coeficiente complejo promediado determinado.

5. Un aparato (100) de acuerdo con la reivindicación 4, en donde el procesador de señal (110) está adaptado para generar el coeficiente complejo promediado determinando un bloque periódico promediado que comprende coeficientes complejos procesados de una señal subbanda.

6. Un aparato (100) de acuerdo con la reivindicación 2, en donde el procesador de señal (110) está adaptado para determinar información de fase como la información específica de receptor, en donde el procesador de señal (110) está adaptado para determinar un valor de fase promediado como la información de fase en base a un coeficiente complejo promediado determinado, y en donde el procesador de señal (110) está adaptado para generar el coeficiente complejo promediado determinando un bloque periódico promediado que comprende coeficientes complejos procesados de una de la pluralidad de señales subbanda.

7. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones 5 o 6, en donde el procesador de señal (110) está adaptado para determinar el coeficiente complejo promediado combinando por lo menos dos de los coeficientes complejos procesados de las señales subbanda, en donde los coeficientes complejos procesados tienen iguales valores de fase o valores de fase que difieren en no más que 0,5 medido en radianes.

8. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones 5 a 7, en donde el procesador de señal (110) está adaptado para generar los coeficientes complejos procesados cambiando el valor de fase de por lo menos uno de los coeficientes complejos por un número predeterminado.

9. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde el procesador de señal (110) está adaptado para determinar una posición de impacto de sincronización que indica una posición de un coeficiente complejo que codifica un primer bit de una firma de marca de agua.

10. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones 5 a 8, en donde el procesador de señal (110) está adaptado para determinar una posición de impacto de sincronización que indica una posición de por lo menos uno de los coeficientes complejos procesados que codifica un primer bit de una firma de marca de agua de la marca de agua incrustada, incrustada en las señales de audio con marca de agua registradas (rx1, rx2).

11. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde el proveedor de información de dirección (120) está adaptado para emplear un número complejo para determinar y proveer información de dirección.

12. Un aparato (100) de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde el proveedor de información de dirección (120) está adaptado para proveer un vector dirección de llegada como información de dirección.
- 5 13. Un aparato para estimación de posición espacial, que comprende:
- un aparato (100) para proveer información de dirección de acuerdo con una cualquiera de las reivindicaciones 1 a 11, y
- 10 un estimador de posición (300) para estimar una posición del aparato para estimar posición espacial, en donde el estimador de posición (300) está adaptado para estimar la posición del aparato para estimación de posición espacial en base a la información de dirección provista por el aparato para proveer información de dirección.
14. Un método para proveer información de dirección en base a una señal de audio reproducida (awm) con una marca de agua incrustada, que comprende:
- 15 registrar ondas de sonido de la señal de audio reproducida (awm) mediante por lo menos dos receptores de audio (221, 222) en diferentes posiciones espaciales para obtener al menos dos señales de audio con marca de agua registradas (rx1, rx2), donde cada una de las por lo menos dos señales de audio con marca de agua registradas (rx1, rx2) comprende la marca de agua incrustada,
- 20 procesar las por lo menos dos señales de audio con marca de agua registradas (rx1, rx2) para obtener información de fase como información específica de receptor para cada señal de audio con marca de agua registrada (rx1, rx2), en donde la información específica de receptor depende de las marcas de agua incrustadas, incrustadas en las señales de audio con marca de agua registradas (rx1, rx2), y
- 25 proveer información de dirección en base a la información específica de receptor para cada una de las por lo menos dos señales de audio con marca de agua registrada (rx1, rx2).
15. Un método para estimación de posición espacial de acuerdo con la reivindicación 14,
- 30 en donde la etapa de recibir señales de audio con marca de agua se lleva a cabo empleando una disposición ordenada de por lo menos dos receptores de audio, cada señal de audio con marca de agua registrada (rx1, rx2) comprendiendo una marca de agua incrustada, y
- 35 en donde el método comprende además la etapa de estimar una posición de la disposición ordenada de por lo menos dos receptores de audio, en donde la posición es determinada en base a la información de dirección.
16. Programa de ordenador para realizar el método de acuerdo con la reivindicación 14 o la 15, cuando el programa de ordenador es ejecutado por un ordenador o un procesador.
- 40

