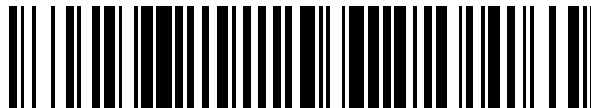


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 401 975**

51 Int. Cl.:

H04M 3/56

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.12.2006 E 06841205 (5)**

97 Fecha y número de publicación de la concesión europea: **21.11.2012 EP 2103098**

54 Título: **Conferencia en cuyo transcurso la mezcla se somete a un control temporal por parte de un dispositivo de representación**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.04.2013

73 Titular/es:

**TELECOM ITALIA S.P.A. (100.0%)
PIAZZA DEGLI AFFARI, 2
20123 MILANO, IT**

72 Inventor/es:

VARESIO, ANDREA

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 401 975 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

Conferencia en cuyo transcurso la mezcla se somete a un control temporal por parte de un dispositivo de representación

Antecedentes de la invención

5 **Sector de la invención**

[0001] La presente invención se refiere generalmente al sector de las telecomunicaciones, y concretamente a las conferencias de audio o de audio/ vídeo. Específicamente, la invención se refiere a un terminal de telecomunicaciones que aloja a un mezclador de conferencia de audio, o de audio/ video.

Descripción de la técnica relacionada

10 [0002] En el sector de las telecomunicaciones, la difusión de servicios de Voz sobre Protocolo de Internet (VoIP), y de dispositivos compatibles con ellos, está creciendo rápidamente. Un crecimiento rápido similar han experimentado los servicios y dispositivos de apoyo por comunicación de vídeo (VCD).

15 [0003] Muy a menudo, los servicios de este tipo implican dos pares que se intercomunican, pero una extensión interesante la representa la conferencia "virtual" de audio y/ o video, donde más de dos partes ("pares") están involucradas en la sesión de comunicación de audio y/ o vídeo, y pueden interactuar unas con otras por escucha/ habla y/ o visualización.

20 [0004] Los aparatos que permiten conferencias virtuales de audio y/ o video se conocen como "mezcladores de conferencia". Esencialmente, un mezclador de conferencia reúne los contenidos de audio y/ o de vídeo generados mediante dispositivos de captura locales (micrófonos, videocámaras) previstos en terminales de usuario (los "puntos finales") en cada una de las partes de conferencia, mezcla adecuadamente los contenidos de audio y/ o de vídeo generados, y redistribuyen los contenidos mezclados a cada parte en la conferencia virtual.

25 [0005] Convencionalmente, los mezcladores de conferencias son aparatos distintos y alejados de los puntos finales de las partes en conferencia, que son aparatos de red de núcleo básicos (denominados como "Unidades de Control Maestro", con las siglas MCUs).

[0006] Se conocen en la técnica soluciones para funciones de mezcla en los puntos finales de los pares de conferencias.

30 [0007] Por ejemplo, en la solicitud de patente publicada de EE.UU. 2003/0142662 se da a conocer un terminal de paquete de datos, en particular un ordenador personal, asistente digital personal, teléfono, radioteléfono móvil, dispositivo de acceso a la red, periféricos de Internet y similares, que inicia, coordina y controla la prestación de servicios a la carta de llamada de conferencia, con poco o sin apoyo de la red. Un convertidor de digital a analógico para la conversión convierte flujos de paquete de datos primero y segundo en representación analógica separada; un mezclador selectivo manipula las representaciones analógicas para proporcionar una salida mixta; un circuito multiplexor distribuye los flujos de paquetes de datos a una pluralidad de sesiones de llamadas.

[0008] EP 778714 divulga una rutina de puenteo implementada en software para un sistema de teléfono audio totalmente doble.

[0009] EP 855827 divulga A un procedimiento de conferencia que no requiere hardware adicional o dedicado en un terminal telefónico controlado por procesador.

40 [0010] La presente invención difiere de lo divulgado en estos dos documentos al menos por el hecho de que un primer mezclador de porciones de datos de audio está configurado para realizar las operaciones de:

- a) obtener porciones de corriente de datos de audio almacenadas en unos primeros búfers de datos;
- b) mezclar las porciones de corriente de datos de audio obtenidas a partir de los primeros búfers de datos para producir una primera porción de datos de audio mezclada;

45 y

c) suministrar la primera porción de datos de audio mezclada a un dispositivo de representación de un terminal de telecomunicaciones, a partir de la recepción de una notificación desde dicho dispositivo de representación que indica que el dispositivo de representación está preparado para representar una nueva porción de datos de audio mezclada.

[0011] JP 05 037655 divulga un sistema de comunicación multi-punto de voz en el que el deterioro en la calidad de voz se reduce notablemente.

[0012] EP 617537 divulga una técnica de puenteo de conferencia para señales de información comprimidas que monitoriza la energía de señal transmitida por cada conferenciante.

5 [0013] En estos dos documentos, la función de mezclado se activa cuando se inserta una nueva corriente de audio; este proceso introduce un transitorio pero bastante retardo.

Resumen de la invención

10 [0014] El Solicitante observa por un lado que la implementación de servicios de conferencia virtual audio o audio/ vídeo basada en la provisión de componentes de red básicos dedicados (los MCUs) no es satisfactoria, sobre todo porque afecta a la estructura de la red de teléfono/ telecomunicaciones, e implica costes para los operadores de red. Por lo tanto, el solicitante cree que una implementación diferente de los servicios de conferencias virtuales, en los que una funcionalidad de mezcla de conferencia audio o audio/ vídeo se aloja en el punto final de al menos uno de los pares que participan en la conferencia virtual de audio o audio/ vídeo es mejor, debido a que esencialmente no tiene impacto en la red de telefonía/ telecomunicaciones.

15 [0015] No obstante, el solicitante ha observado que un aspecto importante que queda por considerar cuidadosamente, es la reducción, en la medida de lo posible, del retardo de extremo a extremo que es experimentado por los pares que participan en una conferencia virtual.

20 [0016] El solicitante ha abordado el problema de cómo reducir el retardo de extremo a extremo en servicios de conferencias virtual de audio o de audio/ vídeo que deben ser realizados a través de un mezclador de conferencia alojado en un extremo de uno de los pares de la conferencia.

[0017] El solicitante ha encontrado que el retardo de extremo a extremo experimentado en conferencias virtuales de audio o audio/ vídeo puede ser reducido, siempre que las operaciones de mezcla estén temporizadas en el o los dispositivos de representación y/ o el o los dispositivos de captura del extremo que aloja el mezclador de conferencia.

25 [0018] Según un aspecto de la presente invención, se proporciona un terminal de telecomunicaciones, que aloja un mezclador de conferencia adaptado para habilitar al menos una conferencia audio entre un primer par de conferencia y al menos dos pares de conferencia adicionales. El mezclador de conferencia comprende:

30 - para cada uno de los al menos dos pares de conferencia adicionales, un primer búfer de datos respectivo configurado para almacenar en búfer porciones de al menos una corriente de datos de audio recibida desde el par de conferencia respectivo;

- un primer mezclador de porciones de datos de audio suministrados por los primeros búfers de datos y configurado para:

a) obtener porciones de corriente de datos de audio almacenadas en los primeros búfers de datos;

35 b) mezclar las porciones de corriente de datos de audio obtenidas a partir de los primeros búfers de datos para producir una primera porción de datos de audio mezclada;

y

c) suministrar la primera porción de datos de audio mezclada a un dispositivo de representación del terminal de telecomunicaciones,

40 en el que dicho primer mezclador de porciones de datos de audio está configurado para realizar las operaciones a), b) y c) a partir de la recepción de una notificación desde dicho dispositivo de representación que indica que el dispositivo de representación está preparado para representar una nueva porción de datos de audio mezclada.

[0019] Para los propósitos de la presente invención, por "conferencia de audio" se entiende una conferencia virtual entre tres o más pares, que incluye al menos audio. Posiblemente, la conferencia audio también podría incluir vídeo, es decir que podría ser una conferencia virtual audio/ vídeo.

45 [0020] Según otro aspecto de la presente invención, se trata de un procedimiento para realizar al menos una conferencia audio entre un primer par de conferencia y al menos dos pares de conferencia adicionales, comprendiendo el procedimiento:

50 - en un terminal de telecomunicaciones del primer par de conferencia, realizar un primer almacenamiento en búfer de porciones de al menos una corriente de datos de audio recibida desde cada uno de los al menos dos pares de conferencia adicionales;

y

- a partir de la recepción de una notificación desde un dispositivo de representación del terminal de telecomunicaciones, que indica que el dispositivo de representación está preparado para representar una nueva porción de datos de audio mezclada:

- 5 - mezclar las porciones de corriente de datos de audio almacenadas en dicho primer almacenamiento en búfer, para producir la porción de datos de audio mezclada;

y

- suministrar la porción de datos de audio mezclada al dispositivo de representación.

[0021] Dicho dispositivo de representación puede incluir un altavoz.

- 10 **[0022]** Dicha realización de un primer almacenamiento en búfer puede incluir almacenar al menos una porción de la corriente de datos de audio recibida desde el par respectivo.

[0023] El procedimiento también puede incluir recibir dicho al menos una porción de la corriente de datos de audio a ser almacenada desde una función de representación respectiva adaptada para recibir la corriente de datos de audio desde el par de conferencia respectivo.

- 15 **[0024]** Dicha al menos una porción puede tener un tamaño igual a un mayor divisor común de los tamaños de fragmentos de corriente de datos de audio generados por las funciones de representación de los al menos dos pares adicionales.

- 20 **[0025]** Dicha al menos una porción puede tener un tamaño igual a un mínimo múltiplo común de los tamaños fragmentos de corriente de datos de audio generados por las funciones de representación de los al menos dos pares adicionales.

[0026] El procedimiento también puede comprender:

- realizar un segundo almacenamiento en búfer, para cada uno de los al menos dos pares de conferencia adicionales, de las porciones de al menos una corriente de datos de audio recibida desde el otro par de conferencia;

- 25 y

- a partir de la recepción de una notificación desde un dispositivo de captura del terminal de telecomunicaciones que indica que el dispositivo de captura ha capturado una nueva porción de corriente de datos de audio:

- 30 - mezclar las porciones de corriente de datos almacenadas en dicho segundo almacenamiento en búfer con una porción de corriente de datos de audio capturada por el dispositivo de captura para producir una segunda porción de datos de audio respectiva mezclada a enviar a uno respectivo de dichos al menos dos pares de conferencia adicionales.

