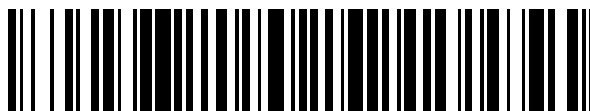


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 585 003**

51 Int. Cl.:

**H04N 7/15** (2006.01)

**H04M 3/56** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.12.2011 E 11851966 (9)**

97 Fecha y número de publicación de la concesión europea: **04.05.2016 EP 2574051**

54 Título: **Método, aparato y sistema para procesar sitios de conferencia en cascada en una conferencia en cascada**

30 Prioridad:

**24.12.2010 CN 201010605183**

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

**03.10.2016**

73 Titular/es:

**HUAWEI DEVICE CO., LTD. (100.0%)  
Building B2 Huawei Industrial Base Bantian  
Longgang District  
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:

**LIANG, LIYAN**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 585 003 T3**

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

## DESCRIPCIÓN

Método, aparato y sistema para procesar sitios de conferencia en cascada en una conferencia en cascada

## CAMPO DE LA INVENCION

5 La presente invención se refiere al campo de las tecnologías de comunicaciones, y en particular, a un método, un aparato, y un sistema para procesar sitios de conferencia en cascada en una conferencia en cascada.

## ANTECEDENTES DE LA INVENCION

10 En una video conferencia común, en general es mantenida una conferencia entre sitios de conferencia ordinaria bajo una unidad de control multipunto (MCU, Multipoint Control Unit), en particular, sitios de conferencia ordinaria en una conferencia están todos conectados sobre una misma MCU. Sin embargo, como la capacidad de conferencia es aumentada o la interconexión en red es cada vez más compleja, se requiere mantener una conferencia en cascada, en particular, no sólo sitios de conferencia bajo cada MCU participan en la conferencia, sino también múltiples conferencias MCU están conectadas a una conferencia a través de un sitio de conferencia en cascada entre múltiples MCU, de modo que consigan el objetivo de que múltiples sitios de conferencia MCU mantengan una conferencia juntos. Por ejemplo, un cierto sistema necesita mantener una conferencia a nivel nacional, y el sistema tiene una MCU y un sitio de conferencia en Beijing, cada capital de provincia, y cada ciudad con nivel de prefectura, y cada país, de modo que la conferencia en cascada de un rango nacional puede ser mantenida, las MCU están previstas en Beijing, cada capital de provincia, y cada ciudad con nivel de prefectura, respectivamente, y cada sitio de conferencia está conectado a una MCU a la que pertenece el sitio de conferencia. Debido a que los sitios de participación de conferencia son numerosos, y están dispersados en diferentes lugares, cada sitio de conferencia sólo se requiere que esté conectado a la MCU más cercana a través de una conferencia en cascada, reduciendo de este modo las demandas para una red.

20 Un ejemplo específico es tomado a continuación para describir un método existente para procesar sitios de conferencia en cascada en una conferencia en cascada. Como se ha mostrado en la fig. 1, en la conferencia en cascada que incluye sitios de conferencia de telepresencia, una MCU1 está conectada a tres sitios de conferencia que son los sitios de conferencia de telepresencia T1 y T3 y un sitio de conferencia ordinaria T2, el sitio de conferencia de telepresencia T1 incluye tres pantallas que son T1L, T1C, y T1R, y el sitio de conferencia de telepresencia T3 incluye tres pantallas que son T3L, T3C, y T3R; una MCU2 está conectada a tres sitios de conferencia que son los sitios de conferencia de telepresencia T4 y T6 y un sitio de conferencia ordinaria T5, el sitio de conferencia de telepresencia T4 incluye tres pantallas que son T4L, T4C, y T4R, y el sitio de conferencia de telepresencia T6 incluye tres pantallas que son T6L, T6C, y T6R.

30 Se ha asumido que cada MCU soporta la reserva de datos de audio de dos partes con voces máximas, en particular, la MCU como máximo selecciona datos de audio de dos sitios de conferencia con las voces máximas desde todos los sitios de conferencia conectados (incluyendo el sitio de conferencia ordinaria, el sitio de conferencia de telepresencia, y el sitio de conferencia en cascada) para realizar la mezcla de sonido, y si hay menos de dos sitios de conferencia conectados, la MCU selecciona datos de todos los sitios de conferencia conectados para realizar la mezcla de sonido.

35 Si la MCU1 y la MCU2 realizan una conferencia en cascada, un canal de audio en cascada es T12, y se ha asumido que en un canal de video en cascada es una corriente de código de video la que está la pantalla intermedia T1C del sitio de conferencia T1. El procesamiento de mezcla de sonido de conferencia es: en el procesamiento de MCU1, se ha asumido que el canal en cascada es la mezcla de sonido de dos partes con las voces máximas, y se ha asumido que cuando dos partes con las voces máximas en la MCU1 son los sitios de conferencia T1 y T2, una corriente de código de mezcla de sonido emitida por la MCU1 a la MCU2 a través del canal de audio en cascada es  $T12 = T1 + T2$ . En la MCU2, se ha asumido que cuando dos partes con las voces masivas son los sitios de conferencia en cascada T12 y T5, si las imágenes presentadas por los sitios de conferencia de telepresencia T4 y T6 son T1C, T5, T6R, y T4L, T1C, T5, respectivamente, un caso de voces oídas en los sitios de conferencia T4 y T6 es como sigue:

45 La voz oída en el sitio de conferencia T4 es  $T12 + T5$ , en particular,  $T1 + T2 + T5$ , y mientras tanto debido a que el sitio de conferencia T4 es un sitio de conferencia de telepresencia, imágenes presentadas por las tres pantallas del sitio de conferencia T4 son imágenes de la pantalla intermedia, en particular, T1C, del sitio de conferencia T1, T5, y la pantalla derecha T6R del sitio de conferencia T6, respectivamente. Debido a que el usuario pretende que la orientación de la imagen de la pantalla corresponda a la orientación de la voz oída en el sitio de conferencia T4, en particular, la voz de T1 es oída en la izquierda, la voz de T5 es oída en la intermedia, y la voz de T6 es oída en la derecha, y debido a que la propia voz de cada sitio de conferencia tiene una cierta orientación, que es innecesariamente consistente con la orientación en que la imagen es presentada, la MCU2 necesita procesar la voz oída en T4, en particular, realizada el ajuste de orientación en la voz de cada sitio de conferencia, ajusta la voz a la orientación de la imagen correspondiente, y luego, realiza la mezcla de sonido para emitir la voz al sitio de conferencia T4, y de este modo, la orientación de la voz oída en T4 puede corresponder a la orientación de la imagen.

55 El sitio de conferencia de telepresencia T6 tiene el mismo problema que en T4, y también se requiere ajustar la orientación de la voz ( $T12 + T5$ ) oída en T6, de modo que la orientación ajustada de la voz corresponde a la orientación

de la imagen vista. Para los sitios de conferencia T4, T5, y T6, debido a que estos tres sitios de conferencia están conectados directamente a la MCU2, la MCU2 puede procesar directamente datos de audio de los sitios de conferencia, de modo que adapta el ajuste de orientación de los sitios de conferencia T4 y T6, respectivamente.

5 En una solución convencional, T12 es un sitio de conferencia en cascada, y los datos de audio son un resultado de la mezcla de sonido de la MCU de nivel previo, en particular, es la suma de datos de los sitios de conferencia T1 y T2. Debido a que ambos T4 y T6 presentan la imagen del sitio de conferencia T1C, pero las ubicaciones de la imagen son diferentes, si la orientación de audio de T1 es ajustada de acuerdo con la ubicación donde la imagen es presentada en cada sitio de conferencia, debido a que los datos de T1 y de T2 no pueden ser separados, es decir, la orientación de voz de T2 es ajustada simultáneamente, y debido a que las orientaciones de la imagen de T1 vistas en dos sitios de conferencia son diferentes, lo que inevitablemente causa que las orientaciones de T2 oídas en T4 y T6 son diferentes, el efecto de correspondencia uno a uno entre la orientación de imagen y la orientación de voz de cada sitio de conferencia en la conferencia en cascada no se puede lograr.