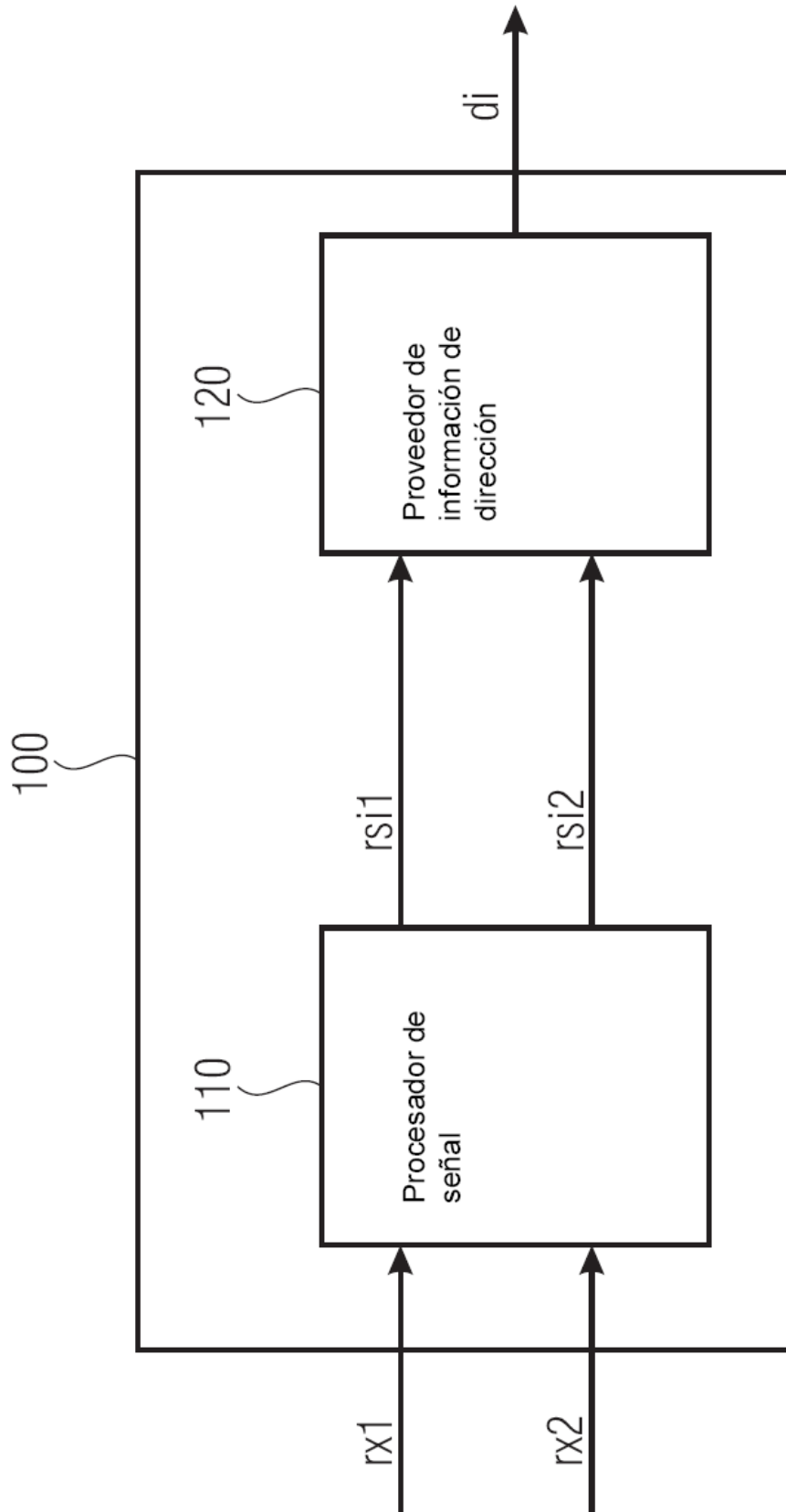


FIG 1

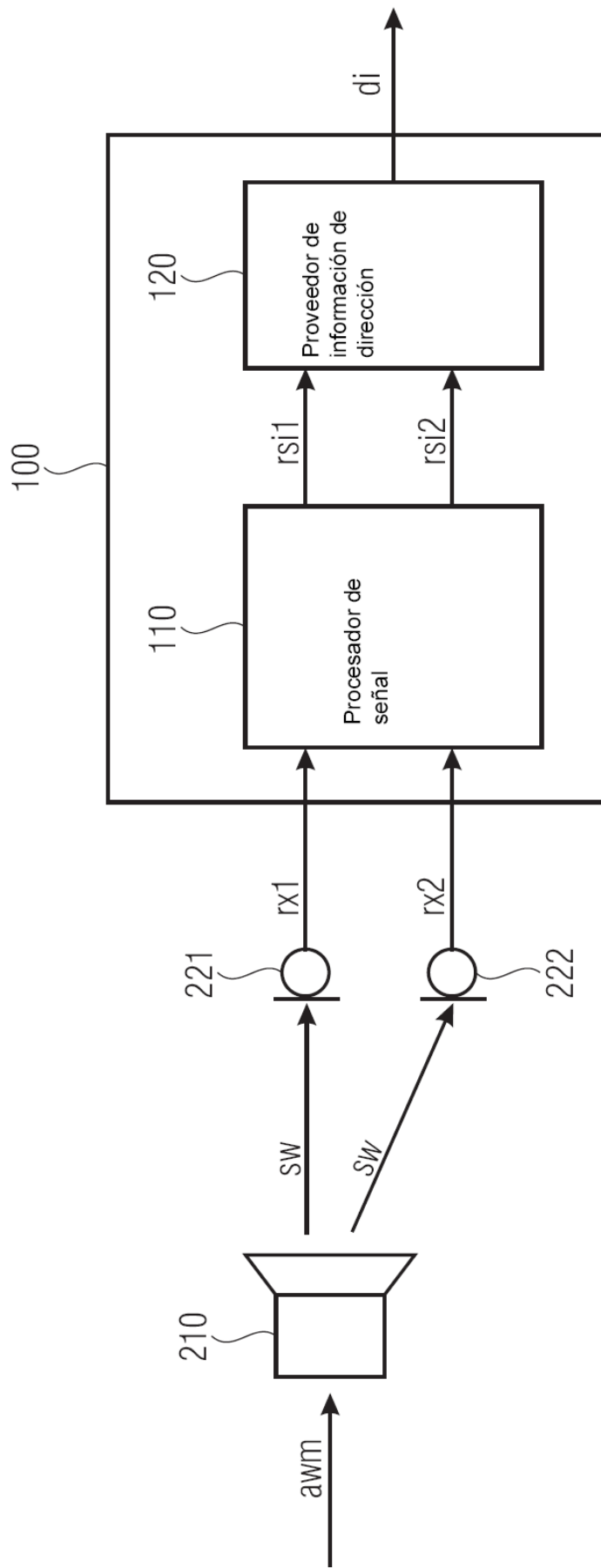


FIG 2



FIG 3A

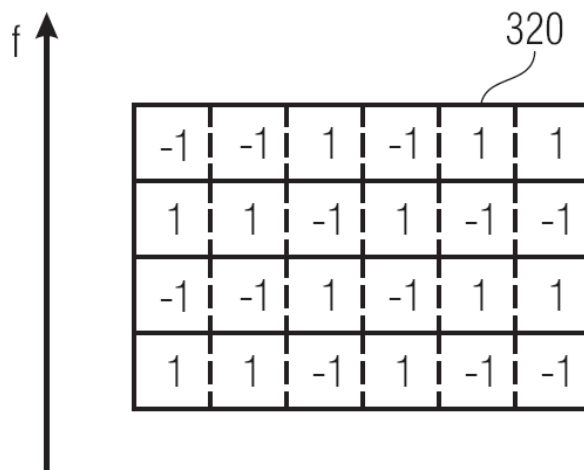


FIG 3B

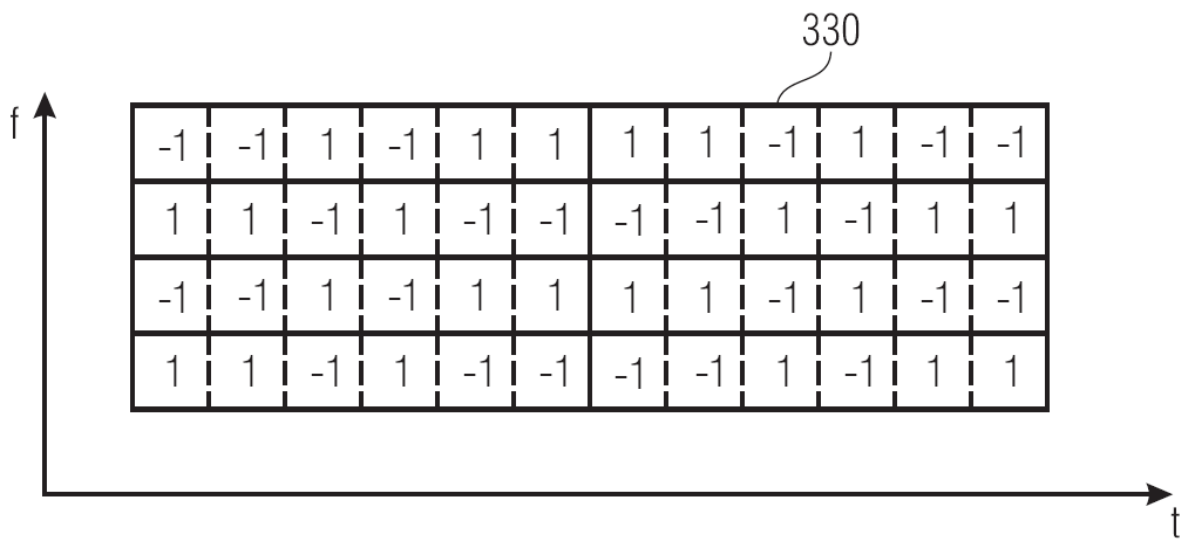


FIG 3C

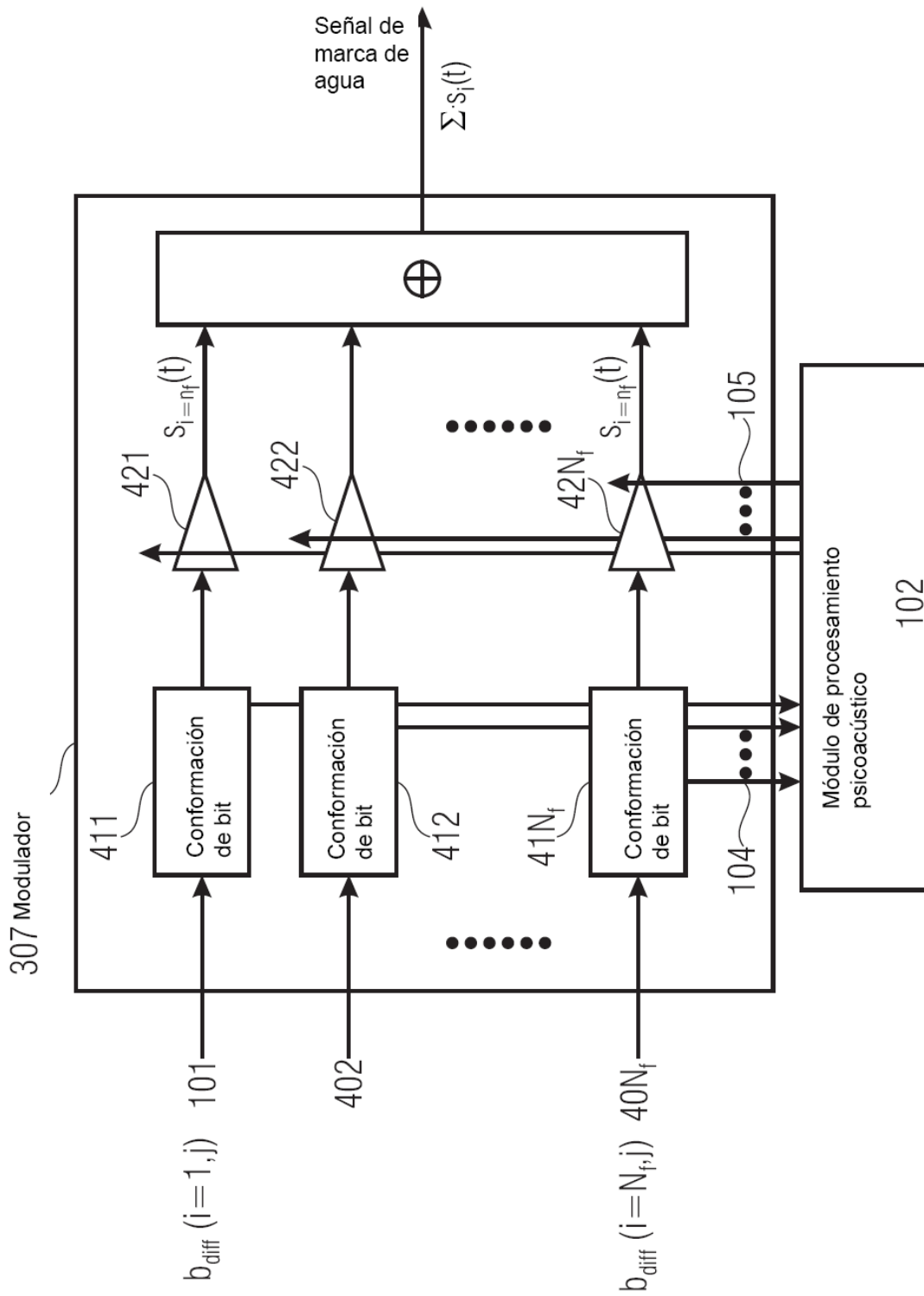


FIG 4

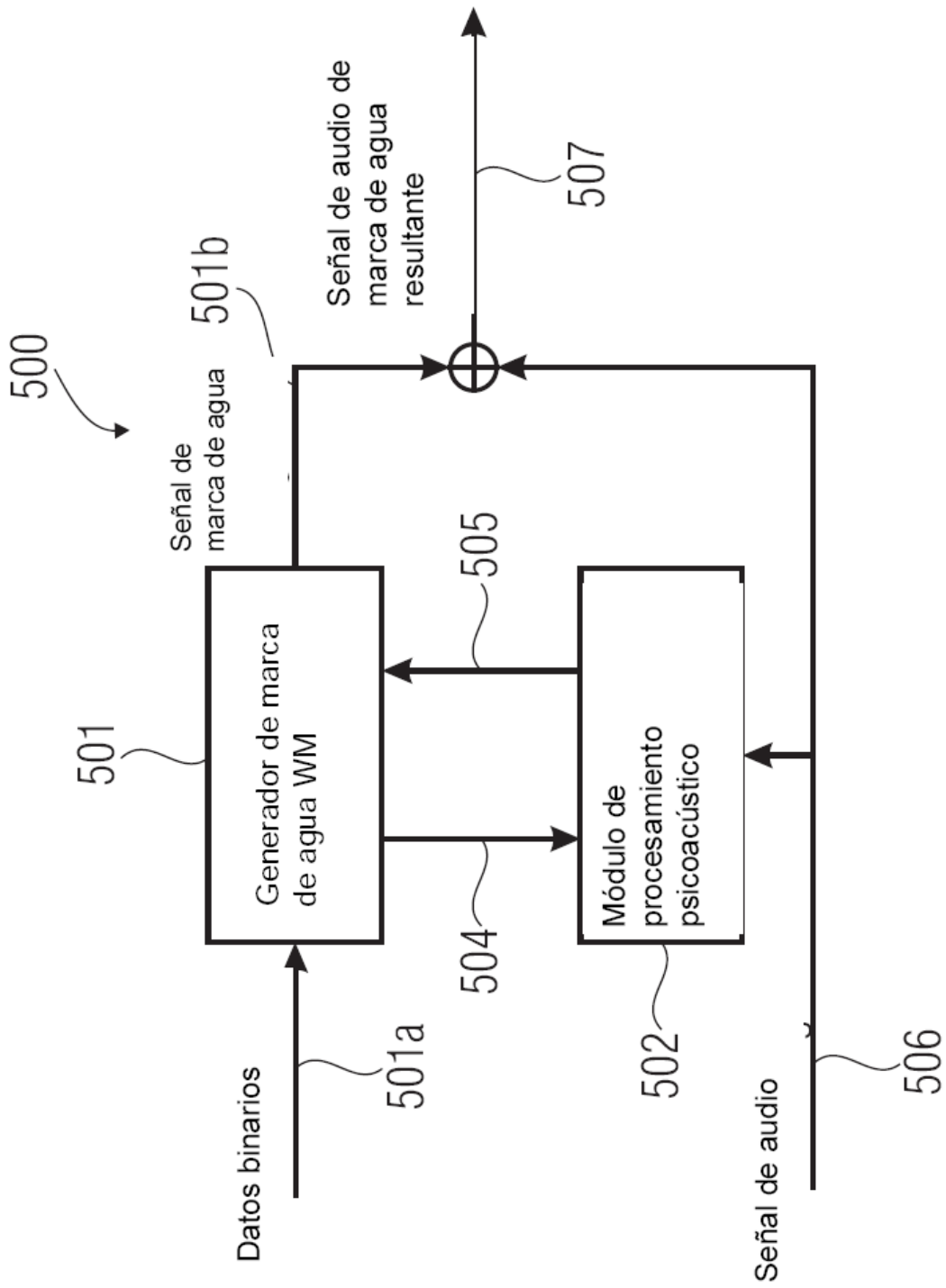


FIG 5

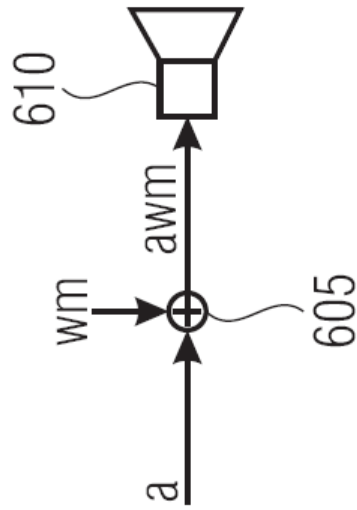
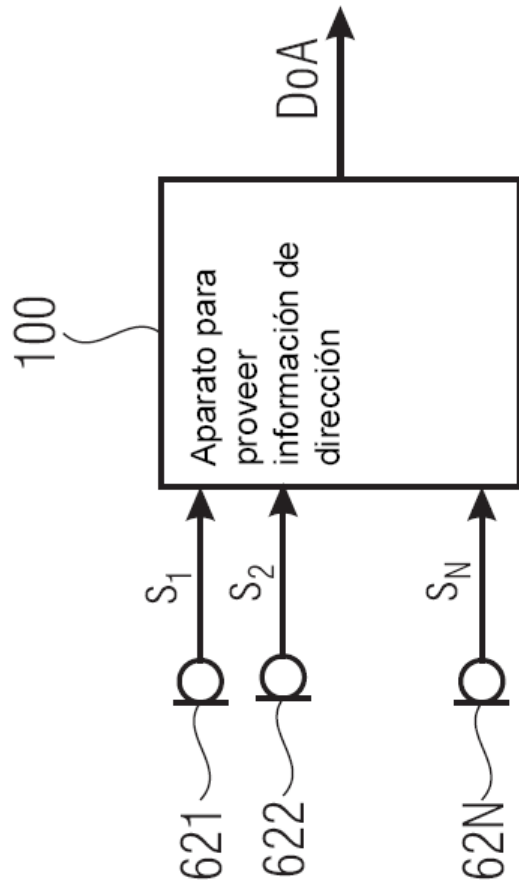