[0027] El procedimiento también puede comprender realizar, para cada uno de los al menos dos pares de conferencia adicionales, un tercer almacenamiento en búfer de las segundas porciones de datos de audio mezcladas.

- 35 **[0028]** El procedimiento también puede comprender suministrar las segundas porciones de datos de audio mezcladas almacenadas en búfer a una función de captura respectiva adaptada para transmitir la corriente de datos de audio al par de conferencia respectivo.

[0029] El procedimiento también puede comprender:

- detectar la presencia de dichos al menos dos pares de conferencia adicionales;

- 40 - habilitar automáticamente dicho primer almacenamiento en búfer y mezclar tras la detección de dichos al menos dos pares adicionales;

- automáticamente disabling dicho primer almacenamiento en búfer y mezclar, en caso de que se detecte, la presencia de únicamente un par adicional.

Breve descripción de los dibujos

- 45 **[0030]** Las características y ventajas de la presente invención se comprenderán mejor con la lectura de la siguiente descripción detallada de un ejemplo de realización de la misma, proporcionada meramente a modo de ejemplo no limitativo, que se llevará a cabo haciendo referencia a los dibujos adjuntos, en los cuales:

La **figura 1** muestra esquemáticamente un escenario donde se aplica la presente invención;

La **figura 2** muestra esquemáticamente, en términos de bloques funcionales, la estructura de un mezclador de conferencia audio según una realización de la presente invención;

5 La **figura 3** muestra esquemáticamente los componentes funcionales principales de un terminal de comunicaciones adaptado para alojar el mezclador de conferencia audio de **la figura 2**;

La **figura 4** muestra con mayor detalle la estructura del mezclador de conferencia audio según una realización de la presente invención;

La **figura 5** representa, en términos de un diagrama de flujo esquemático, el funcionamiento del mezclador de conferencia audio de **la figura 4**, según una realización de la presente invención;

10 La **figura 6** es un diagrama de tiempos que ilustra la temporización de un proceso de llenado de un búfer de fragmentos de una corriente de datos de audio del mezclador de conferencia audio de **la figura 2**, según una realización de la presente invención;

y

15 La **figura 7** es un diagrama de tiempos de un caso de ejemplo en el que unos fragmentos de corrientes de datos de audio que llegan al punto final que aloja al mezclador de conferencia audio de dos pares implicados en una conferencia virtual se ponen a disposición del mezclador de conferencia audio a diferentes velocidades.

Descripción detallada de una realización de la invención

20 [0031] Haciendo referencia a los dibujos, en la **figura 1** se muestra esquemáticamente un escenario donde es aplicable la presente invención. Los números de referencia **105a**, **105b** y **105c** denotan tres personas que participan en una conferencia de audio virtual, que explotan los respectivos terminales de comunicación **110a**, **110b** y **110c**, como por ejemplo video-teléfonos, asistentes digitales personales (PDA), teléfonos móviles, ordenadores personales, interconectados a través de una red de telecomunicaciones **115** que puede incluir una red de telefonía fija y/ o móvil, y una red de paquetes de datos tal como Internet.

25 [0032] Las tres personas (pares) **105a**, **105b** y **105c** que participan en la conferencia de audio virtual pueden cada una hablar y ser vista por los otros dos pares, y cada uno de los tres pares puede escuchar y ver a los otros dos pares. Esto se hace posible mediante un mezclador de conferencia, cuya funcionalidad principal es proporcionar experiencia de "conferencia virtual" al usuario.

30 [0033] Se señala que la elección de considerar sólo tres pares de conferencia está meramente dictada por razones de simplicidad en la descripción: la presente invención no está limitada a ello, y es aplicable a cualquier número de pares de conferencia.

[0034] Haciendo referencia a la **Figura 2**, se muestran esquemáticamente los componentes principales de un mezclador de conferencia de audio de acuerdo con una realización de la presente invención.

35 [0035] El mezclador de conferencia, indicado **205**, está alojado en uno de los terminales (puntos finales) de los tres pares que participan en la conferencia de audio virtual, en el ejemplo mostrado en el punto final **110c** del par **105c** (el "par C"). Es suficiente que el punto final de al menos uno de los pares que participan en la conferencia virtual aloje el mezclador de conferencia (en general, el mezclador de conferencia puede ser alojado en el punto final del par que inicia la conferencia virtual), sin embargo, nada impide que un mezclador de conferencia esté también alojado en uno o más de los puntos finales de los otros pares.

40 [0036] El mezclador de conferencia **205** comprende un primer mezclador **220** configurado para recibir corrientes de datos de audio **225a** y **225c**, respectivamente recibidas desde el punto final **110a** de el par **105a** (el "par A") y generadas por el punto final **110c** de el par C, y para generar una corriente de datos de audio mezclada **225ac** a enviar al punto final **110b** de el par **105b** (el "par B"). Las corrientes de datos de audio **225a** y **225c** son generadas por dispositivos de captura tales como micrófonos **215a** y **215c** de los puntos finales **110a** y **110c**;

45 en el punto final **110b**, la corriente de datos de audio mezclada **225ac** es representada por un dispositivo de representación tal como un altavoz **210b**.

50 [0037] El mezclador de conferencia comprende además un segundo mezclador **230**, configurado para recibir la corriente de datos de audio **225c**, generada por el punto final **110c** de el par C, y una corriente de datos de audio **225b** recibida desde el punto final **110b** de el par B, y para generar una corriente de datos de audio mezclada **225bc** a enviar al punto final **110a** de el par A. La corriente de datos de audio **225b** es generada por un micrófono **215b** de los puntos finales **110b**; en el punto final **110a**, la corriente de datos de audio mezclada **225bc** se representa por un altavoz **210a**.

[0038] El mezclador de conferencia comprende además un tercer mezclador **235**, configurado para recibir la corriente de datos de audio **225a**, recibida desde el punto final **110a** de el par A, y la corriente de datos de audio **225b**, recibida desde el punto final **110b** de el par B, y para generar una corriente de datos de audio mezclada **225ab**, se representa por un altavoz **210c** de el punto final **110c**.

5 [0039] En particular, los flujos de datos de audio están en formato digital, y los mezcladores son mezcladores digitales.

[0040] Se señala que el número de mezcladores de flujos de datos del mezclador de conferencia **205**, así como el número de flujos de datos que cada uno de los mezcladores está configurado para recibir y mezclar, en función del número de pares que participan en la conferencia virtual.

10 [0041] También es importante señalar que la operación realizada por el mezclador de conferencia **205** difiere de la operación realizada por un MCU proporcionado en la red central. El mezclador de conferencia **205** integra las operaciones de capturar, representar y mezclar en un solo dispositivo (el dispositivo **110c**), mientras que una MCU en el núcleo de red lleva a cabo todo el funcionamiento relevante de un dispositivo de representación con la finalidad de tener acceso a las muestras de audio a mezclar, los mezcla y los transmite (posiblemente comprimidos) al par apropiado.

15 [0042] Haciendo referencia a la **Figura 3**, se representa esquemáticamente la estructura de hardware del punto final **110c** que aloja al mezclador de conferencia **205**. Esencialmente, es la estructura general de un aparato de procesamiento de datos, con varias unidades que están conectadas en paralelo a un bus de comunicación de datos interna **305**. En detalle, un procesador de datos (microprocesador o microcontrolador) **310** controla el funcionamiento del terminal **110c**; una RAM (memoria de acceso aleatorio) **315** se utiliza directamente como una memoria de trabajo por el procesador de datos **310**, y una ROM (Read Only Memory) **320** almacena el microcódigo (firmware) a ser ejecutado por el procesador de datos **310**. Un subsistema de comunicación **325** incluye dispositivos de hardware para el manejo de al menos el nivel físico de las comunicaciones a través de la red telefónica/ telecomunicaciones **115**; se proporciona un teclado **330** para marcar los números de teléfono;

25 un subsistema de audio/ vídeo **335** gestiona el dispositivo de altavoz/ pantalla **210c** y el micrófono/ cámara de vídeo **215c**.

[0043] Pasando a la **Figura 4**, se muestran con mayor detalle los componentes funcionales del mezclador de conferencia **205**. En particular, la **Figura 4** muestra el contenido parcial de la memoria de trabajo **315** del terminal **110c** durante una conferencia virtual entre los pares A, B y C; de este modo, los bloques funcionales ilustrados en la figura **4** deben entenderse como módulos software/ firmware, o ejemplos de módulos de software/ firmware. Esto sin embargo no debe interpretarse como una limitación de la presente invención, que puede ser implementada totalmente en hardware, o como una combinación de hardware y software/ firmware.

30 [0044] Los bloques **405a** y **405b** representan ejemplos de un módulo de software de captura multimedia, adaptado para realizar las tareas de captura de las corrientes de datos de audio mezcladas **225bc'** y, respectivamente, **225ac'**, codificarlas para obtener corrientes de datos codificados (*es decir* comprimidos) **225bc** y, respectivamente, **225ac**, y transmitir las por la red telefónica/ telecomunicaciones **115** a los puntos finales **110a** y, respectivamente, **110b** de los pares A y B.

35 [0045] En una realización preferida de la presente invención, la operación de mezclado de las corrientes de datos de audio que genera las corrientes de datos de audio mezcladas **225bc** y **225ac** se realiza en el dominio no comprimido (en valores modulados codificados por pulsos), lo que evita el problema de la compatibilidad entre los diferentes algoritmos de compresión (por ejemplo G.723, AMR, G.722) que hayan sido acordados entre los pares diferentes (es decir, los pares A y C podrían haber acordado un algoritmo de compresión diferente de aquel acordado entre los pares B y C). En este caso, los ejemplos de módulos de software multimedia de captura **405a** y **405b** también son responsables de la codificación de las corrientes de datos de audio mezcladas **225bc** y **225ac** según una norma de codificación respectiva predeterminada, que pueden ser diferente para los diferentes pares A y B de la conferencia virtual.