15 Puede verse a partir de las descripciones anteriores que, los datos de audio del sitio de conferencia en cascada son utilizados como el resultado de la mezcla de sonido de la MCU de nivel anterior, y la orientación de voz de los datos de audio del sitio de conferencia en cascada no es consistente usualmente con la ubicación en la que la imagen del sitio de conferencia es presentada; debido a que es utilizado como el resultado de la mezcla de sonido, los datos no pueden ser separados, y para diferentes pantallas de presentación, cuando la orientación de audio es ajustada, el audio correspondiente a la pantalla de presentación no puede ser ajustado individualmente, mientras el resultado de la mezcla de sonido es ajustado de una manera unificada, en particular, la orientación de audio que no debería ser ajustada es también ajustada, la correspondencia uno a uno entre la orientación de imagen y la orientación de voz de cada sitio de conferencia en la conferencia en cascada no puede ser implementada, reduciendo de este modo la experiencia de usuario de un participante.

20 El documento US 2008/0267282 A1 describe un método para optimizar ancho de banda requerido por un sistema de conferencia de video. De acuerdo con el método, corrientes de video particulares para transmisión a otros participantes son seleccionadas y gestionadas individualmente.

#### RESUMEN DE LA INVENCION

Realizaciones de la presente invención proporcionar un método, un aparato, y un sistema para procesar sitios de conferencia en cascada en una conferencia en cascada, que son utilizados para implementar correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en la conferencia en cascada, y mejorar la experiencia de usuario de un participante.

La invención está definida por las reivindicaciones adjuntas.

35 Un método para procesar sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención incluye: recibir una corriente de código de audio enviada por un sitio de conferencia en cascada, donde la corriente de código de audio enviada por el sitio de conferencia en cascada es enviada a sitios de conferencia diferentes a través de canales de sonido de audio diferentes o de canales de cascada de audio; recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada; seleccionar datos de audio que satisfacen una condición previamente configurada a partir de datos de audio que han de ser seleccionados, donde los datos de audio que han de ser seleccionados incluyen: la corriente de código de audio recibida enviada por el sitio de conferencia en cascada y la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada; y ajustar una secuencia de orientación de los datos de audio que satisface la condición previamente configurada.

40 Otro método para el procesamiento de sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención incluye: recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada; seleccionar datos de audio que satisfacen una condición previamente configurada a partir de datos de audio que han de ser seleccionados, en que los datos de audio que han de ser seleccionados incluyen al menos la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada; procesar los datos de audio que satisfacen la condición previamente configurada a diferentes sitios de conferencia a través de diferentes canales de audio o canales de audio en cascada para obtener datos de audio del sitio de conferencia en cascada, de modo que un primer sitio de conferencia en cascada puede identificar los datos de audio del sitio de conferencia en cascada; codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio; y enviar la corriente de código de audio al primer sitio de conferencia en cascada.

55 Un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención incluye: una unidad de recepción, configurada para recibir una corriente de código de audio enviada por un sitio de conferencia en cascada, donde la corriente de código de audio enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de canales de sonido de audio diferentes o de canales de audio en cascada, y además configurada para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada; una unidad de selección, configurada para seleccionar datos de audio que satisfacen una condición previamente configurada a partir de los datos de audio que han de ser seleccionados, donde los

datos de audio que han de ser seleccionados incluyen: la corriente de código de audio recibida enviada por el sitio de conferencia en cascada y la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada; y una unidad de ajuste de secuencia, configurada para ajustar una secuencia de orientación de los datos de audio que satisface la condición previamente configurada.

- 5 Otro aparato para procesar sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención incluye: una unidad de recepción, configurada para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada; una unidad de selección, configurada para seleccionar datos de audio que satisfacen una condición previamente configurada a partir de los datos de audio que han de ser seleccionados, donde los datos de audio que han de ser seleccionados incluyen al menos la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada; una unidad de procesamiento, configurada para procesar los datos de audio que satisfacen la condición previamente configurada a diferentes sitios de conferencia a través de canales de sonido de audio diferentes o de canales de audio en cascada para obtener datos de audio del sitio de conferencia en cascada, de modo que un primer sitio de conferencia en cascada puede identificar los datos de audio del sitio de conferencia en cascada; una unidad de codificación, configurada para codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio; y una unidad de envío, configurada para enviar la corriente de código de audio al primer sitio de conferencia en cascada.

La presente invención proporciona además un sistema de conferencia en cascada.

Puede verse a partir de las soluciones técnicas anteriores que, las realizaciones de la presente invención tienen las siguientes ventajas:

- 20 Puede verse a partir de las soluciones técnicas anteriores previstas en las realizaciones de la presente invención que, debido a que la corriente de código de audio enviada por el sitio de conferencia en cascada y recibida por el aparato para el procesamiento de los sitios de conferencia en cascada en la conferencia en cascada de las realizaciones de la presente invención es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada, cuando el aparato para procesar los sitios de conferencia en cascada en la conferencia en cascada necesita realizar un ajuste de secuencia de orientación en los datos de audio, la orientación de los datos de audio requerida para ser ajustada puede ser ajustada directamente de forma individual, en particular, el ajuste realizado por el aparato para procesar los sitios de conferencia en cascada en la conferencia en cascada en una única orientación de datos de audio no influye en otros datos de audio, y la correspondencia uno a uno entre la orientación de imagen y la orientación de voz de cada sitio de conferencia en la conferencia en cascada puede ser implementada, mejorando de este modo la experiencia de usuario del participante.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

- 35 Para ilustrar las soluciones técnicas en las realizaciones de la presente invención más claramente, los dibujos adjuntos requeridos para describir las realizaciones son introducidos brevemente a continuación. Aparentemente, los dibujos adjuntos en la siguiente descripción son meramente algunas realizaciones de la presente invención, y expertos en la técnica pueden derivar además otros dibujos de acuerdo con un estos dibujos adjuntos.

La fig. 1 es un diagrama estructural de sitios de conferencia en cascada convencionales en una conferencia en cascada;

La fig. 2 es un diagrama esquemático de una realización de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

- 40 La fig. 3 es un diagrama esquemático de otra realización de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

La fig. 4 es un diagrama esquemático de otra realización de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

La fig. 5 es un diagrama esquemático de otra realización de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

- 45 La fig. 6 es un diagrama estructural de sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

La fig. 7 es un diagrama estructural de sitios de conferencia en cascada en otra conferencia en cascada de acuerdo con una realización de la presente invención;

- 50 La fig. 8 es un diagrama esquemático de un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención;

La fig. 9 es un diagrama esquemático de otro aparato para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención; y

La fig. 10 es un diagrama esquemático de otro aparato para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

5 Realizaciones de la presente invención proporcionan un método, y un aparato, y un sistema para procesar sitios de conferencia en cascada en una conferencia en cascada, que son utilizados para implementar la correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en una conferencia en cascada, y mejorar la experiencia de usuario de un participante.

10 Para hacer los objetivos, características, y ventajas de la presente invención más claros y más comprensibles, las soluciones técnicas en la realizaciones de la presente invención son descritas de forma clara y de forma completa a continuación con referencia a los dibujos adjuntos en la realizaciones de la presente invención. Aparentemente, las realizaciones en la siguiente descripción son meramente una parte en lugar de todas las realizaciones de la presente invención. Todas las demás realizaciones obtenidas por expertos en la técnica basadas en las realizaciones de la presente invención caerán dentro del marco de protección de la presente invención.

15 Un método para procesar sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención es introducido por primera vez. Con referencia a la fig. 2, se ha descrito un proceso de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención. El proceso de procesamiento de un aparato que es para procesar sitios de conferencia en cascada en una conferencia en cascada y es utilizado como un extremo de recepción está descrito en la realización; el aparato para procesar sitios de conferencia en cascada en una conferencia en cascada está conectado a un sitio de conferencia en cascada, y además está conectado a un sitio de conferencia que no está en cascada, tal como: un sitio de conferencia ordinaria y/o un sitio de conferencia de telepresencia. La realización de la presente invención incluye las siguientes operaciones:

201: Recibir una corriente de código de audio enviada por una sitio de conferencia en cascada.