FIG 6A

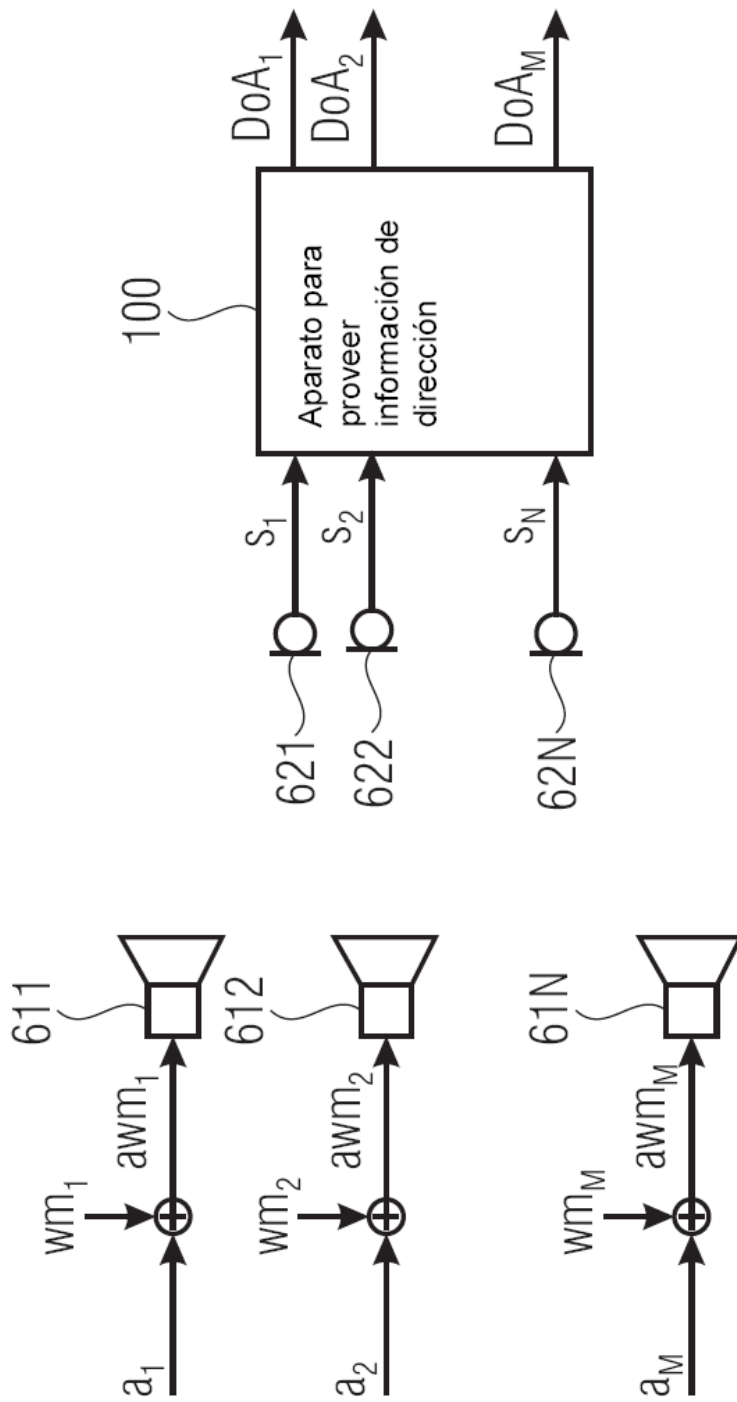


FIG 6B

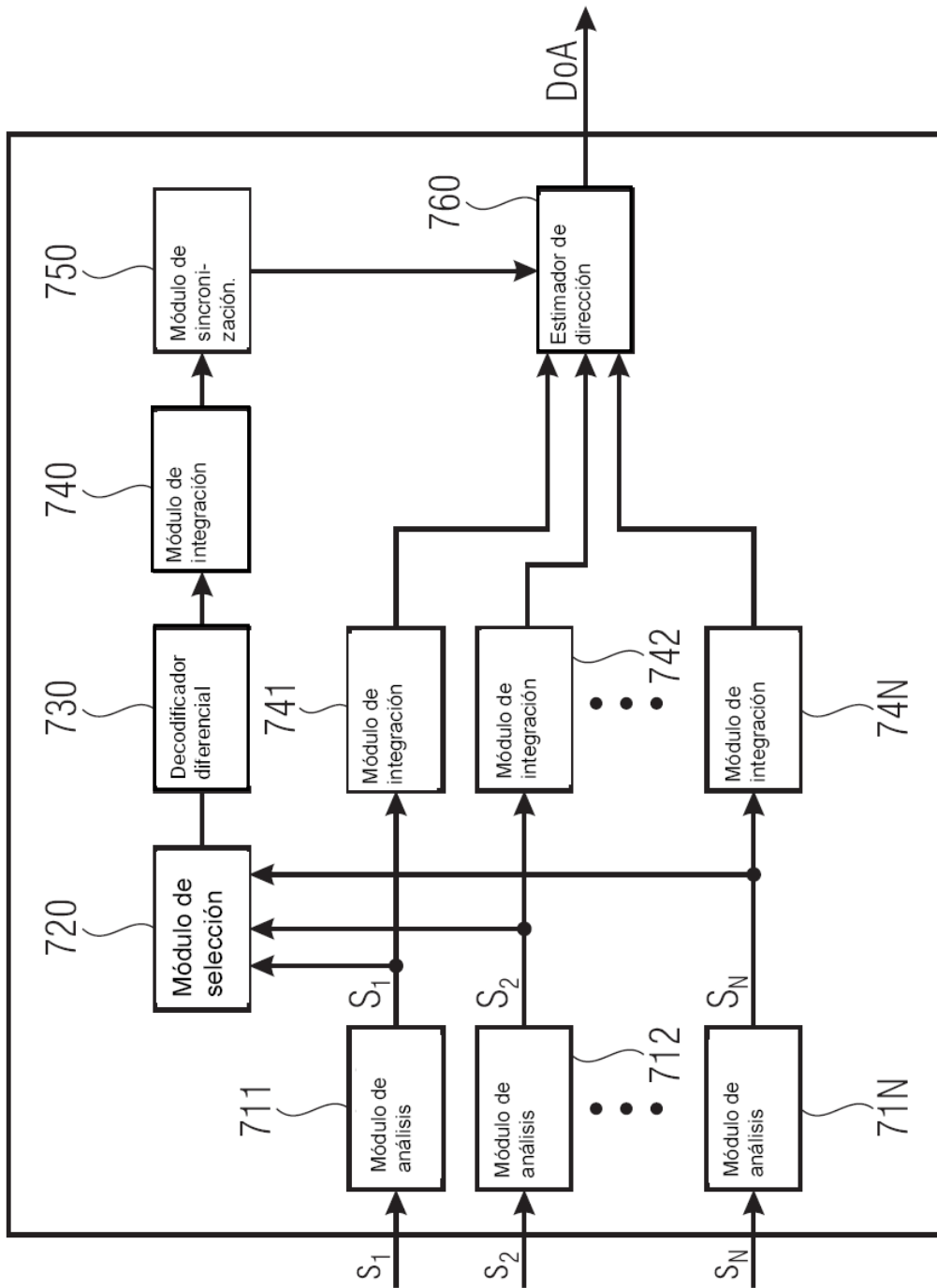


FIG 7

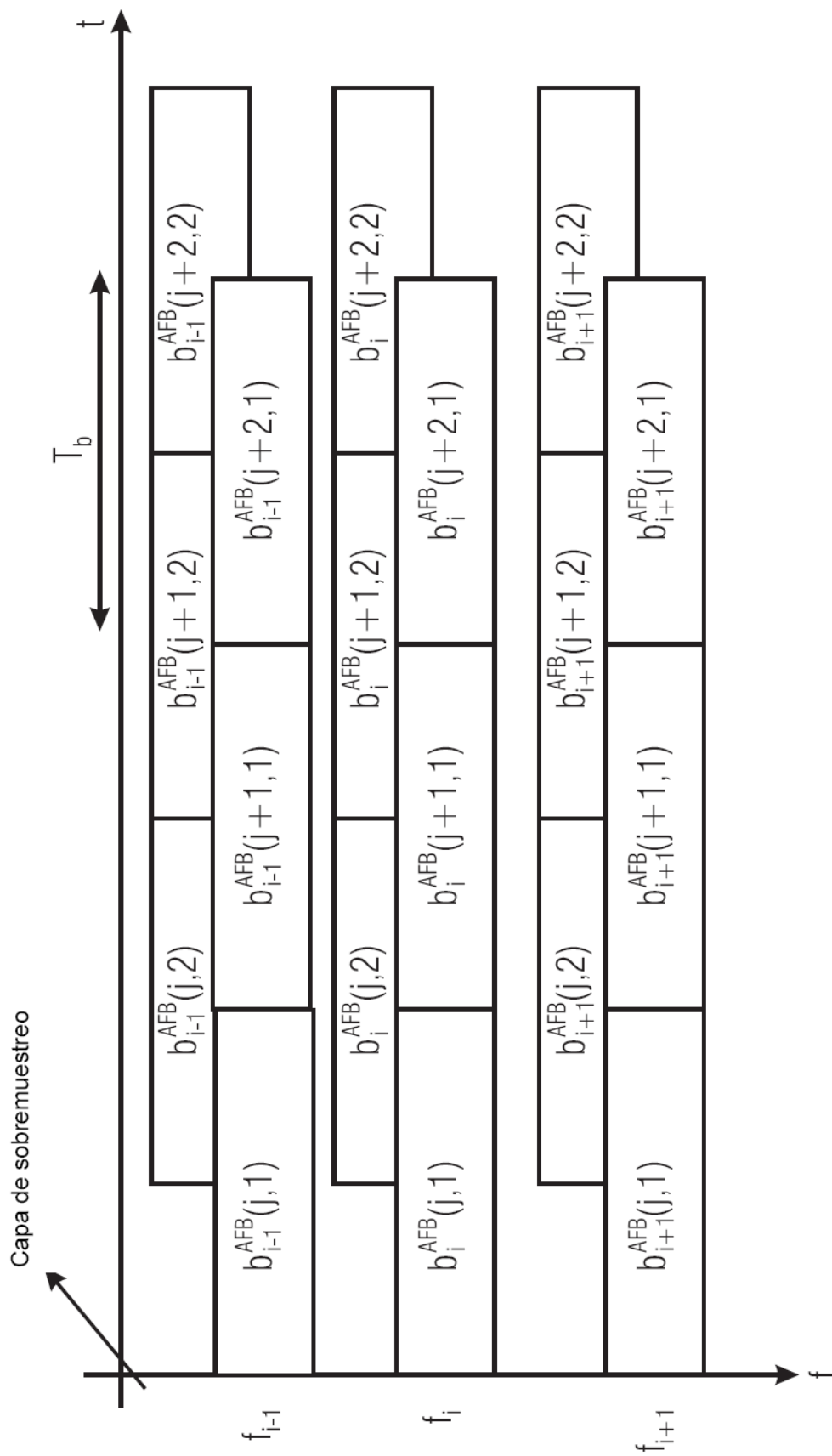


FIG 8A

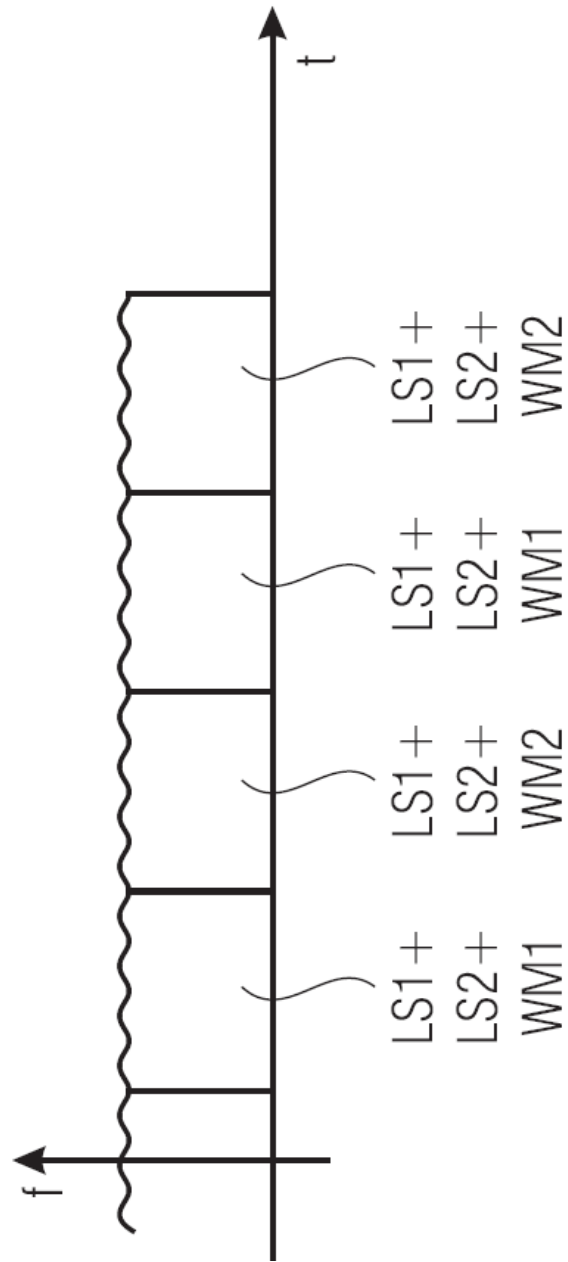


FIG 8B

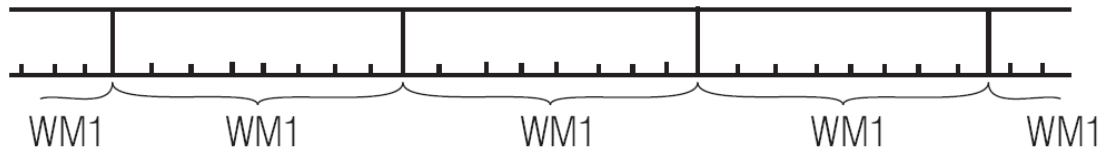


FIG 9A

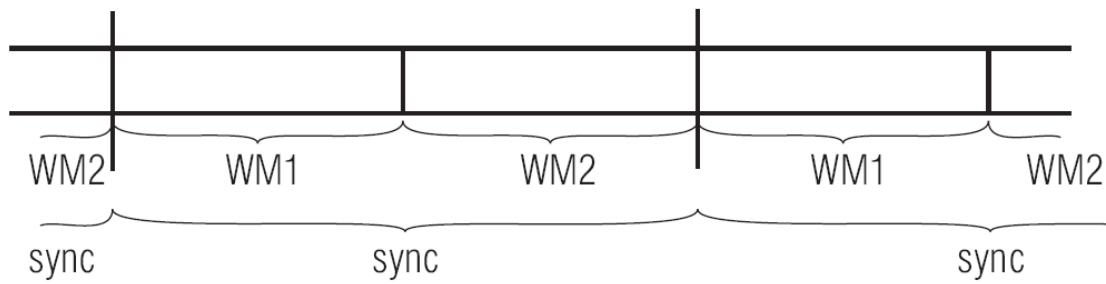


FIG 9B