40 [0046] Los bloques **410a** y **410b** representan ejemplos de un módulo de software multimedia de representación adaptado para realizar las tareas (independientes de las tareas de captura realizadas por los bloques **405a** y **405b**) de recepción de las corrientes de datos de audio **225a** y **225b**, respectivamente, transmitidas por los puntos finales **110a** y, respectivamente, **110b** por la red telefónica/ telecomunicaciones **115**, decodificar las corrientes de datos de audio recibidas **225a** y **225b** para obtener corrientes de datos descodificadas **225a'** y **225b'**, y presentarlas a través de el altavoz **210c** de el punto final **110c**.

45 [0047] Las instancias de módulos de software multimedia de captura y representación **405a**, **405b**, **410a** y **410b** son aplicaciones de espacio de usuario, que se ejecutan en el nivel de capa de aplicación en el punto final **110c**. Como ya saben los expertos en la materia, la memoria de trabajo de un procesador de datos (como la RAM **315** del terminal **110c**) se puede dividir idealmente en dos espacios de memoria básicos: un "espacio de usuario" y un "espacio de núcleo". El espacio de usuario es la región de memoria donde residen y se ejecutan las aplicaciones de software/ firmware o ejecutables del usuario. El espacio de núcleo es la región de memoria

donde residen los módulos de software/ firmware de núcleo o ejecutables; módulos de software/ firmware de núcleo son módulos de software/ firmware que forman el núcleo de un sistema operativo que se inicia en el arranque del procesador de datos, y cuya función es proporcionar un entorno en el que se puedan ejecutar otros programas, proporcionar servicios de hardware a estos (como el suministro de la memoria y el acceso al espacio en dispositivos de almacenamiento), programar su ejecución y permitir que coexistan varios procesos.

[0048] Generalmente, con la finalidad de acceder a recursos de captura y ejecución de audio (captura y representación) (el micrófono **215c**, el altavoz, **210c**), las instancias de módulos de software multimedia de captura y representación **405a**, **405b**, **410a** y **410b** explotan módulos de biblioteca dedicados, típicamente módulos de biblioteca de espacio de usuario.

[0049] Como ya saben los expertos en la materia, hay dos enfoques posibles en el acceso y el uso de dispositivos como micrófonos y altavoces: el uso de módulos de biblioteca que se ejecutan en espacio de usuario, o directamente a través de interfaces de programación de aplicaciones (API) de controladores de espacio de núcleo dispositivo de núcleo. Los módulos de biblioteca utilizan APIs de controlador de dispositivo de uso para controlar el dispositivo de interés. Los controladores de dispositivos de espacio de núcleo se utilizan cuando el sistema operativo impide que las aplicaciones de espacio de usuario accedan directamente a los recursos de hardware; normalmente esto está relacionado con la presencia de dominios de protección jerárquicos (o anillos de protección) en el sistema operativo, que actúan como un método de protección contra fallos generados por la aplicación. Los controladores de dispositivo se pueden ejecutar en el espacio de usuario (y por lo tanto actúan como módulos de biblioteca de espacio de usuario) cuando el sistema operativo no implementa conceptos de anillos de protección.

[0050] En una realización de la presente invención, el mezclador de conferencia **205** está implementado como un proceso de espacio de usuario o hilo, preferiblemente de alta prioridad, que se ejecuta en el punto final **110c**;

En este caso, el mezclador de conferencia **205** puede sustituir a los módulos de biblioteca utilizados por las instancias **405a**, **405b**, **410a** y **410b** de los módulos multimedia de captura y representación de espacio de usuario (capa de aplicación) para acceder a los recursos de captura y representación de audio **210c** y **215c**.

[0051] Como alternativa, el mezclador de conferencia **205** puede ser implementado como un controlador de dispositivo de espacio de núcleo. En este caso, el mezclador de conferencia **205** sustituye al controlador de dispositivo de núcleo normalmente responsable de gestionar las operaciones captura y representación de audio.

[0052] La implementación de la funcionalidad de mezclador de conferencia como controlador de dispositivo de espacio de núcleo permite explotar mejor el beneficio de baja latencia de tratar datos dentro de una rutina de servicio de interrupción, y evitar cargar el sistema con *threads*/ procesos en espacio de usuario de alta prioridad.

[0053] En la figura 4, la línea **415** indica una API expuesta por el mezclador de conferencia **205**, a través de la cual puede ser accedida por las instancias de módulos de software multimedia de captura y representación **405a**, **405b**, **410a** y **410b**. La API **415** replica las mismas funcionalidades suministradas por los módulos de biblioteca utilizadas para acceder a los recursos de entrada/ salida **210c** y **215c**. El comportamiento de las instancias de módulos multimedia de representación y captura **405a**, **405b**, **410a** y **410b** no necesitan ser cambiados para permitir que interactúen con el mezclador de conferencia **205** (cuya presencia es transparente a las instancias de módulos multimedia de representación y captura **405a**, **405b**, **410a** y **410b**).

[0054] El mezclador de conferencia **205** comprende (en el ejemplo aquí considerado de conferencia virtual que implica tres pares) tres mezcladores de fragmentos de corrientes de datos de audio **420**, **425** y **430**, adaptados para mezclar porciones (o fragmentos) de las corrientes de datos de audio. Dos búfers de datos **420-1** y **420-2**, **425-1** y **425-2**, y **430-1** y **430-2** están operativamente asociados con cada uno de los mezcladores **420**, **425** y **430**. Los búfers de datos **420-1** y **425-1** son los búfers de contribución de datos proporcionados con respecto a las corrientes de datos de audio que vienen de los pares A y, respectivamente, B, con propósitos de captura. Los búfers de datos **420-2** y **425-2** son los búfers recipientes de datos para las corrientes de datos de audio mezcladas a enviar a los pares B y, respectivamente, A. Los búfers de datos **430-1** y **430-2** son los búfers de contribución de mezclador para las corrientes de datos de audio recibidas desde los pares A y B, respectivamente, para propósitos de representación.

[0055] En particular, los mezcladores **420**, **425** y **430** son mezcladores digitales.

[0056] En caso de que más de tres pares estén implicados en la conferencia virtual, aumenta el número de mezcladores y de búfers asociados, en particular, para cada par adicional que participa en la conferencia virtual, hay que añadir un mezclador tal como el mezclador **420**, con un par de búfers asociados tales como los búfers **420-1** y **420-2**; también, hay que añadir un búfer tal como el búfer **430-1** o **430-2** para cada par adicional.

[0057] El número de referencia **435** indica un proceso de representación de audio, para enviar fragmentos de corrientes de datos de audio listos para ser representados por el altavoz **210c**, para representarlos. El número de referencia **440** indica un proceso de captura de audio, para recibir datos de audio capturados por el micrófono **215c**.

- 5 [0058] En caso de que el mezclador de conferencia **205** esté implementado como un proceso de espacio de usuario, las operaciones de mezclado realizadas por los mezcladores **420, 425 y 430** pueden ser implementadas como threads/ procesos de alta prioridad. Por procesos de alta prioridad se entienden threads/ procesos que se ejecutan con una prioridad mayor que las tareas normales. Para los propósitos de la presente invención, alta prioridad significa la mayor prioridad posible (cercana a la prioridad en tiempo real), que no ponga en peligro la estabilidad del sistema. Como alternativa, si el mezclador de conferencia **205** es un controlador de dispositivo de espacio de núcleo, las operaciones de mezclado realizadas por los mezcladores **420, 425 y 430** pueden ser implementadas como rutina de interrupción de servicio, que se inician tan pronto como se recibe una interrupción de la dispositivos de representación y captura.
- 10 [0059] En particular, las operaciones de mezclado se realizan con la velocidad de entrada/ salida (I/O) (es decir, a la velocidad a la que los datos son capturados por los dispositivos de captura **215c**, y a la velocidad a la que los datos a representar son consumidos por los dispositivos de representación **210c**).
- [0060] En detalle, cada vez que un nuevo fragmento de datos de audio está disponible en la interfaz de entrada del micrófono **215c**, se lleva a cabo una operación de entrada de mezcla por los mezcladores **420 y 425**:
- 15 se toma el siguiente fragmento disponible de datos presentes en el búfer de contribución de datos **420-1 y 425-1**, respectivamente, y se mezcla con el nuevo bloque de datos de audio recién capturado por el micrófono **215c**.
- [0061] De manera similar, cada vez que se solicita un nuevo fragmento de datos de audio por la interfaz de salida de altavoz **210c**, se lleva a cabo una operación de mezcla de salida por el mezclador **430**, usando el siguiente fragmento disponible de datos de audio presentes en los búfers de contribución de mezclador **430-1 y 430-2**.
- 20 [0062] Esto garantiza el mínimo retardo de extremo a extremo transcurrido entre la disponibilidad de datos (desde el micrófono **215c**, o desde las instancias de módulos de software de representación multimedia **410a y 410b**) y la producción de datos (para las instancias de módulos de software de captura multimedia **405a y 405b** y el altavoz **205c**).
- 25 [0063] El funcionamiento del mezclador de conferencia **205** se explica en detalle a continuación, haciendo referencia al diagrama de flujo de la figura **5**. En la siguiente explicación, se asume que los dispositivos de representación **210c y** los dispositivos de captura **215c** operan sobre una misma base de tiempo, es decir, con un mismo reloj, de modo que (sus interfaces I/ O) generan interrupciones simultáneas (es decir, son síncronos); sin embargo, sin embargo, esto no debe interpretarse como una limitación de la presente invención, que se aplica también en el caso de los dispositivos de representación **210c** y los dispositivos de captura **215c** operan con relojes diferentes pero nominalmente iguales (los diferentes relojes pueden desplazarse), y también en el caso de los dos relojes no sean aún nominalmente iguales. En estos casos, las interrupciones generadas por (interfaces de entrada de) los dispositivos de captura y representación no son, por regla general, simultáneas.
- 30 [0064] Cuando llega el tiempo de representación de un nuevo fragmentos de datos de audio (bloque **505**, rama de salida Y), se toman los fragmentos de datos más antiguos presentes en los búfers de contribución de mezclador **430-1 y 430-2** (bloque **510**), se mezclan juntos en el *fl*y (bloque **515**) y los datos mezclados se suministran al altavoz **210c** para su representación (bloque **520**). La llegada del tiempo de representación del nuevo fragmento de datos de audio es un acontecimiento **507** que puede ser señalado por una notificación, como una interrupción del (interfaz de entrada del) altavoz **210c**, si el mezclador de conferencia **205** está implementado como un controlador de dispositivo de espacio de núcleo, o dicha notificación puede ser asíncrona desde el controlador de altavoz, en caso de que el mezclador de conferencia **205** esté implementado como un proceso de espacio de usuario.
- 35 [0065] Bajo la suposición anterior de que la dispositivos de representación y captura operan basados en el mismo reloj, la llegada del tiempo de representación de un nuevo fragmento de datos coincide con la llegada del tiempo de captura de un nuevo fragmento de datos de audio; sin embargo, en general, cuando el tiempo de captura de un nuevo fragmento de datos de audio llega (también en este caso, este acontecimiento puede ser una interrupción, si el mezclador de conferencia se implementa como un controlador de dispositivo de espacio de núcleo, o puede ser una notificación asíncrona desde el controlador de micrófono), se toma el fragmento recién capturado de datos de audio (bloque **525**), se toman los fragmentos de datos más antiguos presentes en los búfers de contribución de datos **420-1 y**, respectivamente, **425-1** (bloque **530**), son mezclados, por los mezcladores **420 y 425**, con el fragmento recién capturado de datos de audio (bloque **535**), para producir un nuevo fragmento de audio mezclado a enviar a los pares, y los datos mezclados se meten en los búfers de recipiente de datos **420-1 y 425-2**, respectivamente (bloque **540**). Las instancias de módulos de software de captura multimedia **405a y 405b** capturan entonces los fragmentos de datos mezclados de audio más antiguos del respectivo búfer de recipiente de datos **425-2 y 420-2**.
- 40 [0066] La llegada del tiempo de representación del nuevo fragmento de datos de audio también se utiliza para desencadenar la carga de los búfers de contribución de mezclador **430-1 y 430-2** con una nueva trama de audio (es decir, una parte, un fragmento de la corriente de datos de audio) puesta a disposición por las instancias de
- 45
- 50
- 55