25 Específicamente, puede ser recibida la corriente de código de audio enviada por el sitio de conferencia en cascada conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada. En la realización de la presente invención, la corriente de código de audio recibida enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de canales de sonido de audio en cascada o de canales de audio en cascada, que son diferentes de un sistema en el que el sitio de conferencia en cascada realiza la mezcla de sonido en la corriente de código de audio y luego envía la corriente de código de audio.

30 En la realización de la presente invención, un canal de audio en cascada puede solamente ser ajustado, el canal de audio en cascada incluye más de dos canales de sonido de audio, y la corriente de código de audio es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio, en particular, corrientes de código de audio de diferentes sitios de conferencia son transmitidos en diferentes canales de sonido de audio, respectivamente.

35 En la realización de la presente invención, más de los canales de audio en cascada pueden también ser ajustados, y la corriente de código de audio es enviada a diferentes sitios de conferencia a través de diferentes canales de audio en cascada, en particular, corrientes de código de audio de diferentes sitios de conferencia son transmitidas en los canales de audio en cascada, respectivamente. Cuando la corriente de código de audio es enviada a diferentes sitios de conferencia a través de diferentes canales de audio en cascada, la realización de la presente invención incluye además: recibir información de composición del canal de audio en cascada enviada por el sitio de conferencia en cascada, en que la información de composición del canal de audio en cascada es información de cantidad de canales de audio en cascada establecida por el sitio de conferencia en cascada, de modo que obtenga la cantidad de canales de audio en cascada ocupada por la corriente de código de audio recibida enviada por el sitio de conferencia en cascada.

202: Recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada.

45 Específicamente, la corriente de código de audio enviada por un sitio de conferencia ordinaria conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada, y/o la corriente de código de audio que corresponde a cada pantalla de un sitio de conferencia de telepresencia conectada al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada pueden ser recibidas. En la realización de la presente invención, múltiples corrientes de código de audio que corresponden a cada pantalla del sitio de conferencia de telepresencia, respectivamente, son introducidas como una única corriente de código de audio, respectivamente, las múltiples corrientes de código de audio son independientes entre sí, y el sitio de conferencia de telepresencia envía las múltiples corrientes de código de audio de forma independiente.

55 Debería observarse que, en la realización de la presente invención, no hay diferencia de secuencia entre la operación 201 y la operación 202, y la operación 201 puede ser ejecutada en primer lugar y luego la operación 202 es ejecutada, y la operación 202 puede ser ejecutada en primer lugar y luego la operación 201 es ejecutada, y las operaciones 201 y 202 puede ser ejecutadas simultáneamente, lo que no está limitado aquí.

203: Seleccionar datos de audio que satisfacen una condición previamente configurada a partir de datos de audio que han de ser seleccionados.

Los datos de audio que satisfacen la condición previamente configurada son seleccionados a partir de los datos de audio que han de ser seleccionados, donde datos de audio de cada sitio de conferencia utilizado como una pieza de datos de audio participan en la selección acerca de si satisfacer la condición previamente configurada, y los datos de audio que han de ser seleccionados incluyen: la corriente de código de audio recibida enviada por el sitio de conferencia en cascada y la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada. La cantidad de los datos de audio seleccionados es menor o igual que la cantidad predeterminada, donde la cantidad predeterminada es ajustada de antemano de acuerdo con la condición previamente configurada.

En la realización de la presente invención, la condición previamente configurada puede ser específicamente reservar datos de audio de varias partes con las voces máximas en un sitio de conferencia en cascada y en un sitio de conferencia que no está en cascada. Por ejemplo, cuando se reservan datos de audio de tres partes con las voces máximas en el sitio de conferencia en cascada y en el sitio de conferencia que no está en cascada, se seleccionan las primeras tres piezas de datos de audio de todos los datos de audio con volúmenes de voz desde el mayor al menor. De acuerdo con el principio de reservar datos de audio de varias partes con las voces máximas, datos de audio que satisfacen la condición previamente configurada son filtrados o tamizados a partir de los datos de audio, y la cantidad de datos de audio seleccionada es menor que o igual a la cantidad predeterminada.

En la realización de la presente invención la condición previamente configurada puede además ser específicamente para reservar datos de audio de un sitio de conferencia preestablecido, por ejemplo, puede ser preestablecido un sitio de conferencia específico, que puede ser específicamente uno o más de los dos sitios de conferencia designados por un usuario; en la realización de la presente invención, que satisfacen la condición previamente configurada es reservar datos de audio del sitio de conferencia preestablecido, por ejemplo, los datos de audio que han de ser seleccionados incluyen datos de audio de un sitio de conferencia ordinaria T1, datos de audio de un sitio de conferencia ordinaria T2, datos de audio de un sitio de conferencia de telepresencia T3, y datos de audio de un sitio de conferencia de telepresencia T4; se ha asumido que la condición previamente configurada es sólo para reservar datos de audio del sitio de conferencia T1 y datos de audio del sitio de conferencia T2 que son designados por el usuario, los datos de audio del sitio de conferencia T1 y los datos de audio del sitio de conferencia T2 pueden ser reservados a través del filtrado de la condición previamente configurada, mientras que debido a que los datos de audio del sitio de conferencia T3 y los datos de audio del sitio de conferencia T4 no son datos de audio de los sitios de conferencia designados por el usuario, no son reservados.

Debería observarse que, en una aplicación práctica, la condición previamente configurada puede además ser específicamente para satisfacer otra condición, siempre que las condiciones pueden filtrar datos de audio de diferentes sitios de conferencia, que no están limitados aquí.

En la realización de la presente invención, la cantidad predeterminada es preestablecida, y es determinada específicamente de acuerdo con datos de audio reservados de varias partes con las voces máximas seleccionadas por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada. Por ejemplo, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada soporta la reserva de datos de audio de dos partes con las voces máximas, y entonces, la cantidad predeterminada es 2; el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada soporta la reserva de datos de audio de tres partes con las voces máximas, y entonces la cantidad predeterminada es 3. Por ejemplo, cuando el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada soporta la reserva de datos de audio de tres partes con las voces máximas, si la cantidad de sitios de conferencia (incluyendo el sitio de conferencia ordinaria, el sitio de conferencia de telepresencia, y el sitio de conferencia en cascada) conectados al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada es mayor o igual a tres, tres piezas de datos de audio con los volúmenes máximos pueden ser seleccionadas como datos de audio del sitio de conferencia en cascada; y si la cantidad de sitios de conferencia conectados al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada es menor de tres, sólo pueden ser seleccionadas menos de tres piezas de datos de audio.

204: Ajustar una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada.

Después de que los datos de audio que satisfacen la condición previamente configurada son seleccionados, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada ajusta una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada, lo que puede ser implementado específicamente adoptando la siguiente manera:

Si sólo se ha presentado una pantalla de una fuente de video del sitio de conferencia que no está en cascada en una de las múltiples pantallas o en una imagen en una ubicación en una imagen múltiple, una secuencia de orientación de salida de datos de audio que corresponde a todas las pantallas de la fuente de video del sitio de conferencia que no está en cascada es la misma que una secuencia de orientación de presentación de la pantalla presentada en múltiples pantallas o es una secuencia de orientación de la pantalla presentada en la imagen en la ubicación en la imagen múltiple. Por ejemplo, un sitio de conferencia de telepresencia T1 incluye tres pantallas T1L, T1C, y T1R, sólo la pantalla T1L entre

ellas es presentada, orientaciones de salida de datos de audio que corresponden a las tres pantallas T1L, T1C, y T1R del sitio de conferencia de telepresencia son todas las mismas que las secuencias de orientación de presentación de la pantalla T1L presentada en las múltiples pantallas, o son las mismas que en la secuencia de orientación de la pantalla T1L presentada en una imagen en la localización en una imagen múltiple.