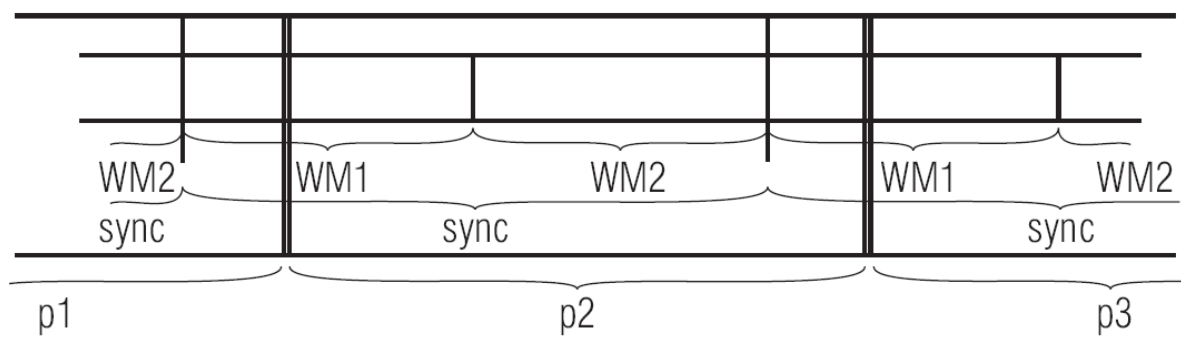


FIG 9C

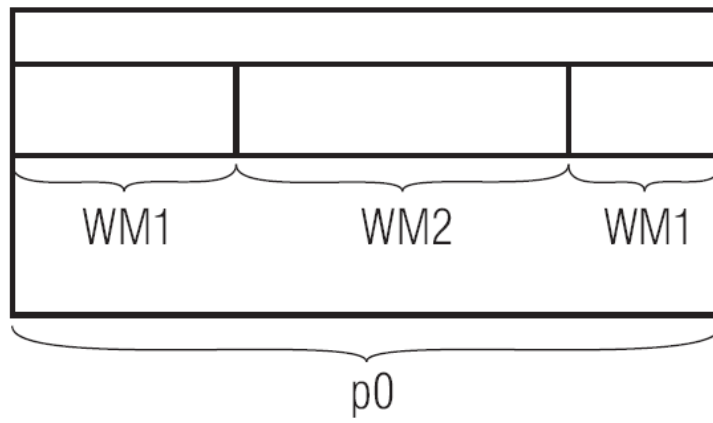


FIG 9D

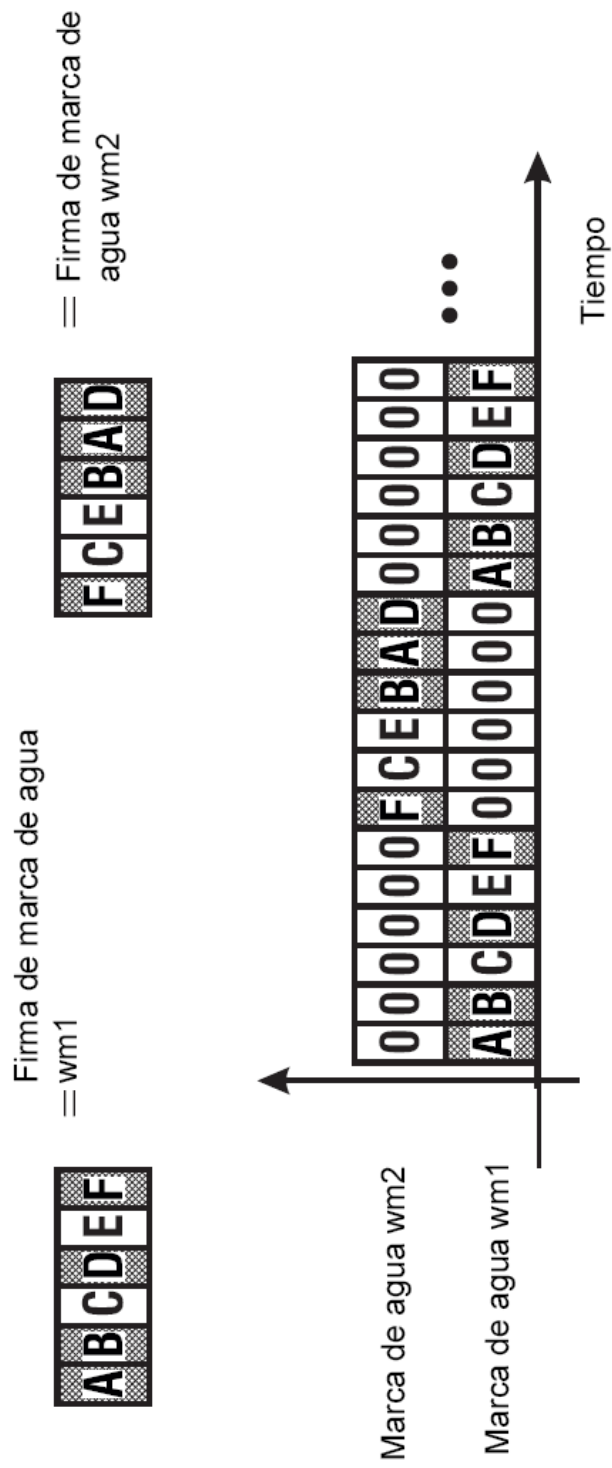


FIG 10

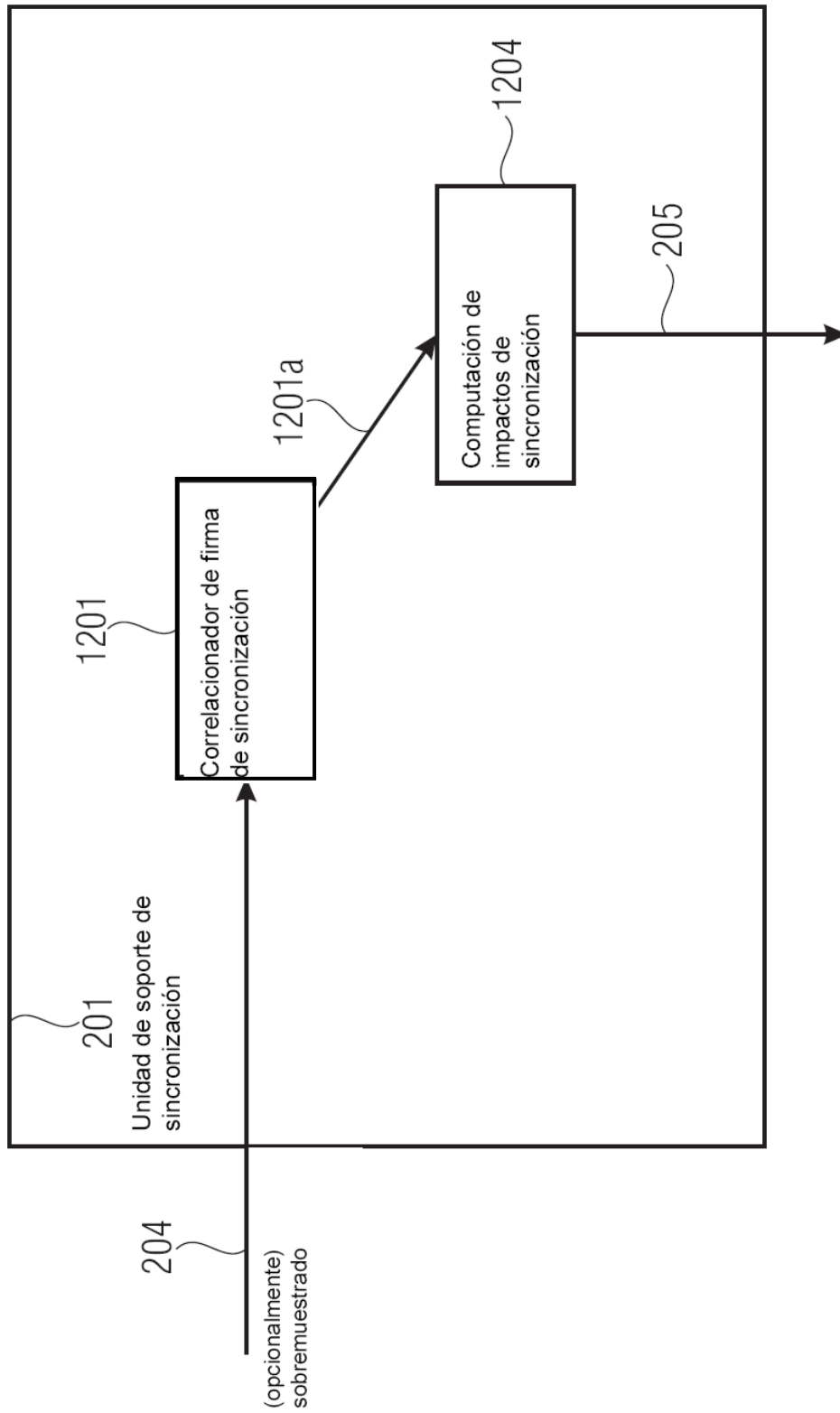


FIG 11A

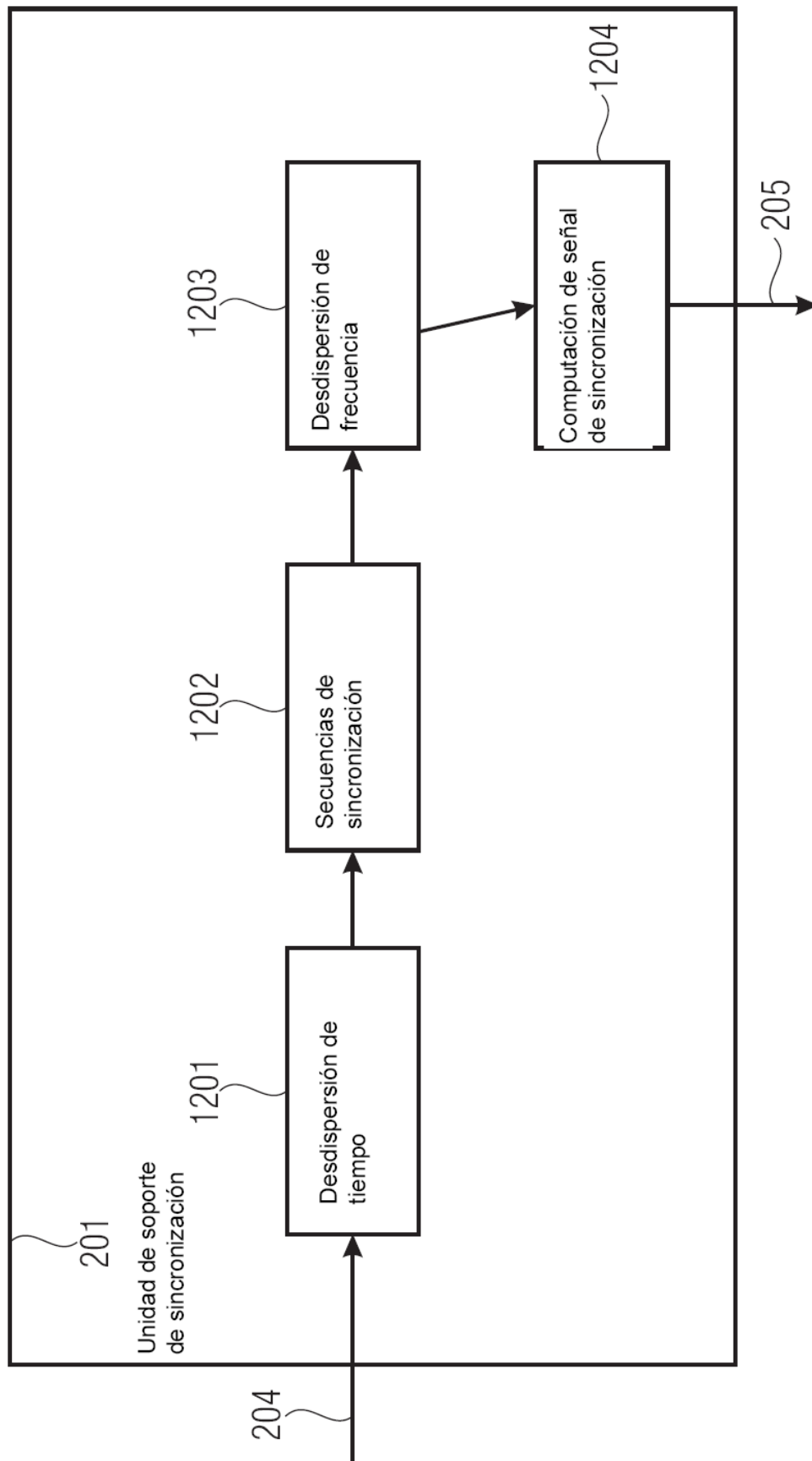


FIG 11B

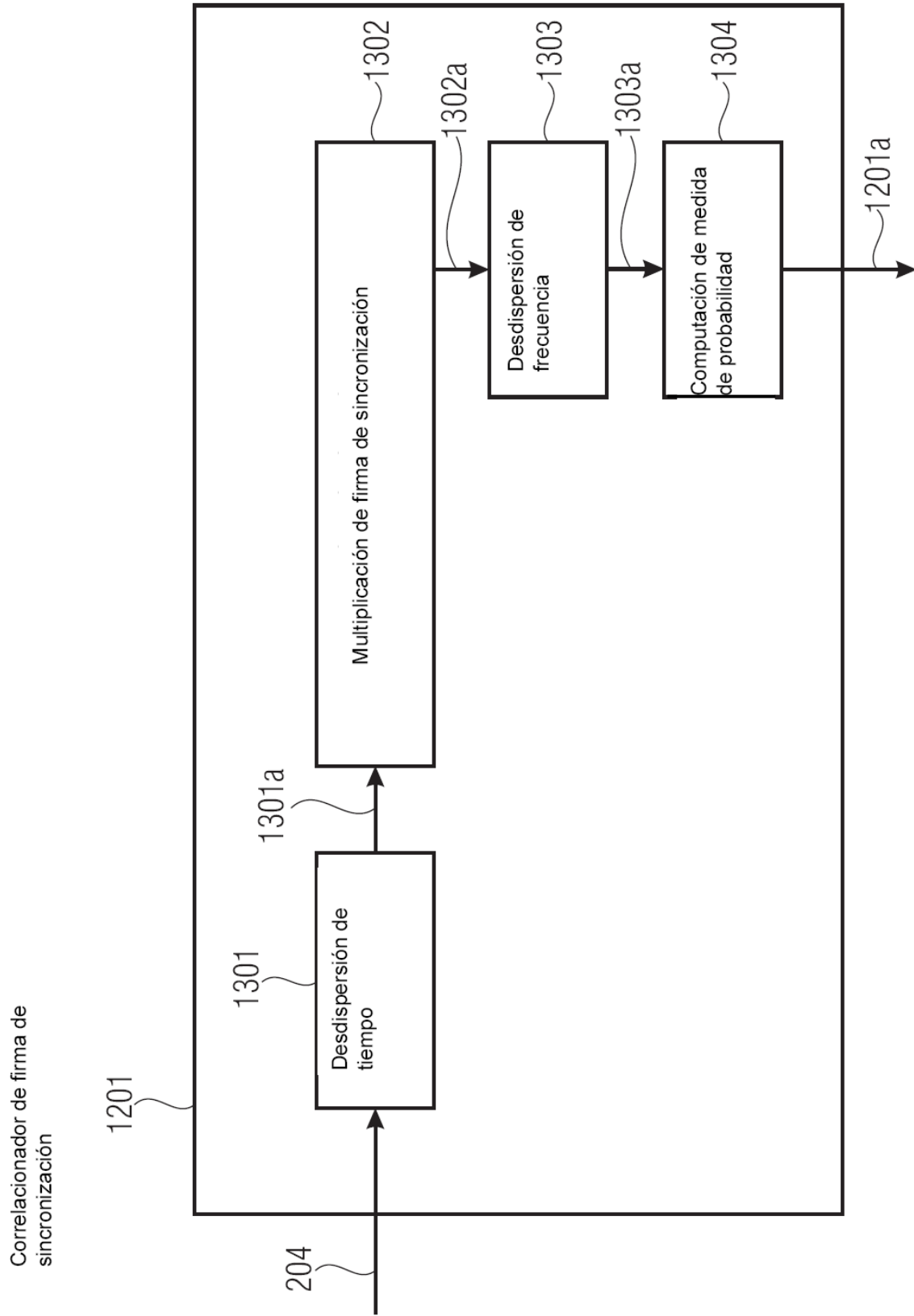


FIG 11C