módulos de software multimedia de representación **410a y 410b** (bloque **545**), que accede a el mezclador de conferencia **205** a través de la API de representación que expone; en esta descripción, el término desencadenar significa indicar a las instancias de módulos de software multimedia de representación **410a y 410b** que hay espacio libre disponible en el búfer asociado **430-1 y 430-2**. Las instancias de módulos de software multimedia de representación pueden entonces (cuando los datos están disponibles) escribir una nueva trama de audio en los búfers **430-1 y 430-2**; la misma trama de audio también se copia en el respectivo búfer de contribución de datos **420-1 ó 425-1**.

[0067] En cuanto al tamaño de los búfers **420-1 y 420-2, 425-1 y 425-2, y 430-1 y 430-2**, hay un compromiso. Con el fin de mantener al mínimo el retardo de extremo a extremo, los búfers deben ser lo más pequeño posible. Sin embargo, esto contrasta con el requisito de tener búfers tan grandes como sea posible, con la finalidad de evitar la introducción de fallos en los flujos de audio, debido a desbordamiento en uno de los búfers de contribución de datos **420-1 y 425-1** (por ejemplo, esto sucede cuando el búfer de contribución de datos **420-1** está vacío o cuando un nuevo fragmento capturado datos de audio llega desde el micrófono **215c**).

[0068] En cuanto al tamaño de los fragmentos de datos de audio que se almacenan en los búfers, los parámetros significativos para su determinación son el número de bits por muestra de audio (por ejemplo, 16 bits), la tasa de muestreo (por ejemplo, 8kHz para algoritmos de compresión de banda estrecha como G.723, o 16kHz para algoritmos de compresión de banda ancha como G.722.2), la duración de la trama de audio, es decir la cantidad mínima de datos manejados por el algoritmo de compresión elegido (por ejemplo, 10ms para G. 729, 30ms para G.723, 20ms para G722). Por lo tanto, el tamaño de fragmentos de datos puede ser:

tamaño de fragmentos de datos = tasa de muestreo * bit por muestra/8* duración de la trama de audio

[0069] Cuando los puntos finales **110a y 110b** de los pares A y B utilizan diferentes algoritmos de compresión de datos, el tamaño de las tramas de audio procedentes de los dos pares (y puestas a disposición por las instancias de módulos de software multimedia de representación) puede ser diferente. Esto introduce casos adicionales de subflujo durante la operación de mezclado: Esto introduce casos adicionales de subflujo durante la operación de mezcla: si la cantidad de datos disponibles de todos las diferentes contribuidores de la operación de mezclado no es suficiente para producir un fragmento de datos completo en la salida, se produce un problema técnico.

[0070] De acuerdo con una realización de la presente invención, el mezclador de conferencia **205** puede calcular un tamaño único de fragmento de datos, que se utiliza para todos los búfers, utilizando el mínimo común múltiplo de todos los tamaños de trama de audio utilizados por los diferentes pares de conferencias. Por ejemplo, suponiendo que el par A transmite (y recibe) tramas de audio de 30 ms, y el par B transmite (y recibe) tramas de audio de 20 ms, el tamaño del fragmento de datos de audio utilizado por el mezclador de conferencia (es decir, el "cuanto" de datos de audio procesado por el mezclador de conferencia **205**) puede ser de 60 ms;

Esto significa que cada 60 ms, llega una interrupción desde el altavoz **210c**, y se genera sobre la marcha un nuevo fragmento de datos de 60 ms por el mezclador 430 y se envía al altavoz 210c (de forma similar, bajo el supuesto de que solo existe una base de tiempo, está listo cada 60 ms un nuevo fragmento de datos de audio capturados por el micrófono 215c).

[0071] El número de fragmentos de datos que el búfer del mezclador de conferencia genérico **205** debe almacenar se diseña teniendo en cuenta que el tiempo de latencia es inferior a la duración o el tiempo de latencia de cambio de tarea proporcionado por el equipo/ sistema operativo que aloja a los módulos multimedia. Bajo el supuesto de que el tiempo de latencia es inferior a la duración (es decir, el tamaño) de un fragmento de datos, el número de fragmentos de datos puede mantenerse en el mínimo (cada búfer tiene dos registros, cada uno del tamaño de un fragmento de datos, que se utilizan en un modo de acceso "ping-pong", de forma alternativa para la lectura y escritura). En otras palabras, cuando el sistema es capaz de procesar un solo fragmento de datos dentro de su duración, se considera un enfoque en el que cada búfer puede almacenar dos fragmentos de datos como un método preferido para la manipulación de datos con el fin de minimizar el retardo de extremo a extremo: mientras que el procesamiento de un nuevo fragmento de datos está en curso, el sistema puede alimentar el fragmento de datos anterior al (desde) dispositivo de representación (captura). Este concepto se describe mejor en la descripción correspondiente a la **Figura 6**.

[0072] El retardo mínimo de extremo a extremo introducido por el mezclador de conferencia **205** es igual a la duración de un fragmento de datos de audio entre los pares A y C y entre los pares B y C (tal como se describe en la figura 6). El retraso adicional de extremo a extremo introducido por el mezclador es igual al tiempo de espera que un fragmento de datos tiene que esperar en el búfer **430-1 ó 430-2** antes de la representación. Con el fin de evitar el subflujo en uno de los búfer antes mencionados, se necesita un tiempo mínimo de espera de un tiempo de fragmento de datos. No se introduce retardo adicional de extremo a extremo en su lugar en la ruta de captura de par C a los pares A o B. Los búfers que introducen retardo de extremo a extremo son los búfers **420-1, 425-1, 430-1 y 430-2**.

[0073] Haciendo referencia a la **Figura 6**, se representa esquemáticamente el retardo introducido por los búfers de contribución de mezclador **430-1 y 430-2**. El nivel mínimo de almacenamiento posible, sin introducir artefactos

durante la operación de mezclado debido a subflujo en uno de los búfers de contribución de mezclador **430-1 y 430-2**, es igual a un fragmento de datos, por lo que el retardo de extremo a extremo es igual a un fragmento de datos. En el dibujo, **INT(n-1)**, **INT(n)**, **INT(n+1)**, **INT(n+2)**, **INT(n+3)** denotan cinco interrupciones consecutivas desde el dispositivo de representación (el altavoz 210c), que se producen en los instantes **Tn-1**, **Tn**, **Tn+1**, **Tn+2**, **Tn+3**. La interrupción **INT(n-1)** inicia la representación de los fragmentos de datos de audio (n-1)ésimo, La interrupción **INT(n)** inicia la representación de los fragmentos de datos de audio énesimos, y así sucesivamente.

[0074] Suponiendo, por motivos de simplicidad que el tamaño de fragmentos de datos es igual al tamaño de la trama de audio, o el genérico de las instancias de módulos de software multimedia de representación **410a o 410b** inicia la escritura (acontecimiento **Ws(n)**) la trama de audio n-ésimo al búfer de contribución de mezclador respectivo **430-1 y 430-2** en el instante **Tn-1**, cuando el proceso de representación de audio **435** recibe la interrupción **INT(n-1)** para empezar a representar el fragmento de datos (n-1)ésimo, finaliza la escritura de la n-ésima trama de audio en la búfer de contribución de mezclador (acontecimiento **We(n)**) antes de la llegada de la siguiente interrupción **INT(n)**, por lo tanto cuando esta interrupción siguiente llega el nuevo fragmento de datos de audio está listo para ser representado; cuando, en el instante **Tn**, el proceso de representación de audio **435** recibe la siguiente interrupción **INT(n)** para empezar a representar el fragmento de datos (n)ésimo, la instancia de módulo de software multimedia de representación **410a o 410b** inicia la escritura (acontecimiento **Ws(n+1)**) de la trama de audio (n+1)ésima en los respectivos búfers de contribución de mezclador **430-1 y 430-2**;

la escritura de la trama de audio (n+1)ésima en el búfer de contribución de mezclador finaliza(acontecimiento **We(n)**) antes de la llegada de la siguiente interrupción **INT(n)**, de modo que el nuevo fragmento de datos está preparado para ser representado cuando llega la siguiente interrupción **INT(n+1)**, y así sucesivamente.