- 5 Si se han presentado más de dos pantallas en múltiples pantallas de la fuente de video del sitio de conferencia que no está en cascada, una secuencia de orientación de salida de datos de audio que corresponde a las más de dos pantallas presentadas de la fuente de video corresponde a una secuencia de orientación de estas pantallas presentadas de la fuente de video una a una. La secuencia de orientación de salida de los datos de audio que corresponde a pantallas no presentadas de múltiples pantallas de la fuente de video es mantenida igual que la de una vez todas las pantallas presentadas de la fuente de video. Por ejemplo, el sitio de conferencia de telepresencia T1 incluye tres pantallas T1L, T1C, y T1R, sólo las pantallas T1L y T1C son presentadas, mientras T1R no es presentada, las secuencias de orientación de salida de datos de audio que corresponden a T1L y T1C presentadas son las mismas que las secuencias de orientación de T1L y T1C presentadas, y los datos de audio que corresponde a la T1R no presentada puede ser los mismos que la secuencia de orientación de una de todas las pantallas presentadas T1L y T1C.
- 10
- 15 Si se presenta simultáneamente una pantalla en la fuente de video del sitio de conferencia que no está en cascada en múltiples imágenes o en una imagen múltiple y una pantalla independiente, las prioridades de secuencias de orientación de salida de datos de audio que corresponden a la fuente de video desde la más alta a la más baja en secuencia son: una orientación de la pantalla independiente, una orientación de una pantalla con una gran imagen secundaria, y una orientación de una pantalla presentada en las prioridades intermedia, izquierda, y derecha. Por ejemplo, el sitio de conferencia de telepresencia T1 incluye tres pantallas T1L, T1C, y T1R, y una pantalla T1L es presentada en múltiples imágenes múltiples o en una imagen múltiple y en una pantalla independiente, las prioridades de secuencia de las orientaciones de salida de datos de audio que corresponden a la pantalla T1L en secuencia son: una orientación de la pantalla independiente, una orientación de una pantalla con una gran imagen secundaria, y una orientación de una pantalla presentada en las prioridades intermedia, izquierda, y derecha.
- 20
- 25 Debería observarse que, la realización anterior es sólo una de las maneras donde el ajuste de orientación es realizado en los datos de audio de acuerdo con la secuencia de orientación de la fuente de video, y también puede adoptarse otra manera de implementación, siempre que la secuencia de orientación de los datos de audio de salida cumpla un requisito de secuencia específico. Por ejemplo, si los datos de audio del sitio de conferencia que no está en cascada satisfacen la condición previamente configurada y son reservados, pero la imagen del sitio de conferencia no se ve, en particular, este sitio de conferencia es un sitio de conferencia invisible, la política de ajuste puede ser: mantener la secuencia de orientación de los datos de audio del propio sitio de conferencia, o, ajustar la secuencia de orientación de los datos de audio en una ubicación fija, tal como, fijar en la orientación intermedia para dispensar el sonido, o fijar en las orientaciones en dos lados, y además, debido a que la imagen del sitio de conferencia no se ve, la secuencia de orientación también se puede fijar en un cierto lugar fuera de la pantalla.
- 30
- 35 En la realización de la presente invención, debido a que la corriente de código de audio enviada por el sitio de conferencia en cascada y recibida por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada, cuando el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada necesita realizar el ajuste de secuencia de orientación en los datos de audio, los datos de audio que no necesitan ser ajustados pueden ser ajustados de forma individual, en particular, el ajuste realizado por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada en una única orientación de datos de audio no influye en otros datos de audio, y la correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en la conferencia en cascada puede ser implementada, mejorando de este modo la experiencia de usuario de un participante.
- 40
- 45 Debería observarse que, en la realización mostrada en la fig. 2, después de la operación 202, el método puede incluir además: descodificar la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada; y los datos de audio que han de ser seleccionados específicamente incluyen: un resultado de descodificar la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada.
- 50
- 55 En la realización anterior, se ha descrito que después de recibir la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada realiza el ajuste de secuencia de orientación en los datos de audio. Se ha descrito a continuación desde la perspectiva de que el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada envía una corriente de código de audio a un primer sitio de conferencia en cascada, puede hacerse referencia a la siguiente realización.

Un método para procesar sitios de conferencia en cascada en una conferencia en cascada previsto en una realización de la presente invención es introducido por primer vez. Con referencia a la fig. 3, se ha descrito un proceso de un método para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención. Un proceso de procesamiento de un aparato para procesar sitios de conferencia en cascada en una

conferencia en cascada se ha descrito en la realización, y el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada está conectado al primer sitio de conferencia en cascada, tal como un sitio de conferencia ordinaria y/o un sitio de conferencia de telepresencia. La realización de la presente invención incluye las siguientes operaciones:

5 301: Recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada.

10 Específicamente, puede ser recibida una corriente de código de audio enviada por un sitio de conferencia ordinaria conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada, y/o una corriente de código de audio que corresponde a cada pantalla de un sitio de conferencia de telepresencia conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada. En la realización de la presente invención, múltiples corrientes de código de audio que corresponden a cada pantalla del sitio de conferencia de telepresencia, respectivamente, son introducidas como una corriente de código de audio y viente, respectivamente, las múltiples corrientes de código de audio son independientes entre sí, y el sitio de conferencia de telepresencia envía las múltiples corrientes de código de audio de forma independiente.

15 302: Seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados.

20 Los datos de audio que ha de ser seleccionados incluyen al menos la corriente de código de audio enviada por el sitio de conferencia que no está en cascada, los datos de audio que satisfacen la condición previamente configurada son seleccionados a partir de los datos de audio que han de ser seleccionados, y la cantidad de los datos de audio seleccionados es menor o igual que la cantidad predeterminada, donde la cantidad predeterminada es preestablecida de acuerdo con la condición previamente configurada, y para la descripción de la condición previamente configurada, puede hacerse referencia a la descripción de la operación 203 en la fig. 2.

303: Procesar los datos de audio que satisfacen la condición previamente configurada para diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada.

25 Después de que el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada seleccione los datos de audio que satisfacen la condición previamente configurada, los datos de audio pueden ser procesados de acuerdo con diferentes sitios de conferencia que ocupan diferentes canales de sonido de audio o de canales de audio en cascada, respectivamente, para obtener datos de audio del sitio de conferencia en cascada, de modo que el primer sitio de conferencia en cascada pueda identificar los datos de audio del sitio de conferencia en cascada.

30 Debería observarse que, en la realización de la presente invención, un sitio de conferencia es tomado como una unidad para el procesamiento en los datos de audio, y el procesamiento es realizado de acuerdo con cada canal de sonido de audio o canal de audio en cascada, respectivamente; se puede asegurar que el procesamiento en una única pieza de datos de audio no influye en otros datos de audio, y el método de procesamiento específico se ha ilustrado en la realización subsiguiente.

304: Codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio.

35 Se codifican los datos de audio del sitio de conferencia en cascada obtenidos en la operación anterior, y un resultado de la codificación es utilizado como la corriente de código de audio.

305: Enviar la corriente de código de audio al primer sitio de conferencia en cascada.

40 Después de que se obtiene la corriente de código de audio después de que se realiza la codificación, la corriente de código de audio es enviada al primer sitio de conferencia en cascada. Debería observarse que, en la realización de la presente invención, el primer sitio de conferencia en cascada es un sitio de conferencia en cascada conectado directamente al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada.

45 Puede verse a partir de la realización de la presente invención que, el aparato para procesar sitios de competencia en cascada en la conferencia en cascada de la realización de la presente invención procesa los datos de audio que satisfacen la condición previamente configurada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada para obtener los datos de audio del sitio de conferencia en cascada, de modo que cuando el ajuste de secuencia de orientación necesita ser realizado en los datos de audio, el primer sitio de conferencia en cascada utilizado como extremo de recepción puede ajustar directamente los datos de audio de forma individual.