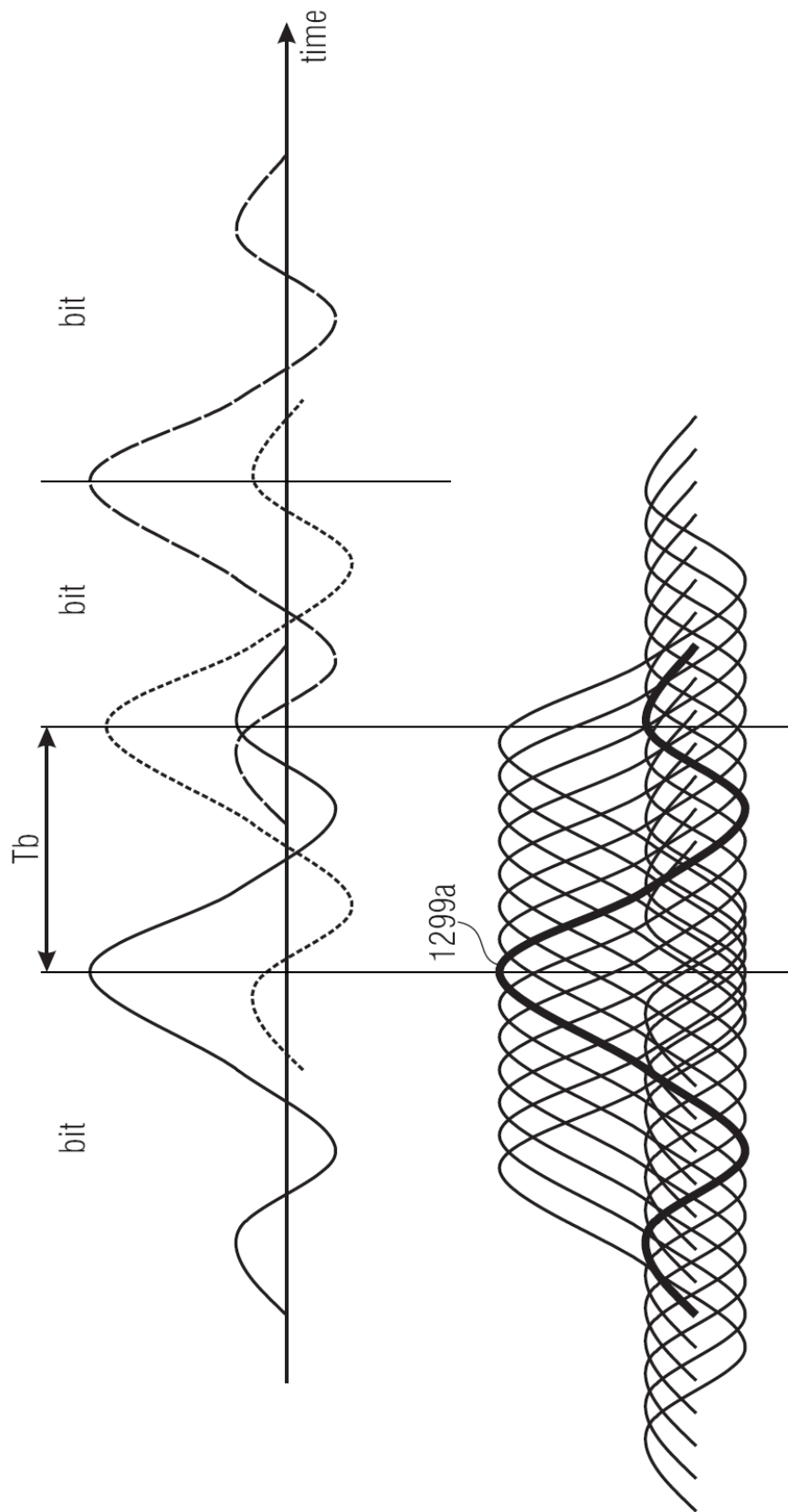


FIG 12A

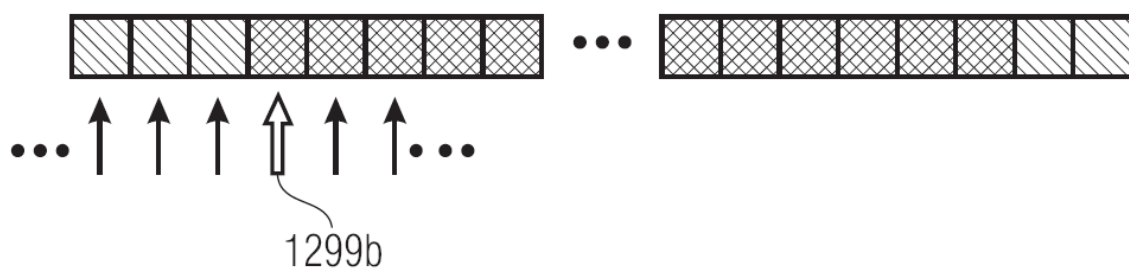


FIG 12B

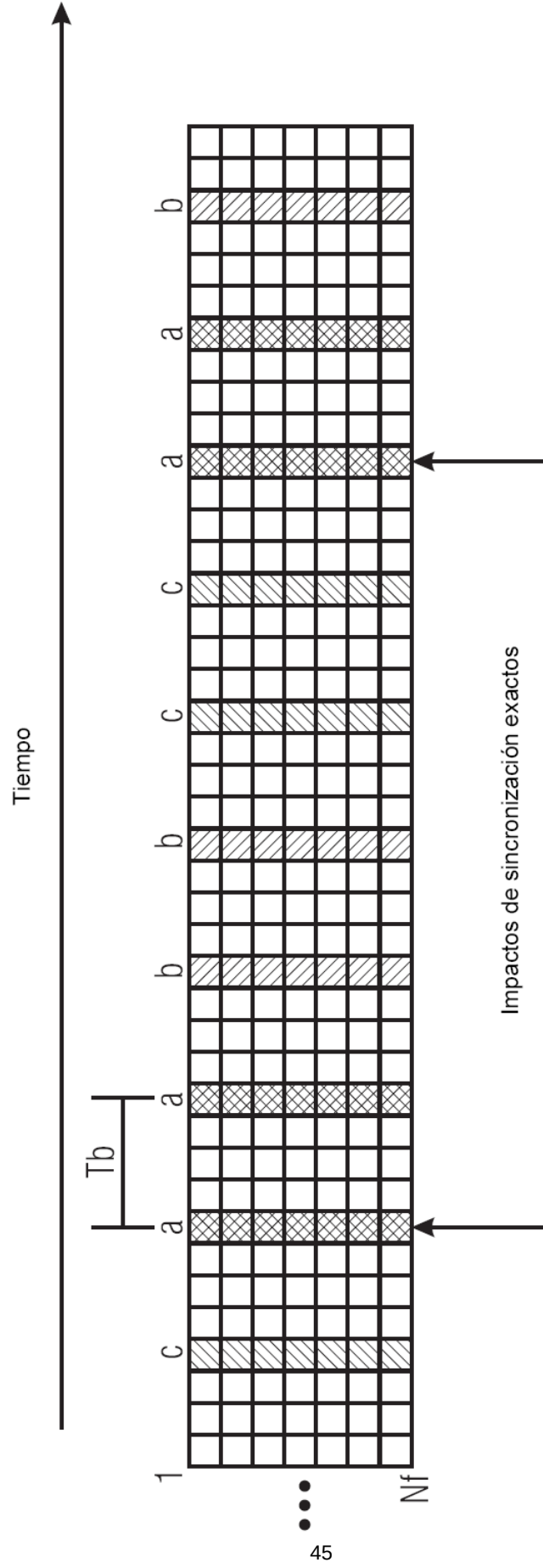


FIG 12C

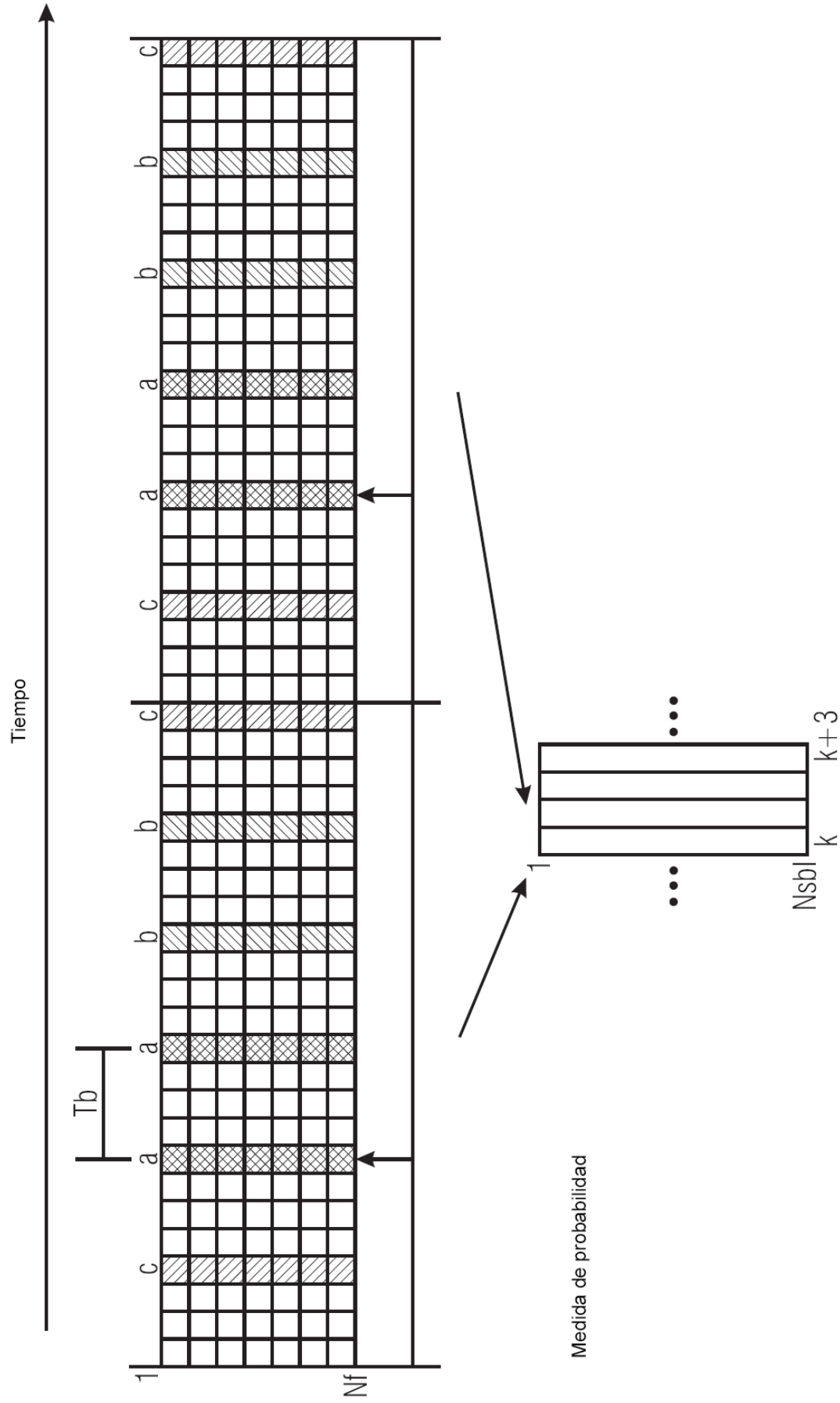


FIG 12D

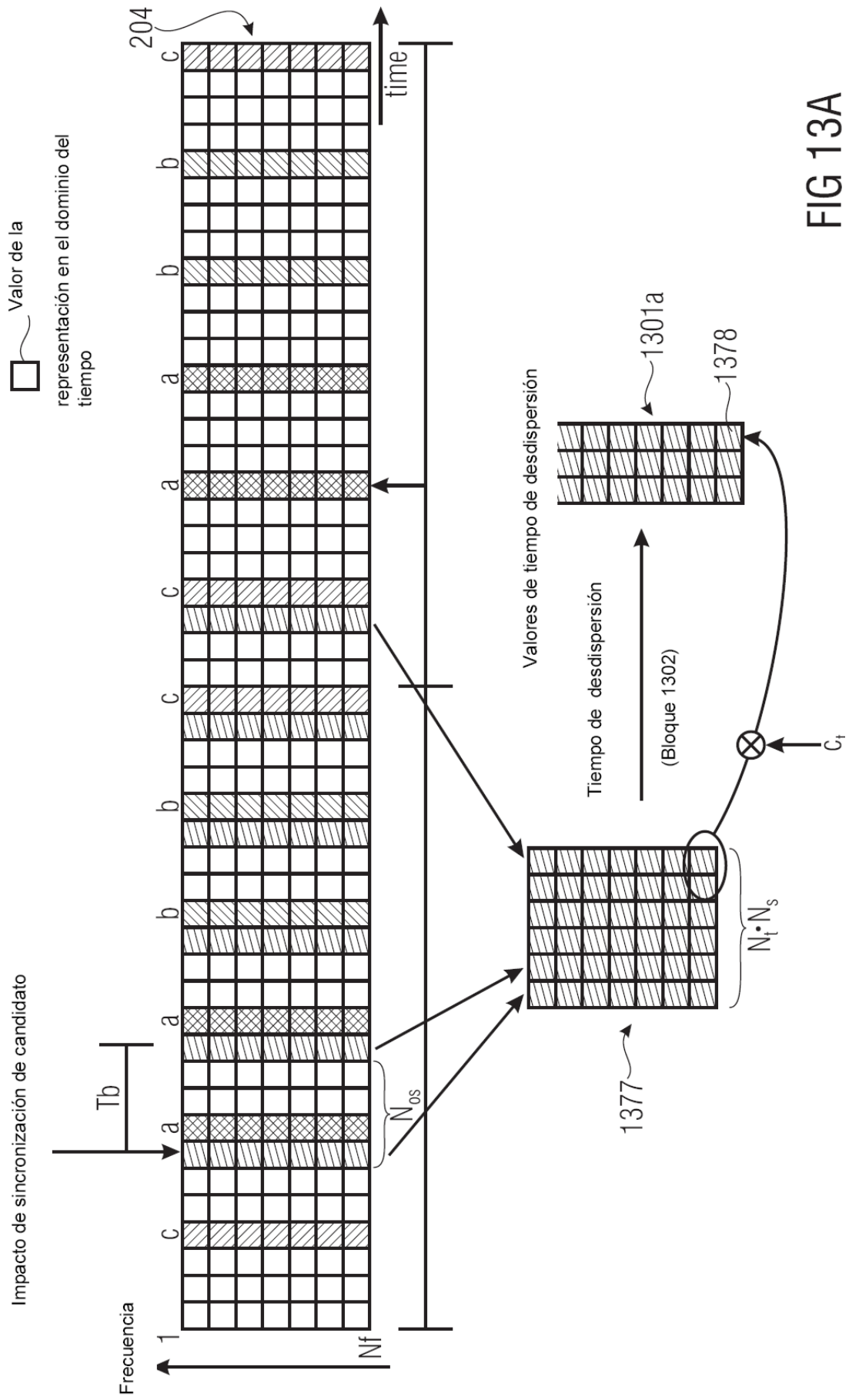


FIG 13A

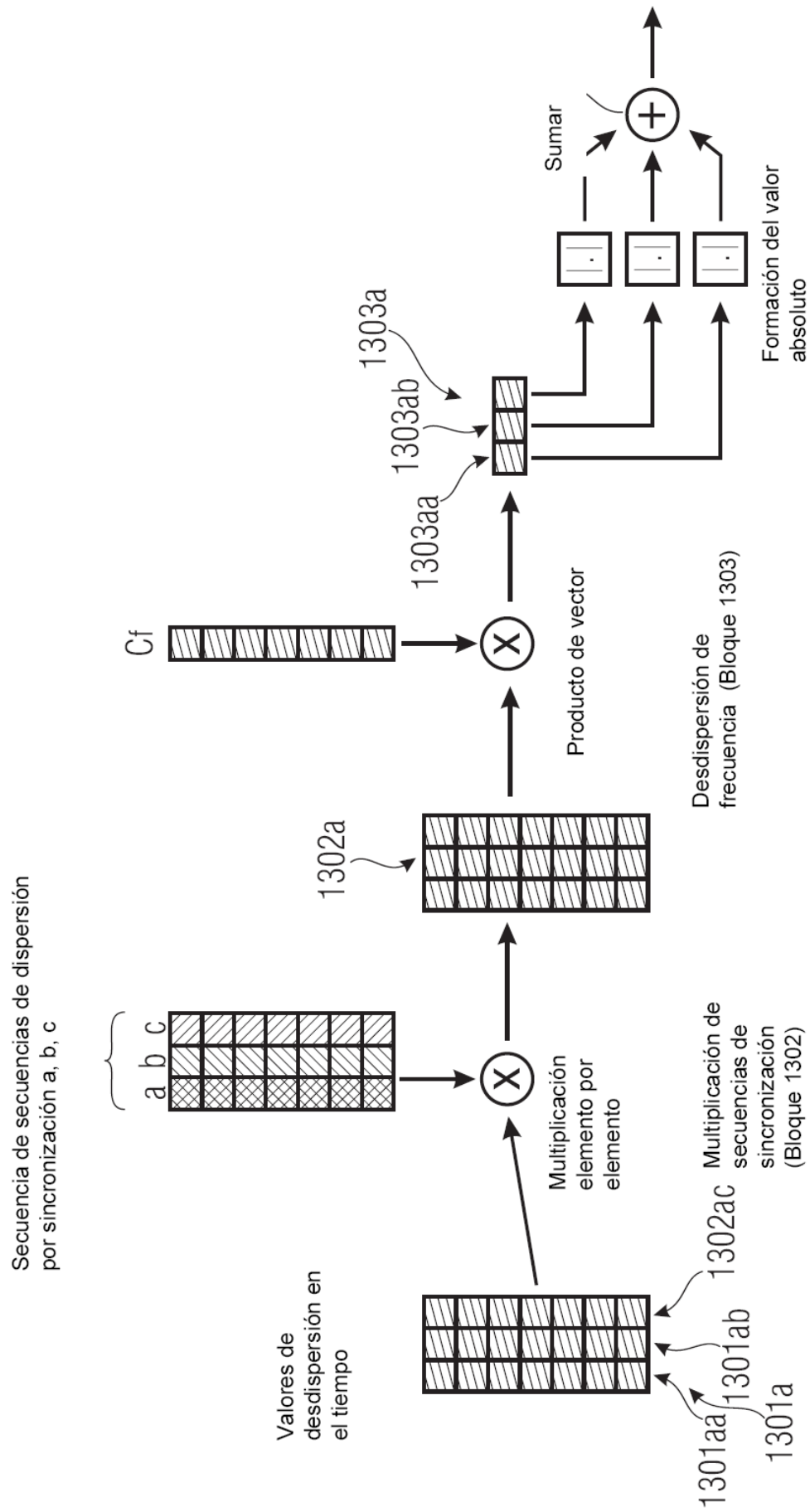


FIG 13B