[0075] Gracias a que la operación de mezclado se realiza en el *fly* en el proceso de representación de audio **435**, empezar la recepción de la interrupción (o de la notificación asíncrona) de los dispositivos de representación, se evitan almacenamientos adicionales en búfer y retardos.

[0076] Como se ha mencionado en lo que antecede, una posibilidad cuando las tramas de audio procedentes de/ destinadas a ser enviadas a los diferentes pares son de diferente tamaño, es trabajar con fragmentos de datos de tamaño igual al menor múltiplo común de los tamaños de las diferentes tramas de audio. Haciendo referencia a la **Figura 7**, en un diagrama de tiempos similar al de la figura **6** se representa un caso en el que los dos pares A y B adoptan diferentes tramas de audio; esto se traduce en diferentes períodos **TA**, **TB** en los que las instancias de módulos de software multimedia de representación **410a o 410b** hacen que estén disponibles los marcos de audio. Suponiendo que, en el instante **Tn-1** en el que la instancia de módulo de software multimedia de representación **410a** comienza a hacer disponible la trama enésima de la corriente de audio que viene del par A, la instancia de módulo de software multimedia de representación **410b** también empieza a hacer disponible la trama enésima de la corriente de audio que viene del par B. En el instante **Tn** la instancia de módulo de software multimedia de representación **410a** completa la trama de audio enésima, y empieza a hacer disponible la trama de audio (n+1); en su lugar la instancia de módulo de software multimedia de representación **410b** completa la trama de audio enésima más tarde, en el instante **Tn+a**. De manera similar, la instancia de módulo de software multimedia de representación **410a** completa la trama de audio (n+1)ésima en el instante **Tn+1**, mientras que la instancia de módulo de software multimedia de representación **410b** completa la trama de audio (n+1)ésimo más tarde, en el instante **Tn+1+a**. La instancia de módulo de software multimedia de representación **410a** completa la trama de audio (n+2)ésima en el instante **Tn+2**, mientras que la instancia de módulo de software multimedia de representación **410b** completa la trama de audio (n+2)ésimo más tarde, en el instante **Tn+3**, cuando la instancia de módulo de software multimedia de representación **410a** completa la trama de audio (n+3)ésima: en este instante, las dos instancias de módulos de software multimedia de representación **410a y 410b** se vuelven a sincronizar. Con la finalidad de impedir el subflujo en el instante **Tn+a**, los búfers de contribución de mezclador **430-1 y 430-2** se pueden diseñar de manera que contengan fragmentos de datos de igual tamaño a cuatro tramas de audio a partir del par A, y tres trama de audio del par B.

[0077] Preferentemente, con el fin de mantener el retardo de extremo a extremo tan bajo como sea posible, el tamaño de los fragmentos de datos utilizados por el mezclador de conferencia **205** puede ser igual a un mayor divisor común, en lugar del mínimo múltiplo común, de las tramas de audio adoptadas por los diferentes pares. Por ejemplo, suponiendo de nuevo que las tramas de audio desde/ al par A son de 30 ms, y aquellas desde/ al par B son de 20 ms, el mezclador de conferencia **205** puede ser diseñado para trabajar con datos fragmentos de 10 ms; Esto significa que cada 10 ms, llega una interrupción desde el altavoz **210c**, y se genera un nuevo fragmento de datos de 10 ms (mezclando fragmentos de datos de 10 ms de las tramas de audio almacenadas en los búfers de contribución de mezclador **430-1 y 430-2**) en el *fly* por el mezclador **430** y enviada al altavoz **210c**; cada dos interrupciones, la instancia de módulo de software multimedia de representación **410b** se notifica que en el búfer **430-2** hay espacio para una nueva trama de audio, mientras que esto sucede cada tres interrupciones para la instancia de módulo de software multimedia de representación **410a**.

[0078] Esto permite realizar la operación de mezclado con un número entero de "cuantos" de fragmento de datos en cada búfer (por el contrario, si el tamaño de fragmentos de datos utilizado en la mezcla es diferente para los diferentes pares, tendrían que almacenarse más datos con el fin de evitar artefactos debidos al subflujo de búfer).

5 [0079] Una ventaja de la aplicación del mezclador de conferencia **205** es que expone las mismas APIs a las instancias de módulos de software multimedia de captura y representación **405a**, **405b**, **410a** y **410b** mientras que las bibliotecas convencionales de representación/ captura o controlador de dispositivo de núcleos las hacen adecuadas para ser insertadas de "manera transparente" en una arquitectura de equipo que ya opera, sin afectar a los módulos implementación de representación y captura multimedia.

[0080] En una realización preferida de la presente invención, el uso del mezclador de conferencia **205** se puede seleccionar por el usuario sobre una base por-conexión: durante una comunicación que comprende sólo dos pares, como por ejemplo pares A y C, o pares B y C, no se necesita mezclado, mientras que la mezcla se hace necesaria durante una sesión multi-pares (tres o más).

10 [0081] Restringir el uso del mezclador de conferencia **205** para aquellos casos en los que la mezcla es realmente necesaria permite reducir los efectos secundarios, tal como el aumento el retardo de extremo a extremo durante una conversación entre sólo dos pares.

[0082] En particular, el mezclador de conferencia **205** puede ser adaptado para "auto habilitarse" cuando un tercer par de una conferencia virtual comienza a solicitar fragmentos de datos de audio.

15 [0083] En principio, siempre y cuando sólo dos pares estén involucrados en una comunicación, el mezclador de conferencia no podría ser utilizado, evitando así la introducción de retardo adicional de extremo a extremo, y cuando un tercer par entra en la sesión de comunicación, el mezclador podría ser insertado. Sin embargo, está inserción en tiempo real del mezclador de conferencia es una operación que consume tiempo que podría provocar problemas técnicos largos. En una realización preferida de la invención, con el fin de evitar tanto la adición de retraso de extremo a extremo cuando la operación de mezclado no es necesaria, y al mismo tiempo evitar los problemas técnicos causados por la demora de inserción del mezclador cuando el tercer par entra en la conferencia, ya se utiliza el mezclador de conferencia **205** desde el comienzo de la comunicación entre los dos primeros pares, y se reduce el almacenamiento en búfer se reduce al mínimo requerido por el número de pares que participan activamente en la conferencia. Por "activamente implicado" se entiende haber "abierto los " dispositivos de representación/ captura. La implicación del mezclador no se basa en el flujo de datos (a o desde sus interfaces de datos) sino en la intención explícita de un módulo de software multimedia (representación o captura) para iniciar una nueva sesión; el flujo de datos, de hecho, puede ser discontinuo, ya que está relacionada con la disponibilidad de datos desde el par remoto. En particular, cuando sólo dos pares, por ejemplo pares A y C están activos, no son necesarias las operaciones de mezclado y el mezclador de conferencia **205** impide el almacenamiento en búfer de fragmentos de datos de audio en el búfer de contribución de datos **420-1** y en el búfer de contribución de mezclador **430-1**. Los fragmentos de datos se escriben directamente por la instancia de módulo de software multimedia de representación **410a** en la salida dispositivo **210c**, sin almacenar en búfer. De esta manera, no se introduce un retardo adicional. En esta condición, el *thread* **435** que, cuando el mezclador está habilitado, trabajando sincrónicamente con las interrupciones recibidas desde el dispositivo de representación de salida **210c**, toma los datos fragmentos de los búfers de contribución de mezclador **430-1** y **430-2**, cuando el mezclador se deshabilita no realiza mezcla alguna, sino que solamente es responsable de señalar a la instancia de módulo de software multimedia de representación **410a** que el dispositivo está preparado para la representación. Se realiza un comportamiento similar por el *thread* de captura. Cuando un nuevo par, como el par B, se vuelve activo, se vuelven a habilitar la operación de mezclado, y las operaciones de almacenamiento en búfer.

[0084] Los módulos de software de representación y captura multimedia no tienen porque cambiar el comportamiento cuando se habilita el mezclador de conferencia. Esto reduce la complejidad del software en sí mismo, especialmente en una parte crucial como la gestión de baja latencia de audio.

45 [0085] La presente invención ha sido aquí descrita teniendo en cuenta algunos ejemplos de realización. Los expertos en la técnica apreciarán fácilmente que varias modificaciones a las realizaciones descritas son posibles, así como diferentes formas de realización, sin apartarse del ámbito de protección definido en las reivindicaciones adjuntas.

50 [0086] En particular, aunque en la anterior descripción se ha hecho referencia a una conferencia de audio, la presente invención también se puede aplicar a las conferencias virtuales de audio/ vídeo, mediante la adición de los búfers y mezcladores para el componente de vídeo de una manera análoga a la descrita más arriba.