50 Además, la presente invención proporciona una realización de otro método para procesar sitios de conferencia en cascada en una conferencia en cascada. Específicamente, con referencia a la fig. 4, un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada está conectado a un primer sitio de conferencia en cascada y a un segundo sitio de conferencia en cascada, y además está conectado a un sitio de conferencia que no está en cascada. En la realización de la presente invención, sólo un canal de audio en cascada es ajustado, y el canal de audio en cascada incluye más de dos canales de sonido de audio, en particular, una corriente de código de audio es transmitida



en cada canal de sonido de audio, respectivamente. La realización de la presente invención incluye:

401: Recibir una corriente de código de audio enviada por el sitio de conferencia que no está en cascada y una corriente de código de audio enviada por el segundo sitio de conferencia en cascada.

5 Específicamente, puede ser recibida una corriente de código de audio enviada por un sitio de conferencia ordinaria conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada, y/o una corriente de código de audio que corresponde a cada pantalla de un sitio de conferencia de telepresencia conectado al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada. En la realización de la presente invención, múltiples corrientes de código de audio que corresponden a cada pantalla del sitio de conferencia de telepresencia, respectivamente, son introducidos como una única corriente de código de audio, respectivamente, las múltiples corrientes de código de audio son independientes entre sí, y el sitio de conferencia de telepresencia envía las múltiples corrientes de código de audio de forma independiente.

10 Cuando el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada está conectado al segundo sitio de conferencia en cascada, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada también recibe la corriente de código de audio enviada por el segundo sitio de conferencia en cascada. Debería observarse que, en la realización de la presente invención, el segundo sitio de conferencia en cascada es un sitio de conferencia en cascada conectado directamente al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada.

402: Seleccionar datos de audio que satisfacen una condición previamente configurada a partir de datos de audio que han de ser seleccionados.

20 Los datos de audio que han de ser seleccionados incluyen al menos la corriente de código de audio enviada por el sitio de conferencia que no está en cascada. En la realización de la presente invención, la condición configurada previamente puede ser específicamente para reservar datos de audio de varias partes con las voces máximas en un sitio de conferencia en cascada y en un sitio de conferencia que no está en cascada. De acuerdo con el principio de reservar datos de audio de varias partes con las voces máximas, datos de audio son filtrados a partir de los datos de audio que han de ser seleccionados, y la cantidad de datos de audio seleccionada es menor o igual que la cantidad predeterminada.

25 En la realización de la presente invención, aquellos que participan en comparación de los datos de audio reservados de varias partes con las voces máximas incluyen la corriente de código de audio enviada por el sitio de conferencia que no está en cascada y la corriente de código de audio enviada por el segundo sitio de conferencia en cascada, donde el sitio de conferencia que no está en cascada y el segundo sitio de conferencia en cascada están conectados al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada; donde la corriente de código de audio del segundo sitio de conferencia en cascada participa en comparación con una corriente de código de audio de un sitio de conferencia. En la operación 402, corrientes de código de audio del segundo sitio de conferencia en cascada pueden ser superpuestas como una corriente de código de audio para participar en comparación. En una aplicación práctica, la información envolvente de la corriente de código de audio del segundo sitio de conferencia en cascada puede además ser enviada al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada de una manera distinta tal como una manera de llenar información de protocolo de transporte en tiempo real (RTP, Real-Time Transport Protocol), y después de que el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada recibe la información envolvente de audio, la información envolvente de audio participa en comparación de los datos de audio reservados de varias partes con las voces máximas.

403: Cuando los datos de audio que satisfacen la condición previamente configurada son datos de audio que corresponden a una pantalla de un sitio de conferencia de telepresencia, utilizar los datos de audio que corresponden a la pantalla del sitio de conferencia de telepresencia como datos de audio que corresponden a un sitio de conferencia separado.

45 En la realización de la presente invención, la pantalla del sitio de conferencia de telepresencia que satisface la condición previamente configurada como un sitio de conferencia separado es transmitida a través de un canal de sonido de audio separado, y el sitio de conferencia en cascada utilizado como un sitio de conferencia también es transmitido a través de un canal de sonido de audio separado.

50 404: Alternativamente, cuando los datos de audio que satisfacen la condición previamente configurada son datos de audio que corresponden a un sitio de conferencia ordinaria y la cantidad de canal de sonido del sitio de conferencia ordinaria no es un único canal de sonido, mezcla los datos de audio del sitio de conferencia ordinaria en los únicos datos de audio del canal de sonido.

55 En la realización de la presente invención, los datos de audio del sitio de conferencia ordinaria que satisfacen la condición previamente configurada y cuya cantidad de canal de sonido no es un único canal de sonido son mezclados en datos de audio con un único canal de sonido y transmitidos a través de un canal de sonido de audio.

405: De acuerdo con la secuencia de orientación de una corriente de código de video que necesita ser enviada, la

secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada es ajustada de acuerdo con diferentes sitios de conferencia, respectivamente.

5 La corriente de código de video que necesita ser enviada es una parte o la totalidad de una corriente de código de video recibida por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada, y cuyas corrientes de código de video que necesitan específicamente ser enviadas al primer sitio de conferencia en cascada pueden ser seleccionadas y decididas por un usuario, y también pueden ser decididas por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada, lo que no está limitado aquí.

En esta operación, los datos de audio que satisfacen la condición previamente configurada y cuya secuencia de orientación es ajustada son utilizados como datos de audio del sitio de conferencia en cascada.

10 En la realización de la presente invención, de acuerdo con la secuencia de orientación de la corriente de código de video que necesita ser enviada, la secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada es ajustada de acuerdo con diferentes sitios de conferencia, respectivamente, de modo que el primer sitio de conferencia en cascada utilizado como un extremo de recepción puede, de acuerdo con la secuencia de orientación de la corriente de código de video, identificar la secuencia de orientación de los datos de audio que satisfacen la condición  
15 previamente configurada.

En una aplicación práctica, si la fuente de video que corresponde a los datos de audio que satisfacen la condición previamente configurada y la corriente de código de video que necesita ser enviada son la misma, la secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada es ajustada a una secuencia de orientación igual que la de la corriente de código de video que necesita ser enviada; si la fuente de video que  
20 corresponde a los datos de audio que satisfacen la condición previamente configurada y la corriente de código de video que necesita ser enviadas son diferentes, la secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada es ajustada de acuerdo con la política descrita en la operación 204 en la fig. 2. Por ejemplo, corrientes de código de video que necesitan ser enviadas son T2, T1C, y T3R. Si los datos de audio que satisfacen la condición previamente configurada son T3R, T2, y T1L, debido a que los datos de audio T2 y la corriente de código de video T2 son los mismos, los datos de audio T2 son ajustados a una secuencia igual que la de la corriente de código de video, en particular, izquierda. Los datos de audio T1L y las corrientes de código de video no son los mismos, pero T1C  
25 existe en las corrientes de código de video, y debido a que tanto la corriente de código de video T1C como la corriente de código de audio T1L son dos corrientes de código de un sitio de conferencia T1, la secuencia de orientación de los datos de audio T1L es ajustada a la secuencia de orientación de la corriente de código presentada T1C, en particular, los datos de audio T1L son ajustados a la secuencia de orientación de la fuente de video T1C, en particular, intermedia. Los datos de audio T3R y la corriente de código de video T3R son los mismos, y los datos de audio T3R son ajustados a la derecha lo mismo que la orientación de la corriente de código de video T3R, y por lo tanto, los datos de audio cuya secuencia de orientación es ajustada son T2, T1L, y T3R.