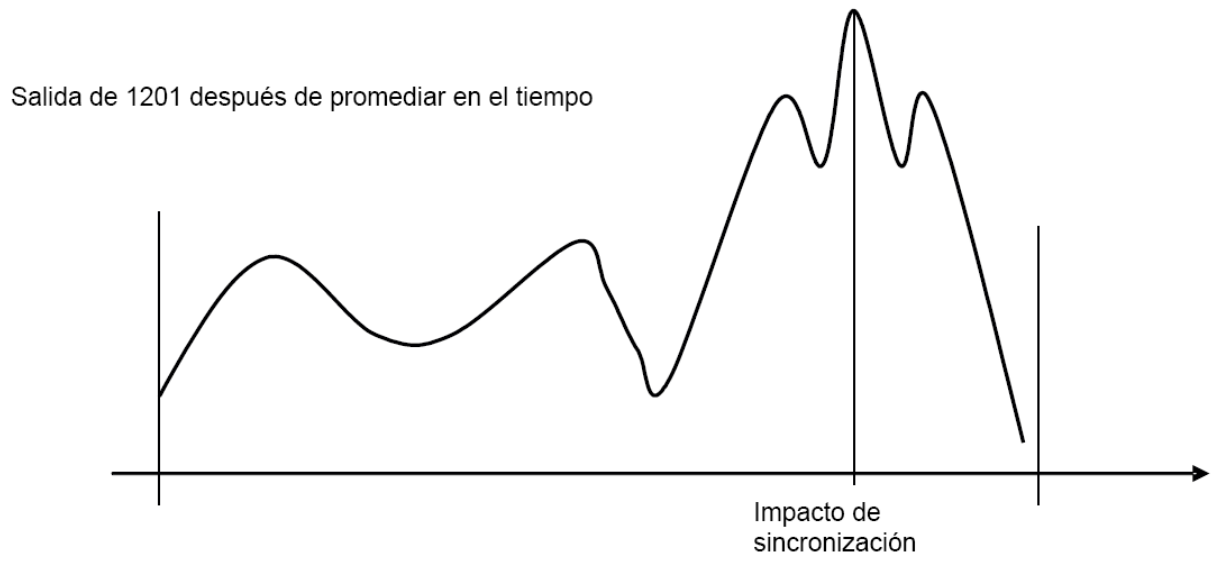


FIG 13C

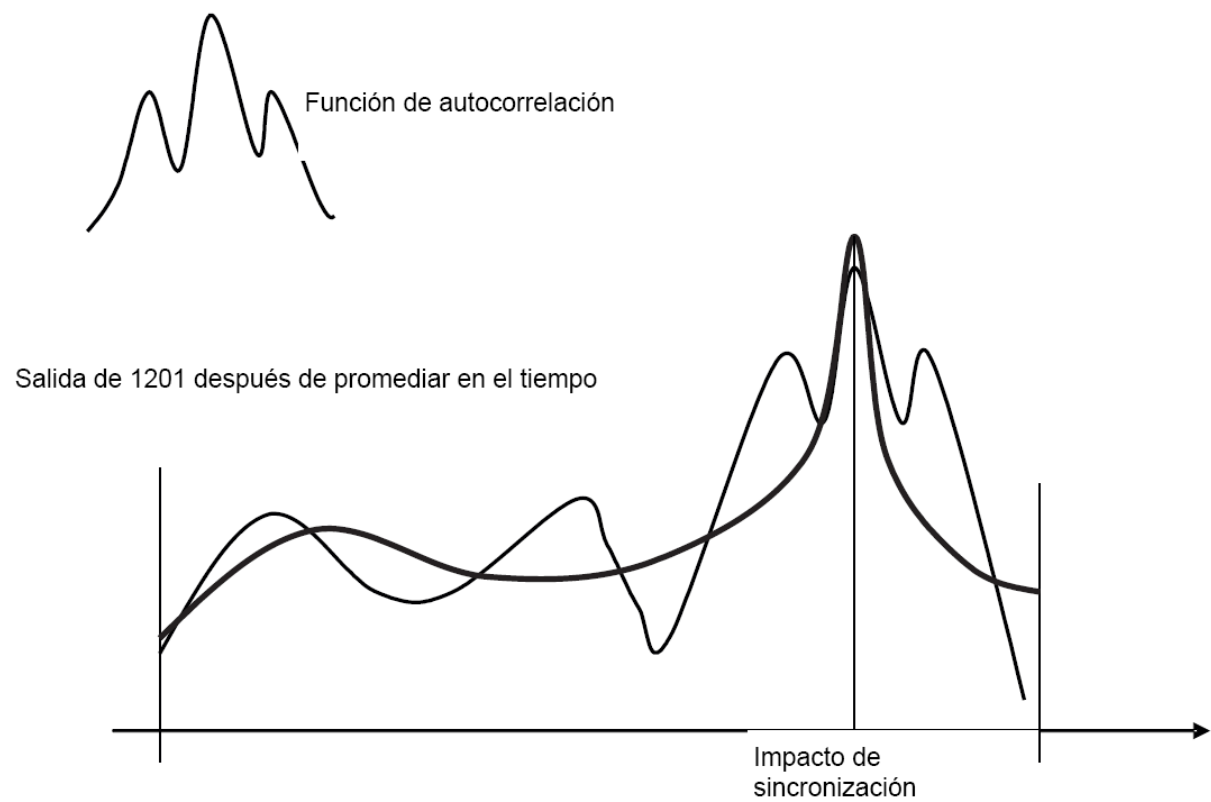


FIG 13D

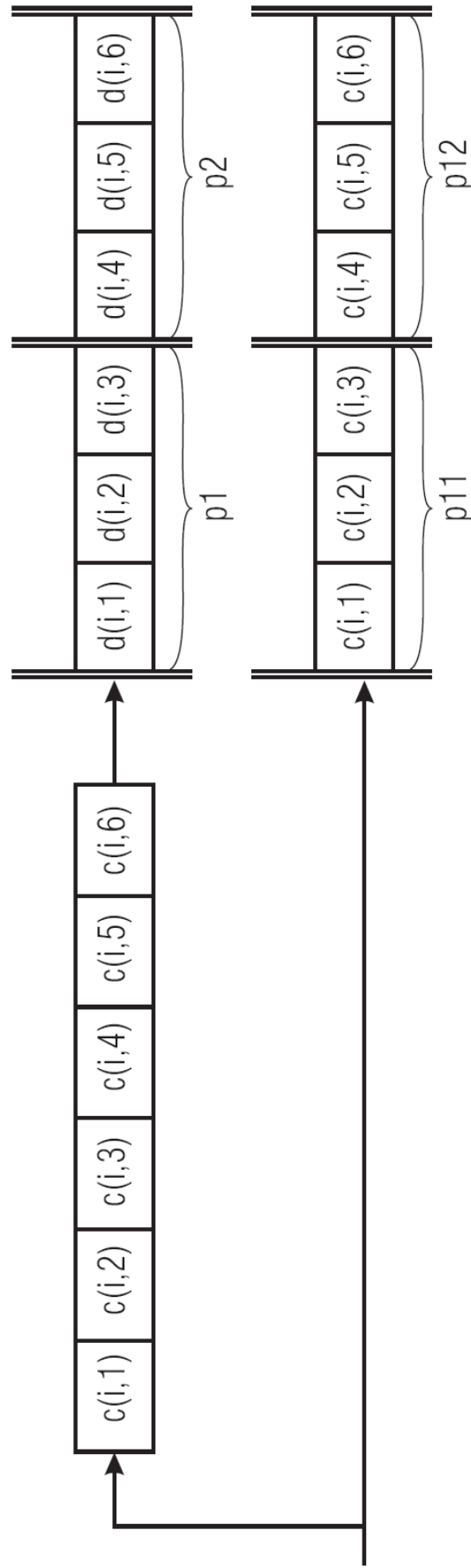


FIG 14

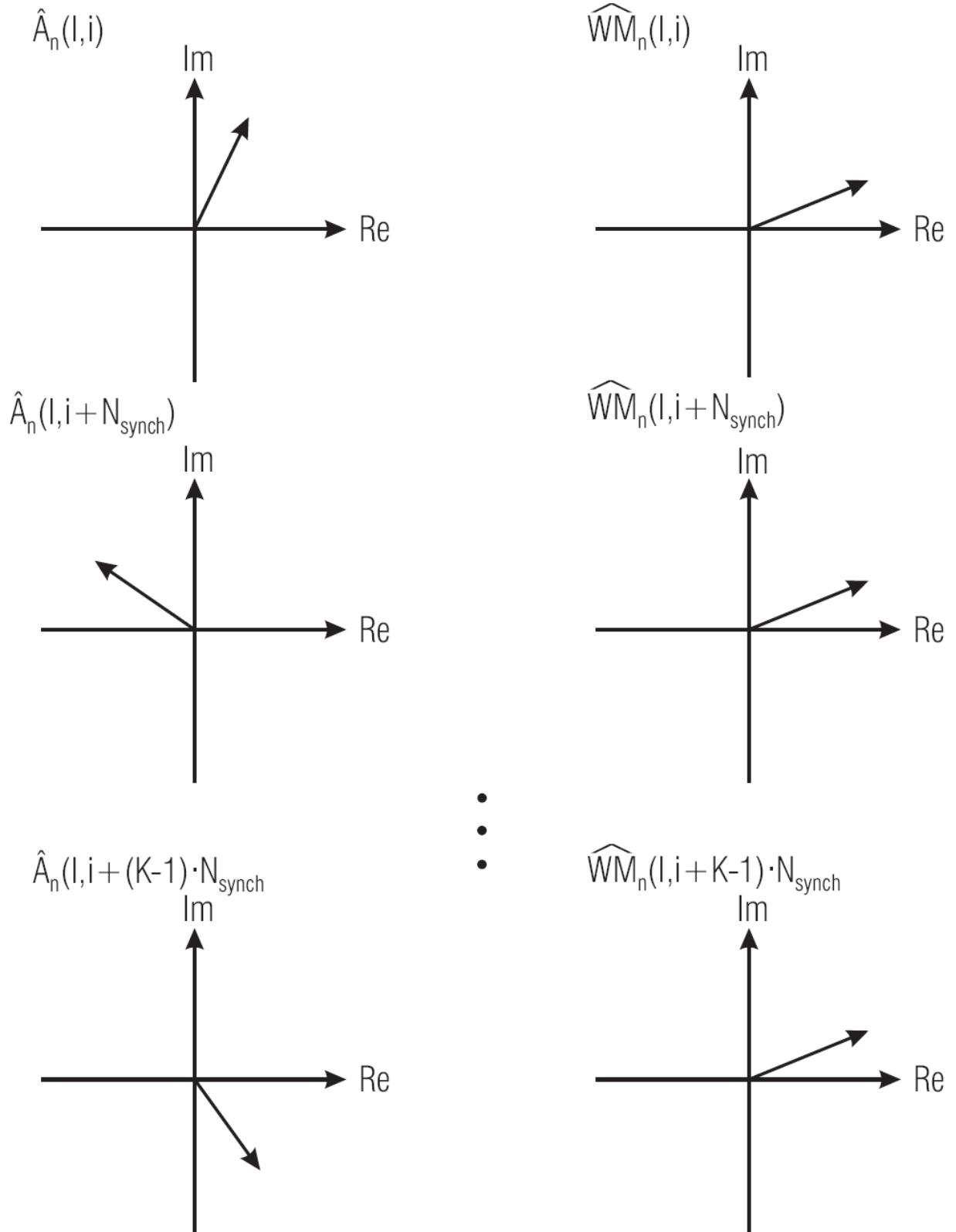


FIG 15A

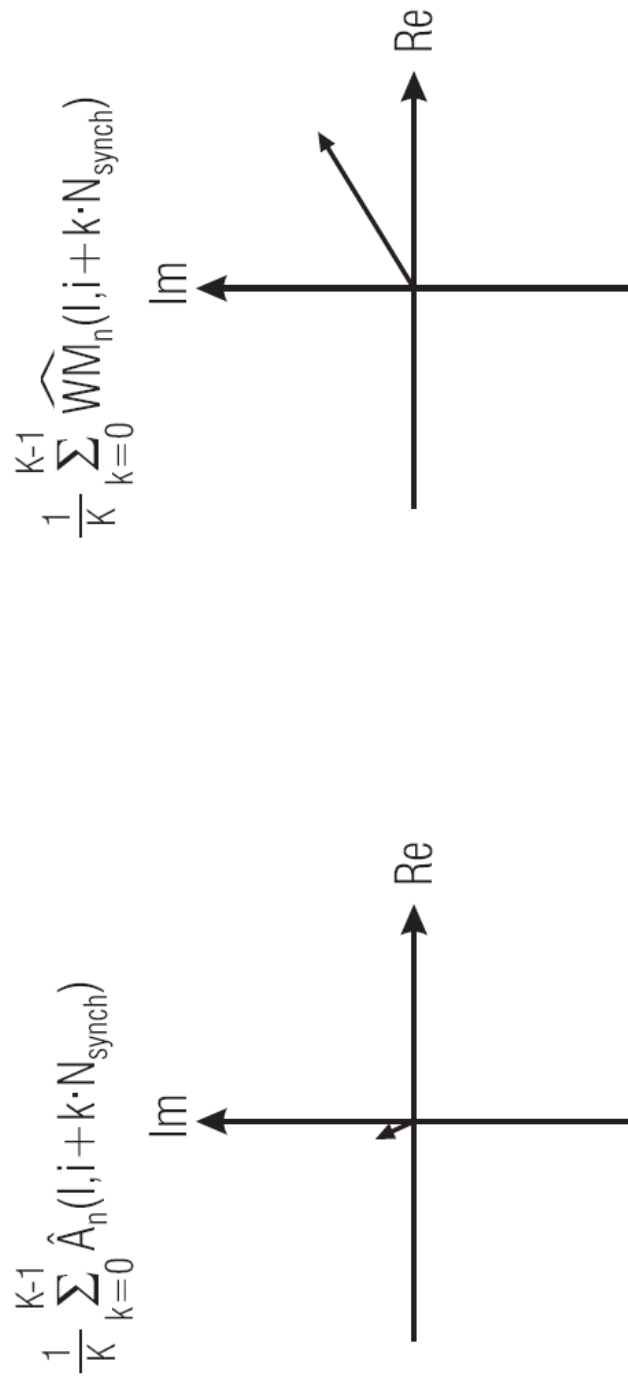


FIG 15B

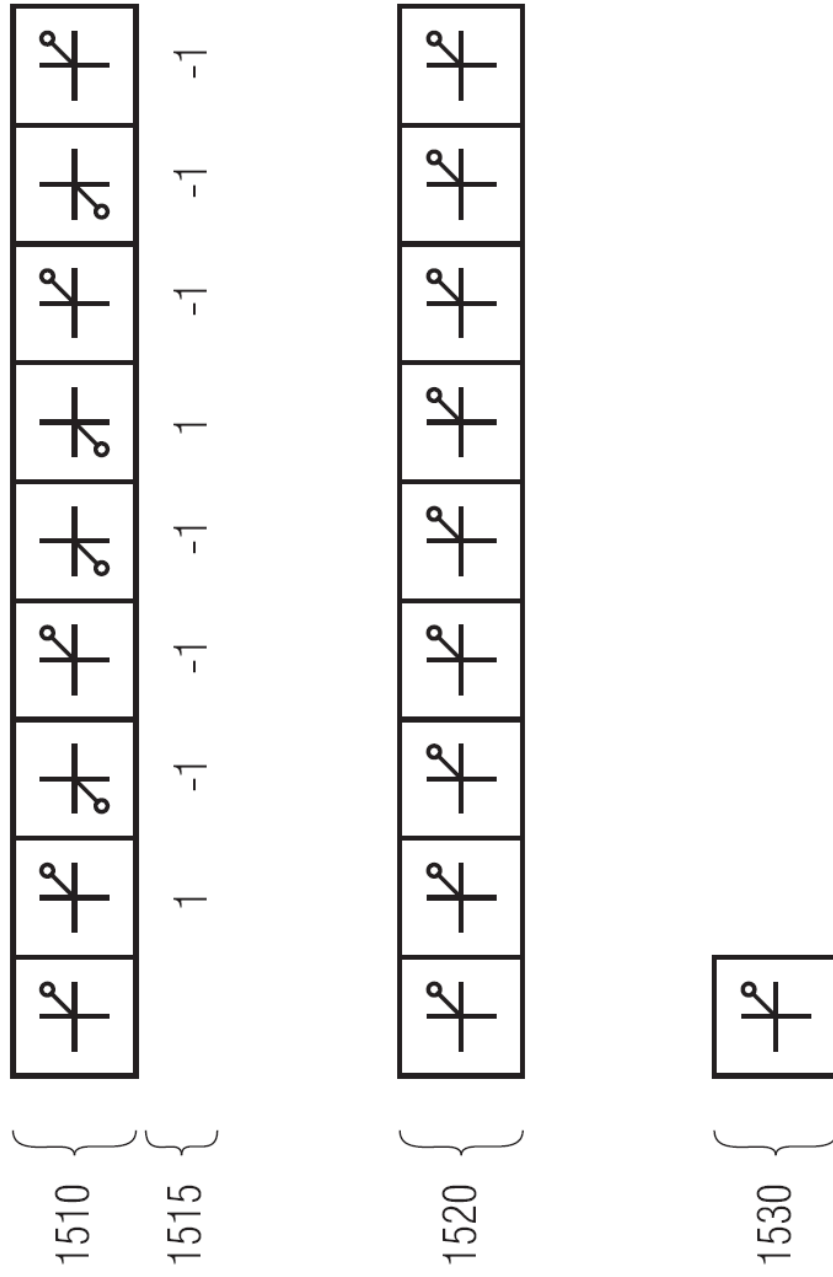


FIG 15C