REIVINDICACIONES

1. Terminal de telecomunicaciones (**110c**) que aloja un mezclador de conferencia (**205**) adaptado para permitir una conferencia comprende al menos audio entre un primer par de conferencia (**105c**) y al menos dos pares de conferencia adicionales (**105a,105b**), comprendiendo el mezclador de conferencia:
- 5 - para cada uno de los al menos dos pares de conferencia adicionales, un primer búfer de datos respectivo (**430-1,430-2**) configurado para almacenar en búfer porciones de al menos una corriente de datos de audio recibida desde el par de conferencia respectivo;
- un primer mezclador de porciones de datos de audio (**430**) suministrados por los primeros búfers de datos y configurado para:
- 10 a) obtener porciones de corriente de datos de audio almacenadas en los primeros búfers de datos;
- b) mezclar las porciones de corriente de datos de audio obtenidas a partir de los primeros búfers de datos para producir una primera porción de datos de audio mezclada; y
- c) suministrar la primera porción de datos de audio mezclada a un dispositivo de representación (**210c**) del terminal de telecomunicaciones, **caracterizado por el hecho de que** dicho primer mezclador de porciones de datos de audio está configurado para realizar las operaciones a), b) y c) a partir de la recepción de una notificación (**507**) desde dicho dispositivo de representación que indica que el dispositivo de representación está preparado para representar una nueva porción de datos de audio mezclada.
- 15 2. El terminal de telecomunicaciones según la reivindicación 1, en el que dicho dispositivo de representación incluye un altavoz.
- 20 3. El terminal de telecomunicaciones según la reivindicación 1 ó 2, en el que cada uno de dichos primeros búfers de datos comprende al menos dos zonas de almacenamiento cada uno adaptado para almacenar al menos un fragmento de la corriente de datos de audio recibida desde el par respectivo.
- 25 4. El terminal de telecomunicaciones según la reivindicación 3, en el que cada uno de dichos primeros búfers de datos está configurado para ser suministrado por una función de representación respectiva (**410a,410b**), estando dicha función de representación adaptada para recibir la corriente de datos de audio desde el par de conferencia respectivo, y para proporcionar al primer búfer de datos respectivo fragmentos de tamaño predeterminado de las corrientes de datos de audio.
- 30 5. El terminal de telecomunicaciones según la reivindicación 4, en el que cada una de dichas zonas de almacenamiento tiene un tamaño igual a un mayor divisor común de los tamaños de los fragmentos de las corrientes de datos de audio generados por las funciones de representación de los al menos dos pares adicionales.
- 35 6. El terminal de telecomunicaciones según la reivindicación 4, en el que cada una de dichas zonas de almacenamiento tiene un tamaño igual a un mínimo múltiplo común de los tamaños de los fragmentos de las corrientes de datos de audio generados por las funciones de representación de los al menos dos pares adicionales.
7. El terminal de telecomunicaciones según cualquiera de las reivindicaciones anteriores, que comprende además:
- para cada uno de los al menos dos pares de conferencia adicionales, un segundo búfer de datos respectivo (**420-1,425-1**) configurado para almacenar en búfer las porciones de al menos una corriente de datos de audio recibida desde el otro par de conferencia;
- 40 - un segundo mezclador de porciones de corriente de datos de audio (**430**) suministrados por los segundos búfers de datos y configurado para:
- d) obtener porciones de corriente de datos de audio almacenadas en los segundos búfers de datos;
- 45 e) obtener porciones de corriente de datos de audio desde un dispositivo de captura (**215c**) del terminal de telecomunicaciones;
- y
- f) mezclar las porciones de corriente de datos de audio obtenidas a partir de los segundos búfers de datos y las porciones de corriente de datos de audio obtenidas a partir del dispositivo de captura para producir una segunda porción de datos de audio respectiva mezclada a enviar a uno respectivo de dichos al menos dos pares de conferencia adicionales, en el que dicho segundo mezclador de porciones de corriente de datos de audio está configurado para realizar las operaciones d), e) y f) a partir de la recepción de una notificación (**507**) desde dicho
- 50

dispositivo de captura que indica que el dispositivo de captura ha capturado una nueva porción de corriente de datos de audio.

8. El terminal de telecomunicaciones según la reivindicación 7, que comprende además:

- 5 - para cada uno de los al menos dos pares de conferencia adicionales, un tercer búfer de datos respectivo (**420-2,425-2**) configurado para almacenar en búfer las segundas porciones de datos de audio mezcladas producidas por el respectivo segundo mezclador de porciones de corriente de datos de audio.

9. El terminal de telecomunicaciones según la reivindicación 8, en el que cada uno de dichos terceros búfers de datos está configurado para suministrar una función de captura respectiva (**415a,415b**), estando dicha función de captura adaptada para transmitir la corriente de datos de audio al par de conferencia respectivo.

- 10 10. El terminal de telecomunicaciones según la reivindicación 4, que comprende además una primera interfaz de programa de aplicación para permitir el acceso a esta por dichas funciones de representación.

11. El terminal de telecomunicaciones según la reivindicación 9, que comprende además una segunda interfaz de programa de aplicación para permitir el acceso a esta por dicha función de captura.

- 15 12. El terminal de telecomunicaciones según la reivindicación 10 o la 11, en el que dicho mezclador de conferencia está configurado para ser implementado como un proceso de espacio de usuario que se ejecuta en un espacio de usuario del terminal de telecomunicaciones.

13. El terminal de telecomunicaciones según la reivindicación 10 o la 11, en el que dicho mezclador de conferencia está configurado para ser implementado como un proceso en espacio de núcleo que se ejecuta en un espacio de núcleo del terminal de telecomunicaciones.

- 20 14. El terminal de telecomunicaciones según cualquiera de las reivindicaciones anteriores, en el que dicho mezclador de conferencia también está configurado para:

- detectar la presencia de dichos al menos dos pares de conferencia adicionales;

- habilitar automáticamente el primer mezclador de porciones de datos de audio tras la detección de dichos al menos dos pares adicionales;

- 25 - deshabilitar automáticamente dicho primer mezclador de porciones de datos de audio en caso de que se detecte la presencia de únicamente un par adicional.

15. El terminal de telecomunicaciones según la reivindicación 14, en el que dicho mezclador de conferencia también está configurado para habilitar automáticamente los segundos mezcladores de porciones de corriente de datos de audio tras la detección de la presencia de dichos al menos dos pares adicionales, y para deshabilitar automáticamente los segundos mezcladores de porciones de corriente de datos de audio en caso de que se detecte la presencia de únicamente un par adicional.

- 30 16. Procedimiento para realizar una conferencia que comprende al menos audio entre un primer par de conferencia (105c) y al menos dos pares de conferencia adicionales (**105a,105b**), comprendiendo el procedimiento:

- 35 - en un terminal de telecomunicaciones (110c) del primer par de conferencia, realizar un primer almacenamiento en búfer de porciones de al menos una corriente de datos de audio recibida desde cada uno de los al menos dos pares de conferencia adicionales;

y caracterizado por el hecho de que:

- 40 - a partir de la recepción de una notificación (**507**) desde un dispositivo de representación (**210c**) del terminal de telecomunicaciones, que indica que el dispositivo de representación está preparado para representar una nueva porción de datos de audio mezclada:

- mezclar las porciones de corriente de datos de audio almacenadas en dicho primer almacenamiento en búfer, para producir la porción de datos de audio mezclada; y

- suministrar la porción de datos de audio mezclada al dispositivo de representación.