Debería observarse que, esta operación 405 también puede ser sustituida por la siguiente operación:

35 Formar una secuencia de los datos de audio que satisfacen la condición previamente configurada de acuerdo con diferentes sitios de conferencia que ocupan diferentes canales de sonido de audio, y utilizar los datos de audio secuenciados que satisfacen la condición previamente configurada como los datos de audio del sitio de conferencia en cascada, donde la secuenciación puede ser realizada disponiendo en secuencia de acuerdo con la secuencia de datos de audio de varias partes con las voces máximas, y también puede ser realizada de acuerdo con otra secuencia, que no está limitada aquí.  
40

Y a continuación, se genera la información de ubicación del sitio de conferencia, donde la información de ubicación del sitio de conferencia de audio es información de secuenciación de ubicación de los datos de audio que satisfacen la condición previamente configurada.

45 Finalmente, la información de ubicación del sitio de conferencia de audio generada es enviada al primer sitio de conferencia en cascada, y puede ser enviada al primer sitio de conferencia en cascada durante la implementación específica adoptando una manera de llenado de datos de información de llenado RTP, e igualmente, puede adoptarse otra manera de implementación, que no está limitada aquí.

406: Codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio.

50 Se codifican los datos de audio del sitio de conferencia en cascada obtenidos en la operación anterior, es decir, la corriente de código de audio puede ser obtenida.

407: Enviar la corriente de código de audio al primer sitio de conferencia en cascada.

55 Después de que se obtiene la corriente de código de audio después de que se realiza la codificación, la corriente de código de audio es enviada al primer sitio de conferencia en cascada. Debería observarse que, en la realización de la presente invención, el primer sitio de conferencia en cascada es otro sitio de conferencia en cascada conectado directamente al aparato para procesar sitios de conferencia en cascada en la conferencia en cascada.

En la operación 406 en la realización de la presente invención, de acuerdo con la secuencia de orientación de la corriente de código de video que necesita ser enviada, la secuencia de orientación de los datos de audio del sitio de conferencia en cascada también es ajustada, de modo que el primer sitio de conferencia en cascada utilizado como un extremo de recepción puede, de acuerdo con la secuencia de orientación de la corriente de código de video, identificar la secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada, y el primer sitio de conferencia en cascada puede ajustar de forma individual los datos de audio.

Se puede ver a partir de la realización de la presente invención que, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada de la realización de la presente invención procesa los datos de audio que satisfacen la condición previamente configurada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada para obtener los datos de audio del sitio de conferencia en cascada, de modo que cuando el ajuste de secuencia de orientación necesita ser realizado en los datos de audio, el primer sitio de conferencia en cascada utilizado como el extremo de recepción puede ajustar directamente los datos de audio de forma individual.

Opcionalmente, la presente invención proporciona una realización de otro método para procesar sitios de conferencia en cascada en una conferencia en cascada. Específicamente, con referencia a la fig. 5, el aparato para procesar sitios de conferencia en cascada en una conferencia en cascada está conectado a un primer sitio de conferencia en cascada y a un segundo sitio de conferencia en cascada, y además está conectado a un sitio de conferencia que no está en cascada. En la realización de la presente invención, más de dos canales de audio en cascada están ajustados, lo que es diferente de un sistema en el que sólo existe un canal de audio en cascada, en particular, una corriente de código de audio es transmitida en cada canal de audio en cascada, respectivamente. La realización de la presente invención incluye:

501: Recibir una corriente de código de audio enviada por el sitio de conferencia que no está en cascada y una corriente de código de audio enviada por el segundo sitio de conferencia en cascada.

El contenido en esta operación es el mismo que el contenido en la operación 401 en la realización anterior, y para más detalles, puede hacerse referencia a la operación, que no se ha descrito aquí de forma detallada.

502: Descodificar la corriente de código de audio enviada por el sitio de conferencia que no está en cascada y la corriente de código de audio enviada por el segundo sitio de conferencia en cascada.

Después de haber obtenido la corriente de de código de audio enviada por el sitio de conferencia de cascada y una segunda corriente de código de audio enviada por el segundo sitio de conferencia en cascada, las corrientes de código de audio pueden ser decodificadas. Debería observarse que, cuando la corriente de código de audio es decodificada, es una manera de implementación opcional.

503: Seleccionar datos de audio que satisfacen una condición previamente configurada a partir de los datos de audio que han de ser seleccionados.

Los datos de audio que han de ser seleccionados incluyen específicamente: un resultado de decodificar la corriente de código de audio enviada por el sitio de conferencia que no está en cascada y la corriente de código de audio enviada por el segundo sitio de conferencia en cascada. En la realización de la presente invención, de acuerdo con el principio de reservar datos de audio de varias partes con las voces máximas, los datos de audio pueden ser seleccionados a partir de los datos de audio que han de ser seleccionados, y la cantidad de los datos de audio seleccionada es menor o igual que la cantidad predeterminada.

504: Procesar los datos de audio que satisfacen la condición previamente configurada de acuerdo con diferentes sitios de conferencia que ocupan diferentes canales de audio en cascada para obtener datos de audio del sitio de conferencia en cascada.

En esta operación, los datos de audio que satisfacen la condición previamente configurada son procesados de acuerdo con diferentes sitios de conferencia que ocupan diferentes canales de audio en cascada, es decir, un canal de audio en cascada es solamente utilizado para transmitir datos de audio de un sitio de conferencia, y los datos de audio que satisfacen la condición previamente configurada son utilizados como los datos de audio del sitio de conferencia en cascada. En la realización de la presente invención, el sitio de conferencia que satisface la condición previamente configurada utilizado como un sitio de conferencia separado transmite datos de audio a través del canal de audio en cascada, y sitios de conferencia en la cascada de conferencia son ajustados con múltiples canales de audio en cascada, y luego, en la realización de la presente invención, datos de audio de cada sitio de conferencia pueden ser procesados de acuerdo con cada canal de audio en cascada, y la cantidad de canal de sonido de cada canal de audio en cascada puede ser un único canal de sonido, canales de sonido dobles, tres canales de sonido o incluso más canales de sonido, lo que no está limitado aquí.

505: Generar información de composición de canal de audio en cascada, donde la información de composición de canal de audio en cascada es información de cantidad de canales de audio en cascada establecida por el sitio de conferencia en cascada, de modo que se obtenga la cantidad de canales de audio en cascada ocupados por la corriente de código de audio recibida enviada por el sitio de conferencia en cascada.

506: Codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio.

Los datos de audio que satisfacen la condición previamente configurada en la operación anterior son codificados y utilizados como la corriente de código de audio.

5 507: Enviar por la corriente de código de audio al primer sitio de conferencia en cascada, y enviar la información de composición de canal de audio en cascada al primer sitio de conferencia en cascada.

10 Después de que se obtiene la corriente de código de audio que necesita ser enviada después de que se realice la codificación, la corriente de código de audio es enviada al primer sitio de conferencia en cascada. Debido a que, en la realización de la presente invención, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada es ajustado con múltiples canales de audio en cascada, y los datos de audio que satisfacen la condición previamente configurada de los canales de audio en cascada están siempre cambiando de forma constante, la información de composición del canal de audio en cascada necesita además se enviada al primer sitio de conferencia en cascada en la realización de la presente invención.

15 Puede verse a partir de la realización de la presente invención, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada de acuerdo con la realización de la presente invención selecciona los datos de audio que satisfacen la condición previamente configurada para enviar de acuerdo con diferentes canales de audio en cascada, respectivamente, de modo que cuando el ajuste de secuencia de orientación necesita ser realizado en los datos de audio, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada utilizado como un extremo de recepción puede ajustar directamente los datos de audio de forma individual.

20 Una realización de la presente invención se ha descrito a continuación con un ejemplo específico. Como se ha mostrado en la fig. 6, se ha descrito una estructura de sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención; que un canal de audio en cascada es ajustado y el canal de audio en cascada que incluye más de dos canales de sonido de audio es tomado como un ejemplo, con referencia a lo que se ha mostrado en la fig. 6:

25 Un sitio de conferencia en cascada incluye dos MCU, que son una MCU1 y una MCU2, donde la MCU1 está conectada a la MCU2.