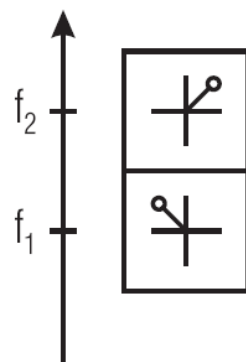


FIG 15D

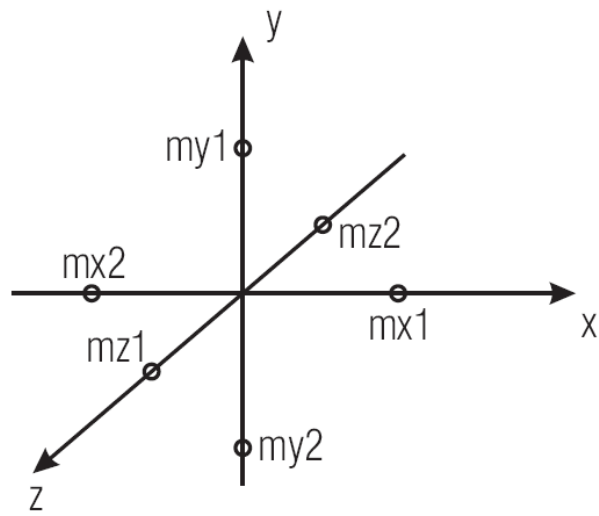


FIG 16

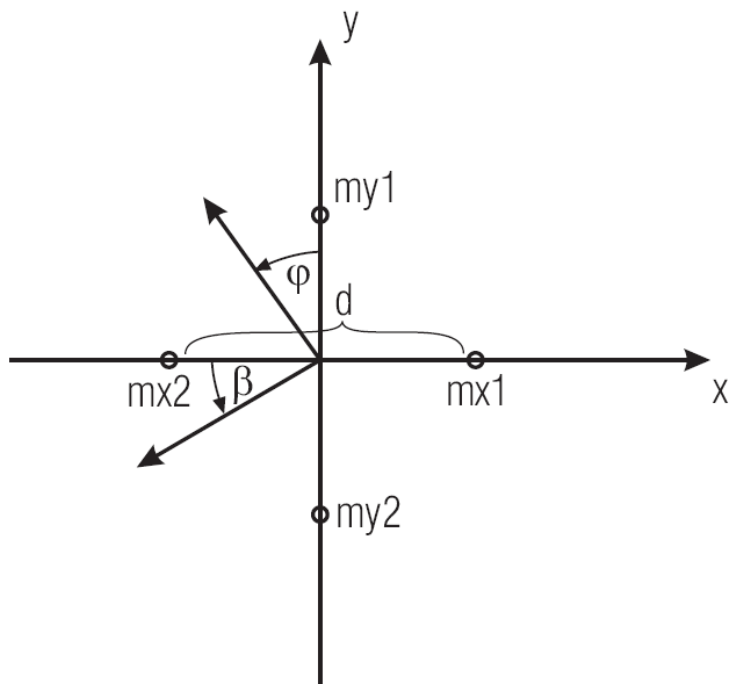


FIG 17

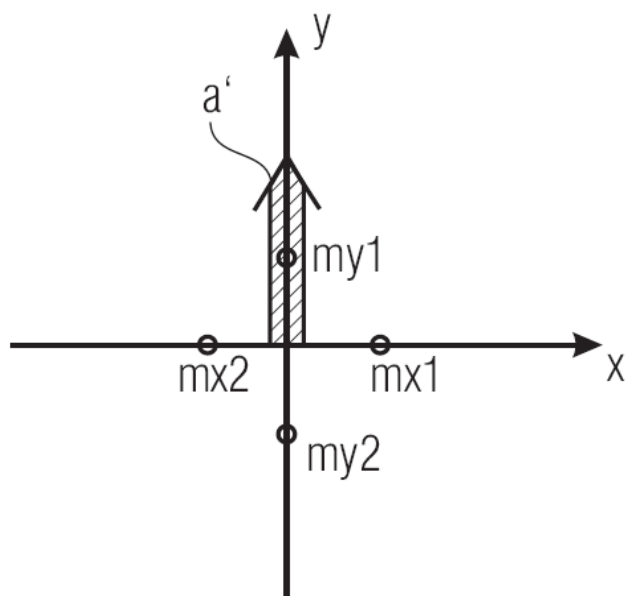


FIG 18A

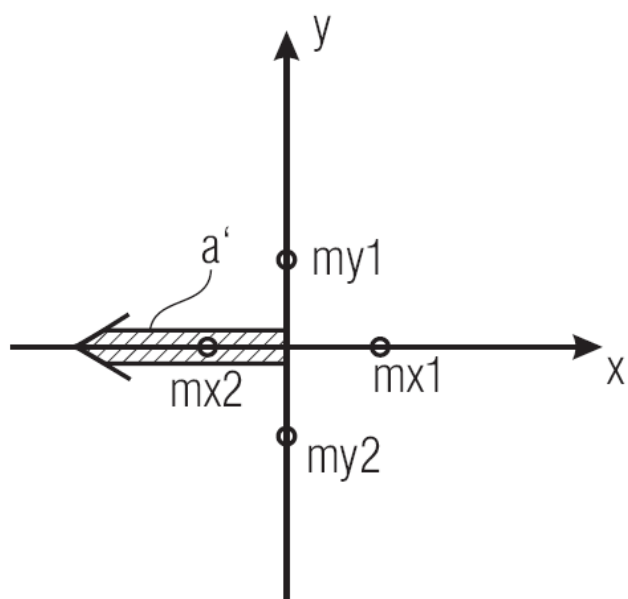


FIG 18B

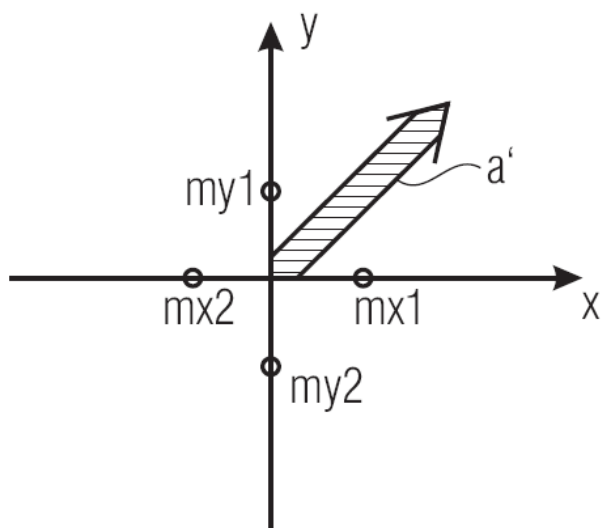


FIG 18C