45

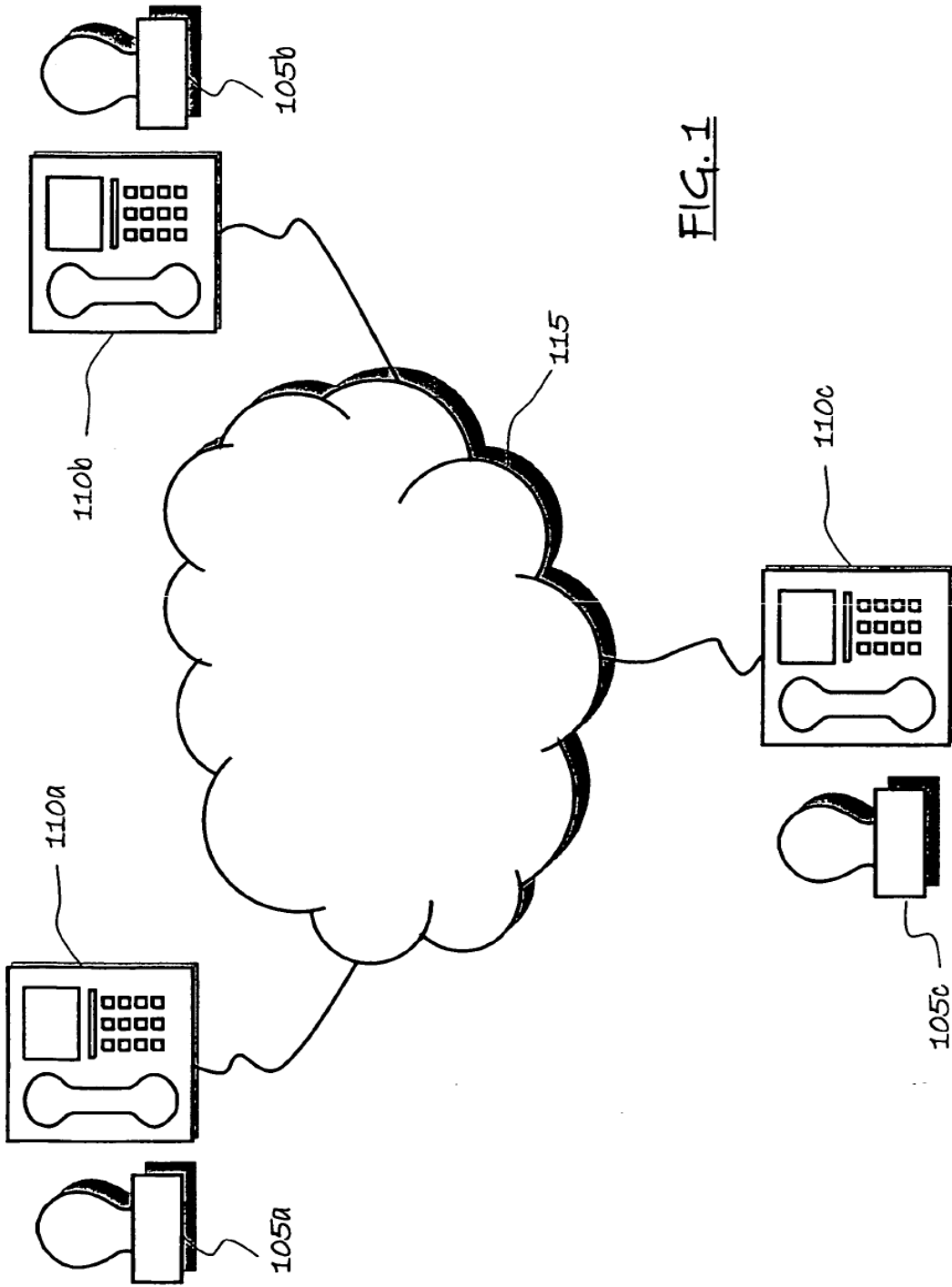


FIG. 1

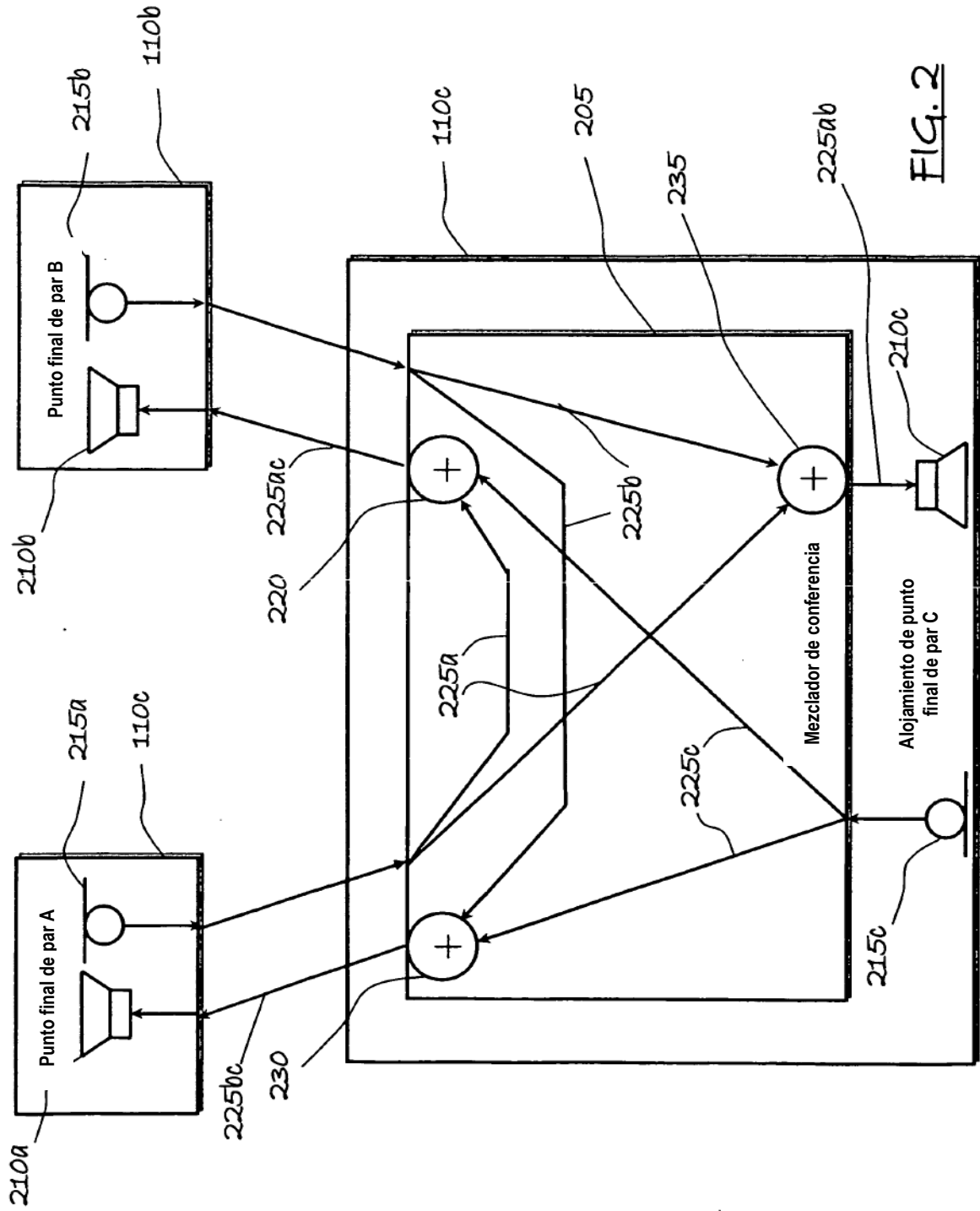


FIG. 2

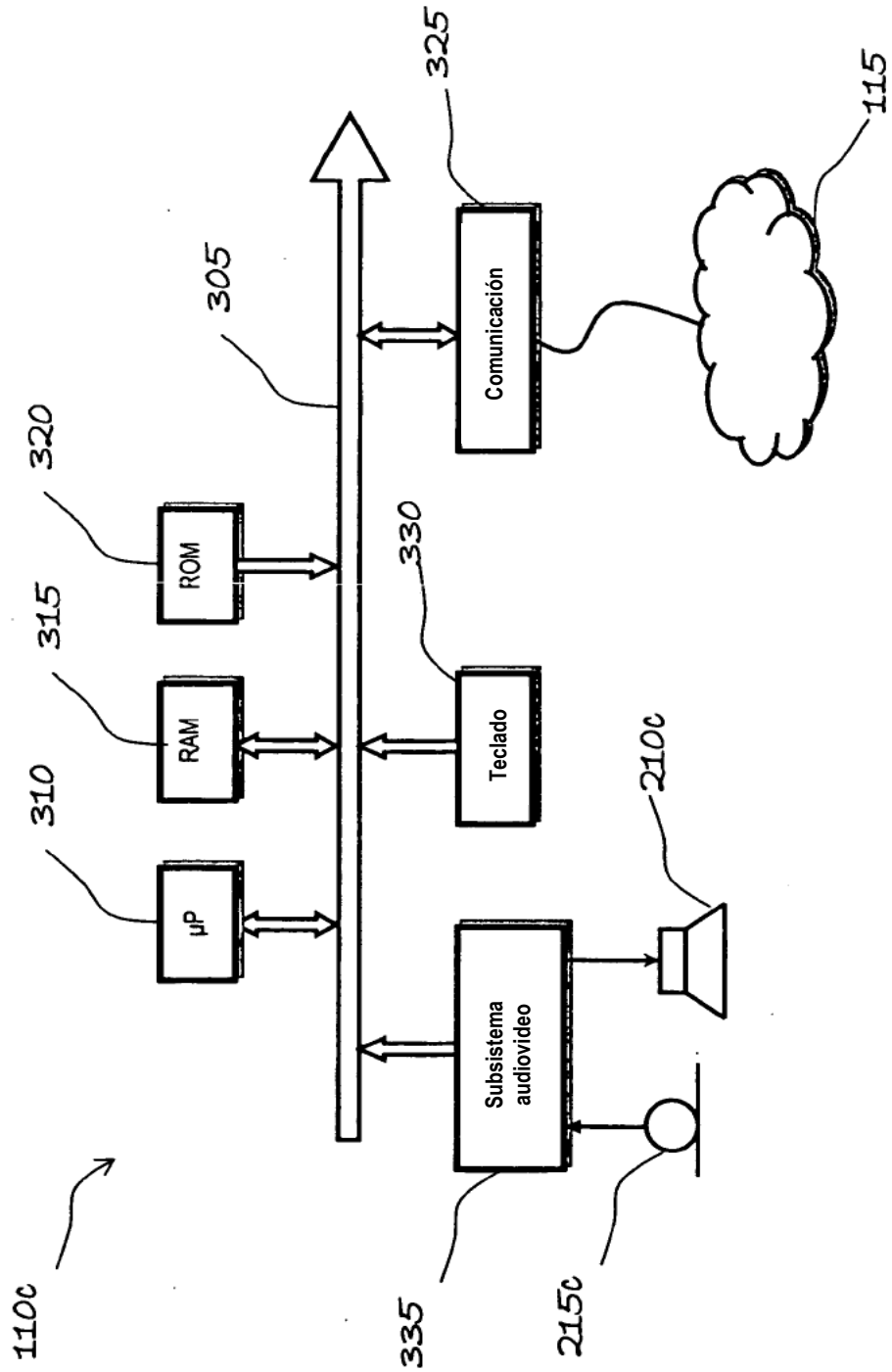


FIG. 3