30 La MCU1 está conectada a un sitio de conferencia ordinaria y a dos sitios de conferencia de telepresencia. La fig. 6 es un diagrama esquemático formado antes de ajustar una secuencia de corriente de código de audio. En la fig. 6, un sitio de conferencia ordinaria es T2, dos sitios de conferencia de telepresencia son T1 y T3, y los sitios de conferencia de telepresencia T1 y T3 tiene tres pantallas, respectivamente, que son T1L, T1C, y T1R, y T3L, T3C, y T3R, respectivamente; mientras tanto, la MCU2 utilizada como un sitio de conferencia en cascada de la MCU1 también está conectada a la MCU1, la MCU2 está conectada a dos sitios de conferencia ordinaria y a un sitio de conferencia de telepresencia. Como se ha mostrado en la fig. 6, dos sitios de conferencia ordinaria son T5 y T6, y un sitio de conferencia de telepresencia es T4 que incluye tres pantallas que son T4L, T4C, y T4R. Cada MCU soporta la reserva de datos de audio de tres sitios de conferencia de audio con las voces máximas, en particular, la MCU selecciona datos de audio de tres sitios de conferencia con las voces máximas a partir de todos los sitios de conferencia conectados (incluyendo el sitio de conferencia ordinaria, el sitio de conferencia de telepresencia, y el sitio de conferencia en cascada) para realizar la codificación.

40 La MCU1 puede recibir corrientes de código de audio enviadas por T1L, T1C, T1R, y T2, y T3L, T3C, y T3R, y la MCU2 puede recibir corrientes de código de audio enviadas por T4L, T4C, T4R, T5 y T6, y una corriente de código de audio del sitio de conferencia en cascada y una corriente de código de video del sitio de conferencia en cascada que son enviadas por la MCU1. Como se ha mostrado en la fig. 6, en la realización de la presente invención, corrientes de código de video del sitio de conferencia en cascada enviadas por la MCU1 a la MCU2 son T2, T1C, y T3R, y en esta realización, no se ha descrito un proceso de que la MCU1 envía datos de medios del sitio de conferencia en cascada a T1L, T1C, T1R, y T2, y T3L, T3C, y T3R, y sólo se ha descrito un proceso por el que la MCU1 envía la corriente de código de audio del sitio de conferencia en cascada a la MCU2.

50 La MCU1 descodifica las corrientes de código de audio, respectivamente, para obtener datos de audio del sitio de conferencia que corresponden a T1L, T1C, T1R, y T2, T3L, T3C, y T3R, los datos de audio del sitio de conferencia que corresponden a T1L, T1C, T1R, y T2, y T3L, T3C, y T3R son utilizados como datos de audio que han de ser seleccionados, y de acuerdo con el principio de reservar datos de audio de tres partes con las voces máximas, los datos de audio son filtrados a partir de los datos de la unidad que han de ser seleccionados. Se ha asumido que los datos de audio filtrados de acuerdo con los datos de cambios reservados de tres partes con las voces máximas son T1C, T2, y T3R, una fuente de video de la corriente de código de video del sitio de conferencia en cascada y una fuente de audio de los datos de audio presentados son completamente las mismas, en particular, asistentes de conferencia que corresponden a las corrientes de código de video del sitio de conferencia en cascada T2, T1C, y T3R y asistentes de conferencia que corresponden a los datos de audio presentados T1C, T2, y T3R son completamente los mismos, la secuencia de orientación de los datos de audio presentados puede ser ajustada correspondientemente de acuerdo con la secuencia de la corriente de código de video del sitio de conferencia en cascada, y después del ajuste, la orientación de

secuencia de los datos de audio presentados y la orientación de secuencia de la corriente de código de video del sitio de conferencia en cascada son las mismas, de modo que los datos de audio cuya secuencia es ajustada son utilizados como los datos de audio del sitio de conferencia en cascada y son codificados para obtener corrientes de código de audio T1C, T2, y T3R de la MCU1, y la MCU2 puede identificar las corrientes de código de audio.

5 Si los datos de audio filtrados de acuerdo con los datos de audio reservados de tres partes con las voces máximas no son T1C, T2, o T3R, en particular, la fuente de los datos de audio presentados y la fuente de la corriente de código de video del sitio de conferencia en cascada son diferentes, el sitio de conferencia donde los datos de audio presentados están ubicados es definido como un sitio de conferencia invisible, y los datos de audio que corresponden al sitio de conferencia invisible pueden ser ajustados de acuerdo con la política descrita en la operación 204 en la fig. 2.

10 Se ha descrito a continuación desde la perspectiva de que la MCU2 se utiliza como un extremo de recepción. En primer lugar, se recibe una corriente de código de audio enviada por la MCU1, se reciben corrientes de código de audio enviadas por los sitios de conferencia que no están en cascada T4, T5, y T6 conectados a la MCU2, y luego, se seleccionan los datos de audio que satisfacen la condición previamente configurada de los datos de audio que han de ser seleccionados; el proceso de seleccionar los datos de audio es el mismo que el proceso de procesamiento en la MCU1, lo que no se ha descrito aquí de nuevo reiteradamente. Finalmente, puede ajustarse la secuencia de orientación de los datos de audio seleccionados, y la política de ajuste específica está descrita en la realización en la fig. 2, lo que no se ha descrito aquí de nuevo reiteradamente. En la realización de la presente invención, el ajuste en una única orientación de datos de audio no influye en otro datos de audio, y la correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en la conferencia en cascada puede ser implementada, mejorando de este modo la experiencia de usuario de un participante.

Una realización de la presente invención está descrita a continuación con otro ejemplo específico. Como se ha mostrado en la fig. 7, se ha descrito una estructura de sitios de conferencia en cascada en una conferencia en cascada de acuerdo con otra realización de la presente invención, y en la realización de la presente invención, más de dos canales de audio en cascada son ajustados, en particular, una corriente de código de audio es transmitida en cada canal de audio en cascada, respectivamente, con referencia a lo que se ha mostrado en la fig. 7.

Un sitio de conferencia en cascada tiene dos MCU que son una MCU1 y una MCU2, donde la MCU1 está conectada a la MCU2, y cuatro canales de audio en cascada y cuatro canales de video en cascada están conectados entre la MCU1 y la MCU2. La MCU1 está conectada a los dos sitios de conferencia ordinaria y a los dos sitios de conferencia de telepresencia. Como se ha mostrado en la fig. 7, dos sitios de conferencia ordinaria son T2 y T7, y dos sitios de conferencia de telepresencia son T1 y T3; mientras tanto, la MCU2 utilizada como un sitio de conferencia en cascada de la MCU1 también está conectada a la MCU1, y la MCU2 está conectada a dos sitios de conferencia ordinaria y a un sitio de conferencia de telepresencia. Como se ha mostrado en la fig. 7, los dos sitios de conferencia ordinaria son T5 y T6. Las fuentes de video en cascada entre la MCU1 y la MCU2 son T2, T1C, T3R, y T7. Cada MCU soporta la reserva de datos de audio de cuatro sitios de conferencia de audio con las voces máximas, en particular, la MCU selecciona datos de audio de cuatro sitios de conferencia con las voces máximas de todos los sitios de conferencia conectados (incluyendo el sitio de conferencia ordinaria, el sitio de conferencia de telepresencia, y el sitio de conferencia en cascada).

En la realización de la presente invención, los canales de audio en cascada de la MCU1 y de la MCU2 son múltiples en número, lo que puede ser determinado específicamente de acuerdo con los requisitos de la conferencia en cascada. Mientras tanto, un canal de audio en cascada filtra de manera similar los datos de audio de acuerdo con el principio de reservar datos de audio de varias partes con las voces máximas. En la realización de la presente invención, los canales de audio en cascada son cuatro en número, y la cantidad de canal de sonido de cada canal de audio en cascada puede ser un único canal de sonido, canales de sonido dobles, tres canales de sonido o incluso más canales de sonido, lo que no está limitado aquí. Debido a que los canales de audio en cascada son múltiples en número, en la realización de la presente invención los datos de audio del sitio de conferencia de telepresencia pueden ser utilizados como datos de audio de un sitio de conferencia, pero la cantidad de canal de sonido de los canales de audio en cascada puede ser ajustada a canales de sonido dobles o tres canales de sonido, de modo que asegure que un canal de audio en cascada puede ser cargado con todo el audio de un sitio de conferencia en telepresencia.