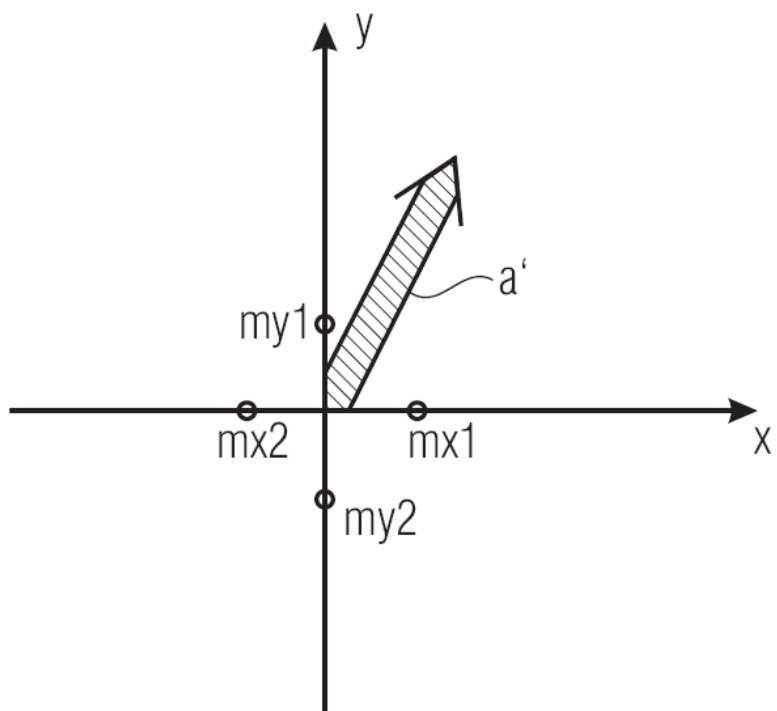


FIG 18D

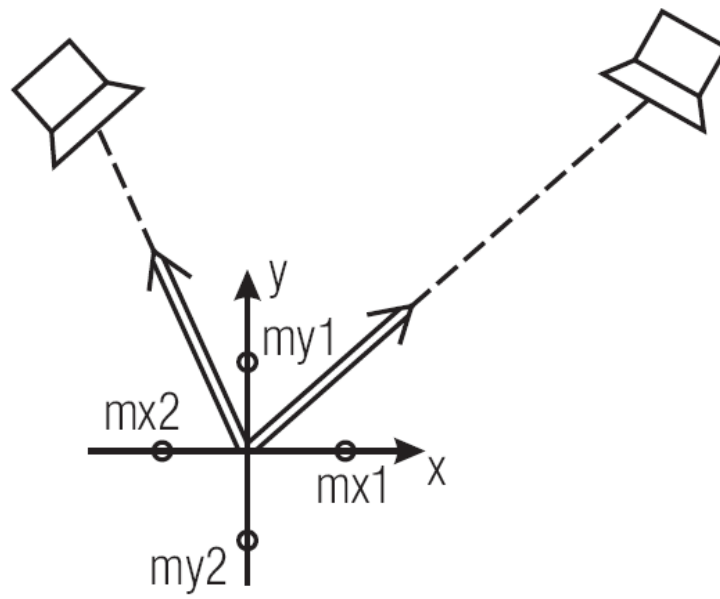


FIG 19

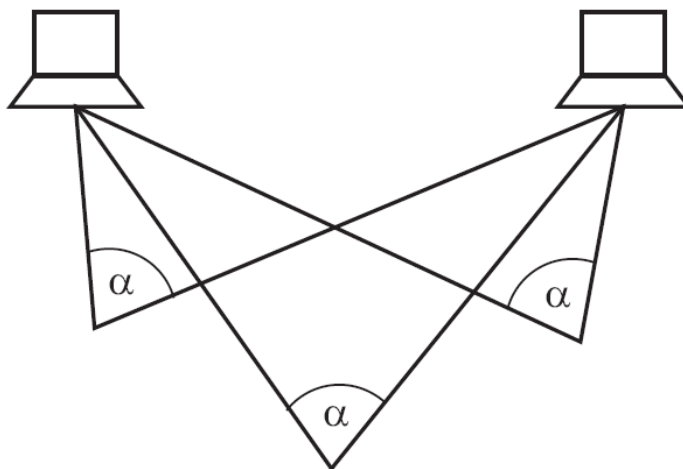


FIG 20A

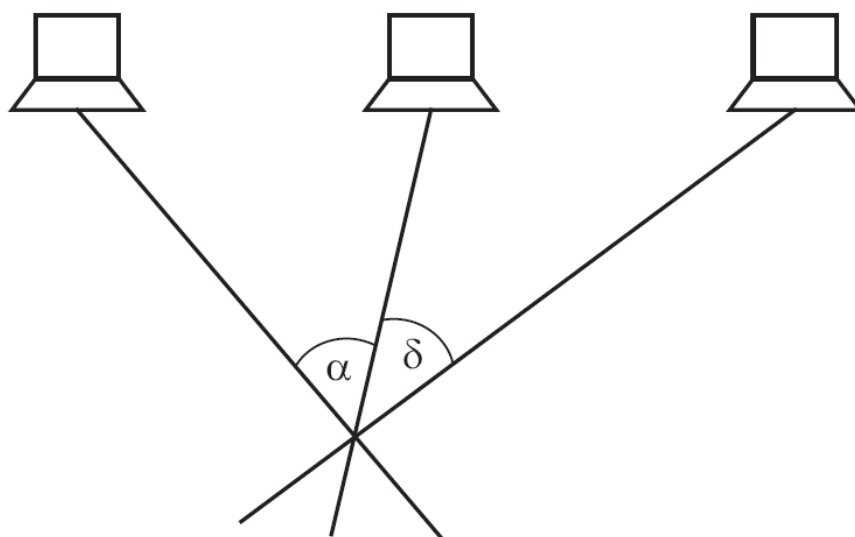


FIG 20B

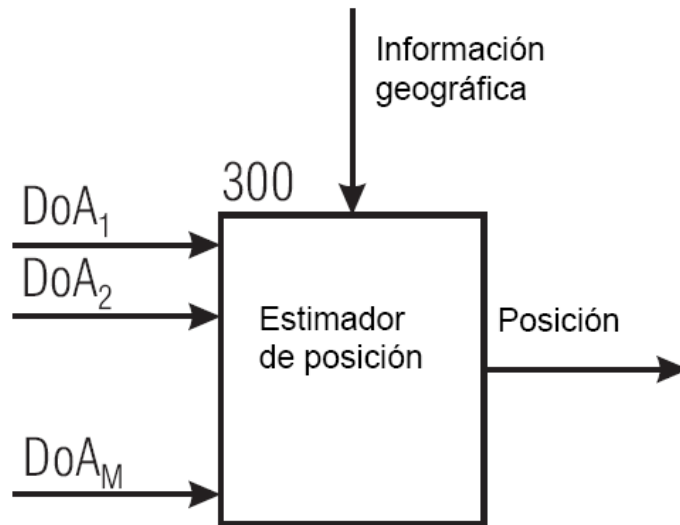


FIG 21

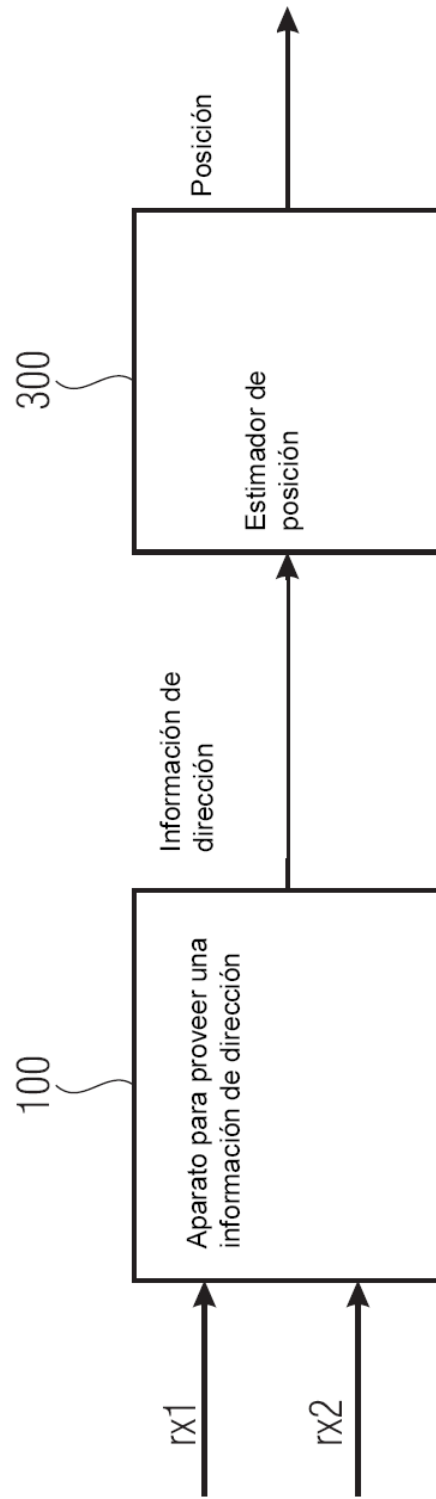


FIG 22