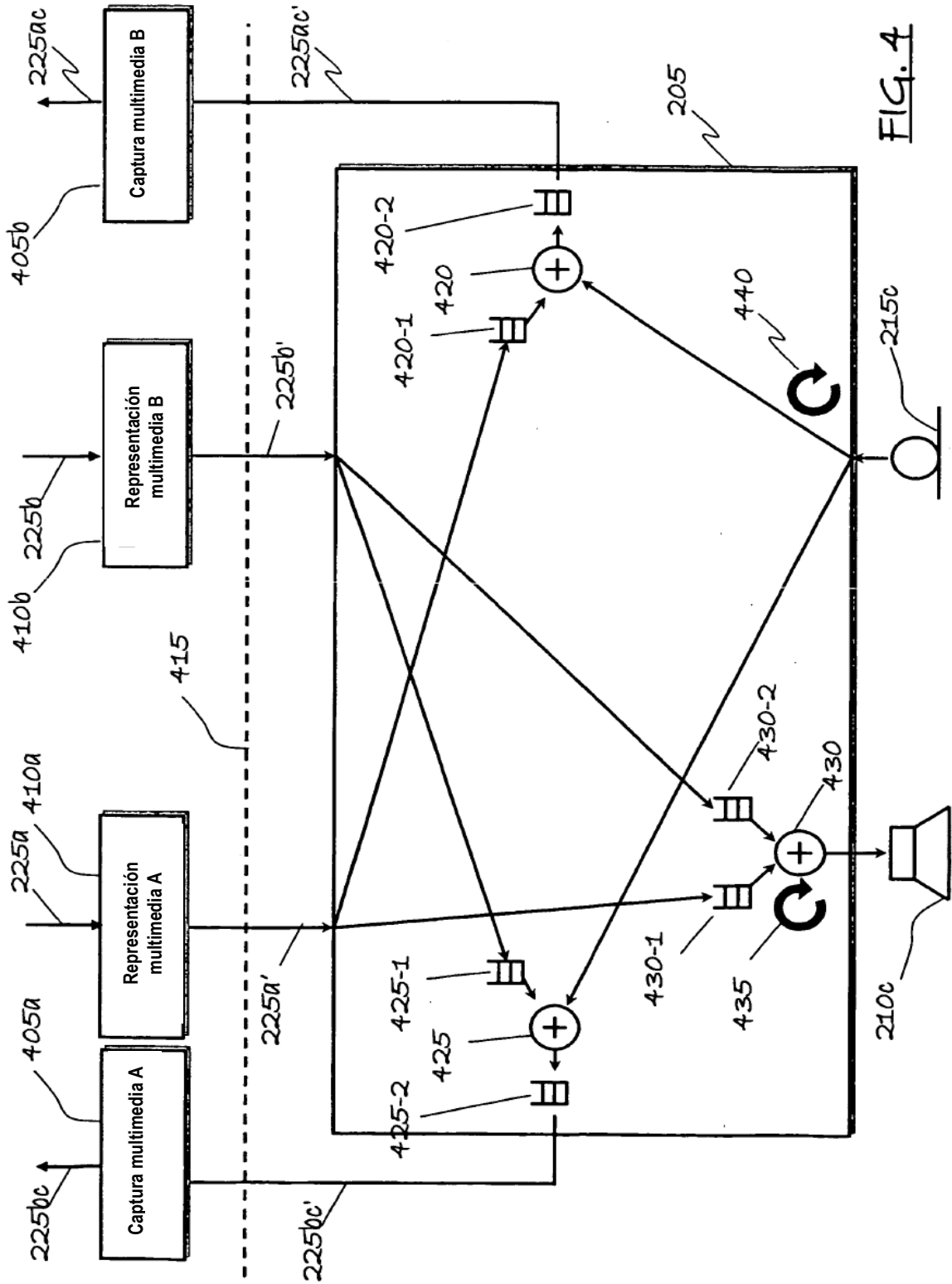
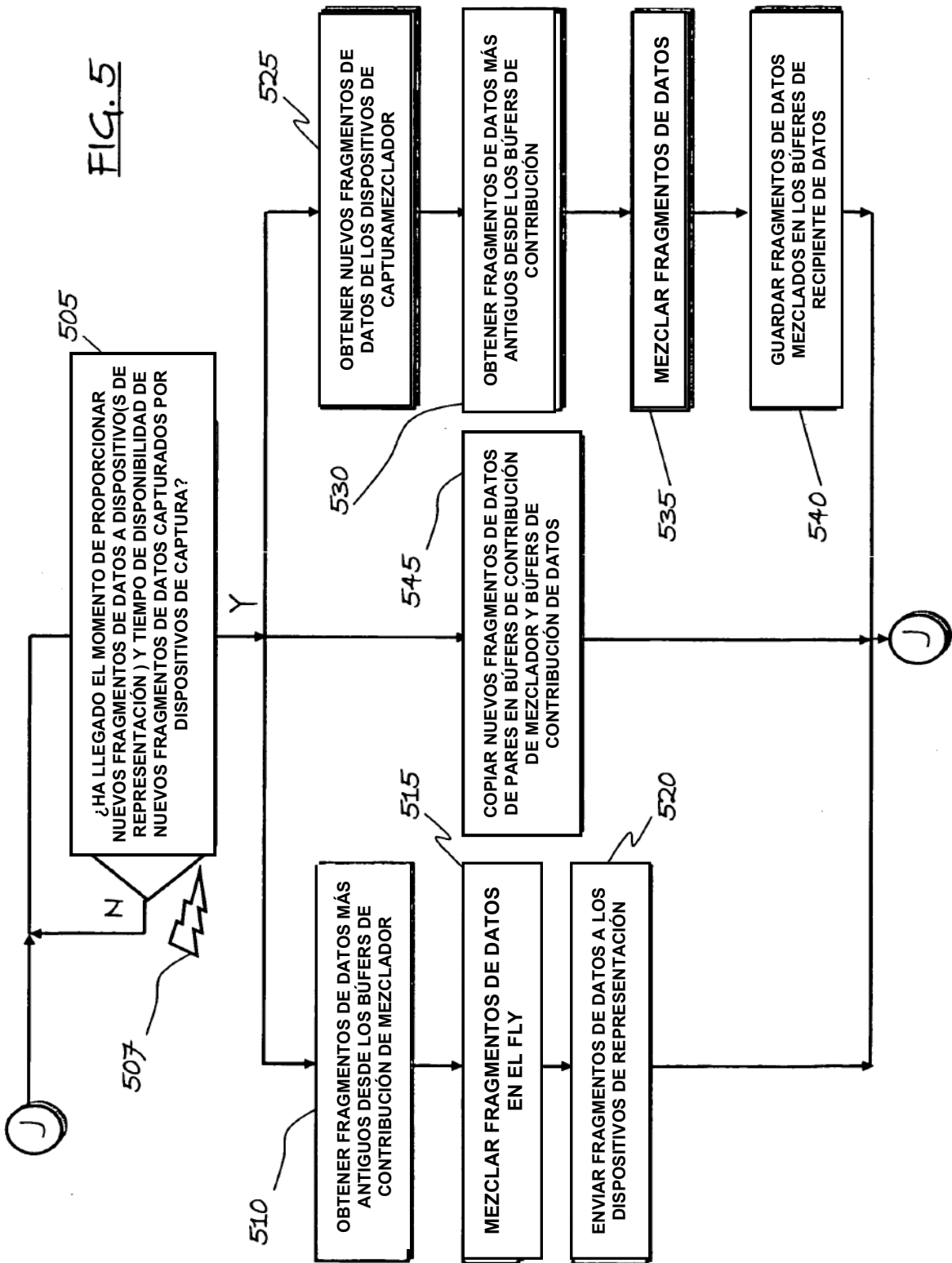


FIG. 4



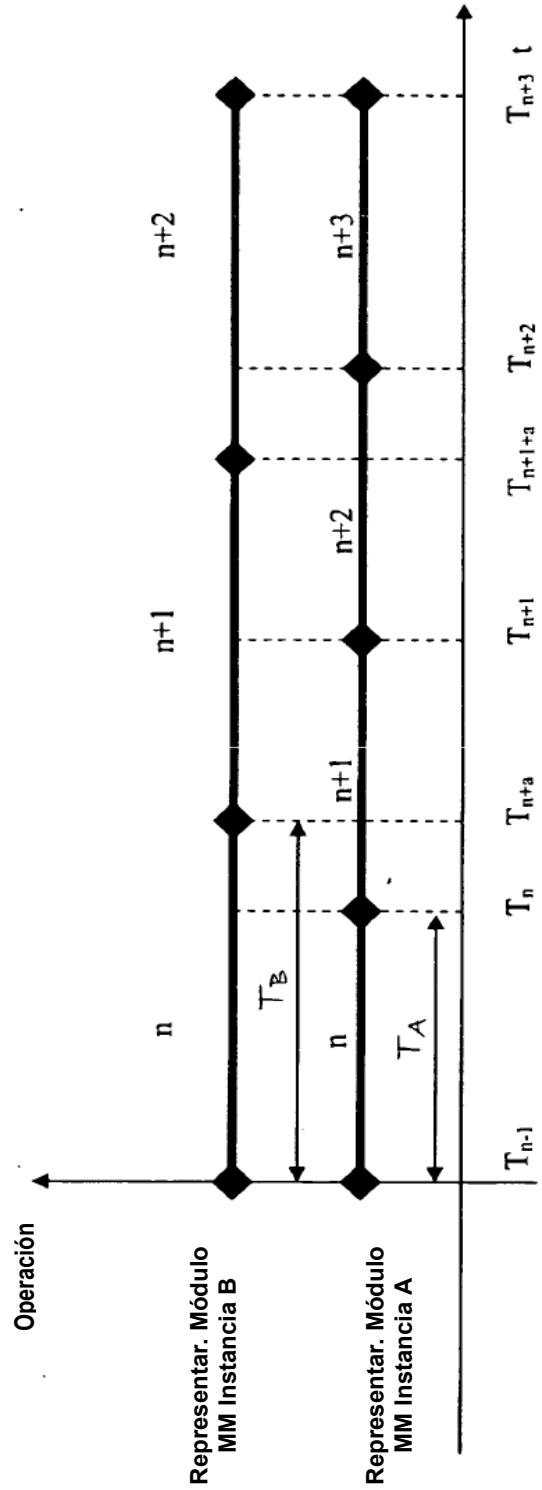
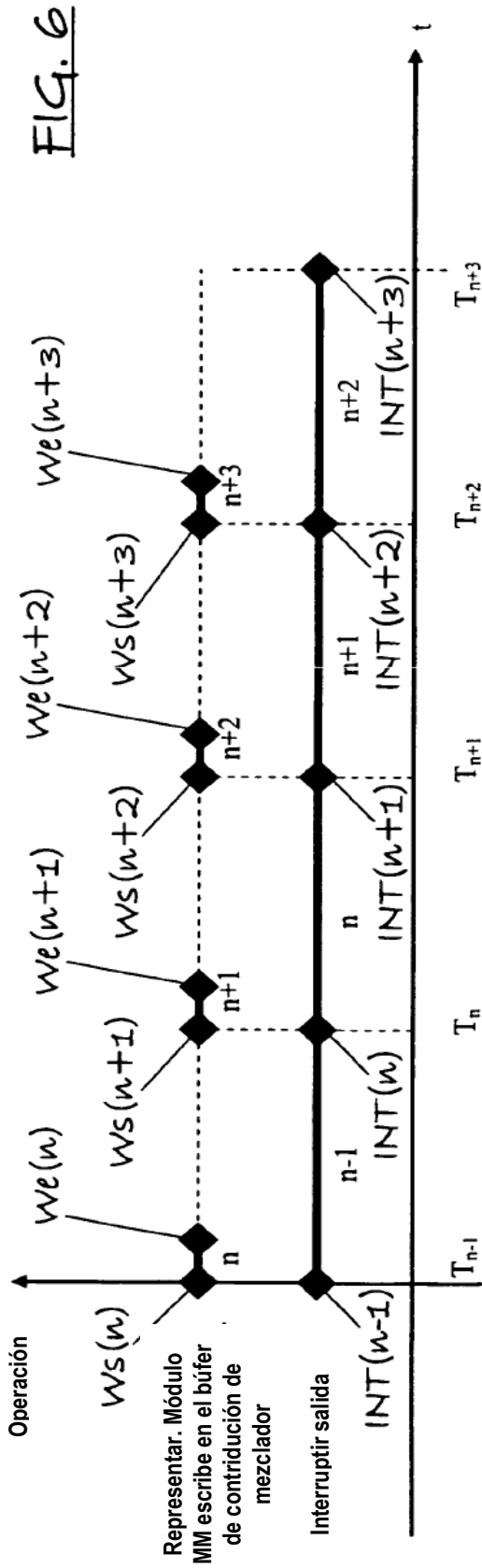


FIG. 7