La MCU1 puede recibir corrientes de código de audio enviadas por T1, T2, T3, y T7, y la MCU2 puede recibir corrientes de código de audio enviadas por T4, T5, y T6, y una corriente de código de audio del sitio de conferencia en cascada y una corriente de código de video del sitio de conferencia en cascada que son enviadas por la MCU1. Como se ha mostrado en la fig. 7, en la realización de la presente invención, corrientes de código de video del sitio de conferencia en cascada enviadas por la MCU1 a la MCU2 son T2, T1C, T3R, y T7, y en esta realización, no se ha descrito un proceso en el que la MCU1 envía datos de medios del sitio de conferencia en cascada a T1, T2, T3 y T7, y sólo se ha descrito un proceso en el que la MCU1 envía la corriente de código de audio del sitio de conferencia en cascada a la MCU2.

La MCU1 descodifica las corrientes de código de audio, respectivamente, para obtener datos de audio del sitio de conferencia que corresponden a T1, T2, T3 y T7, los datos de audio del sitio de conferencia que corresponden a T1, T2, T3 y T7 son utilizados como datos de audio que han de ser seleccionados, y de acuerdo con el principio de reservar datos de audio de cuatro sitios de conferencia con las voces máximas, los datos de audio T1, T2, T3 y T7 son filtrados a

partir de los datos de audio que han de ser seleccionados, y son utilizados como los datos de audio del sitio de conferencia en cascada. Y luego, las cuatro trayectorias de datos de audio son cargadas en los canales de audio en cascada correspondientes, respectivamente, y si los datos de audio presentados son datos de audio de un sitio de conferencia de telepresencia, los datos de audio presentados son cargados correspondientemente en un canal de audio en cascada de múltiples canales de sonido.

Después de que se obtienen datos de audio del sitio de conferencia en cascada, los datos de audio del sitio de conferencia en cascada son codificados para obtener una corriente de código de audio del sitio de conferencia en cascada; y luego, la corriente de código de audio del sitio de conferencia en cascada es enviada al sitio de conferencia en cascada MCU2, y una corriente de código de video del sitio de conferencia en cascada es enviada al sitio de conferencia en cascada MCU2.

Una parte con el máximo audio del canal en cascada cambia constantemente, y por lo tanto, se requiere enviar información de composición de canal de audio en cascada en una corriente de código a través del canal de audio en cascada, donde la corriente de código es enviada a la MCU2, y la información de composición de canal de audio en cascada es información de cantidad de canales de audio en cascada establecidos por el sitio de conferencia en cascada, de modo que obtenga la pluralidad de canales de audio en cascada ocupados por la corriente de código de audio recibida. Específicamente, la información de composición de canal de audio en cascada puede ser enviada de una manera de llenado de información RTP pero la presente invención no está limitada meramente a esa manera, que no está limitada aquí.

Después de que la MCU2 reciba datos de un sitio de conferencia con el próximo del canal de audio en cascada, además de datos de audio del sitio de conferencia conectados directamente al MCU2, la MCU2 es realmente equivalente para recibir datos independientes de cada sitio de conferencia T7, T1, T2, T3, T4, T5, y T6, y la orientación de la corriente de código de audio del sitio de conferencia correspondiente es ajustada y procesada de acuerdo con el caso de la corriente de código del video que necesita ser presentada en el sitio de conferencia en esta MCU, de modo que la secuencia de orientación de cada corriente de código de video del sitio de conferencia corresponde a la secuencia de orientación de la corriente de código de audio correspondiente una a una.

Debería observarse que, si el sitio de conferencia del canal de audio en cascada y un sitio de conferencia de video de un canal de video en cascada no son completamente los mismos, el procesamiento es realizado de acuerdo con la manera de procesamiento en la realización de la presente invención como se ha mostrado en la fig. 2, que no se ha descrito aquí de nuevo reiteradamente.

En el contenido anterior, se ha descrito el método para procesar sitios de conferencia en cascada en la conferencia en cascada en la realización de la presente invención, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada está descrito a continuación, y el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada previsto en la realización de la presente invención puede ser utilizado como una MCU.

Con referencia a la fig. 8, un ejemplo de un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención incluye:

una unidad de recepción 801, configurada para recibir una corriente de código de audio enviada por un sitio de conferencia, donde la corriente de código de audio enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada, y configurada además para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada;

una unidad de selección 802, configurada para seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados, donde los datos de audio que han de ser seleccionados incluyen: la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada; y

una unidad de ajuste de secuencia 803, configurada para crear una secuencia de orientación de los datos de audio seleccionados por la unidad de selección 802.

En la realización de la presente invención, debido a que la corriente de código de audio enviada por el sitio de conferencia en cascada indefinida por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada, cuando el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada necesita realizar un ajuste de secuencia de orientación en los datos de audio, la orientación de los datos de audio que necesita ser ajustada puede ser ajustada directamente de forma individual, en particular, el ajuste realizado por el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada en una única orientación de datos de audio no influye en otros datos de audio, y la correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en la conferencia en cascada puede ser incrementada, mejorando de este modo la experiencia del usuario de un participante.

Debería observarse que, en la realización mostrada en la fig. 8 de la presente invención, cuando la corriente de código

de audio enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de audio en cascada, la unidad de recepción 801 está además configurada para recibir información de composición de canal de audio en cascada enviada por el sitio de conferencia en cascada, donde la información de composición del canal de audio en cascada es información de cantidad de canales de audio en cascada establecida por el sitio de conferencia en cascada, de modo que obtenga la cantidad de canales de audio en cascada ocupados por la corriente de código de audio recibida.

El aparato para procesar sitios de conferencia en la conferencia en cascada puede incluir además: una unidad de descodificación, configurada para descodificar la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada.

En la realización anterior, se ha descrito que el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada recibe la corriente de código de audio del sitio de conferencia en cascada, y ajusta la secuencia de orientación de los datos de audio que satisfacen una condición previamente configurada. Se ha descrito a continuación desde la perspectiva de que el aparato para procesar sitios de conferencia en cascada en una conferencia en cascada envía una corriente de código de audio a un sitio de conferencia en cascada, y puede hacerse referencia a la fig. 9.

Una unidad de recepción 901 está configurada para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada y configurada además para recibir una corriente de código de audio enviada por segundo sitio de conferencia en cascada.

Una unidad de descodificación 902 está configurada para descodificar la corriente de código de audio recibida por la unidad de recepción 901.

Una unidad de selección 903 está configurada para seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados, donde los datos de audio que han de ser seleccionados incluyen específicamente un resultado de la descodificación de la unidad de descodificación 902.

Una unidad de procesamiento 904 está configurada para procesar los datos de audio seleccionados por la unidad de selección 903 en diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada para obtener datos de audio del sitio de conferencia en cascada, de modo que un primer sitio de conferencia en cascada pueda identificar datos de audio del sitio de conferencia en cascada.

Una unidad de codificación 905 está configurada para codificar un resultado del procesamiento de la unidad de procesamiento 904 para obtener una corriente de código de audio.

Una unidad de envío 906 está configurada para enviar la corriente de código de audio al primer sitio de conferencia en cascada.

Se puede ver a partir de la realización de la presente invención que, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada de la realización de la presente invención procesa los datos de audio que satisfacen la condición previamente configurada para diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada para obtener los datos de audio del sitio de conferencia en cascada, de modo que cuando el ajuste de secuencia de orientación necesita ser realizado en los datos de audio, el primer sitio de conferencia en cascada utilizado como un extremo de recepción puede ajustar directamente los datos de audio de forma individual.

Además, con referencia a la fig. 10, en una realización de la presente invención, sólo un canal de audio en cascada es ajustado, y el canal de audio en cascada incluye más de dos canales de sonido de audio, en particular, una corriente de código de audio es transmitida en cada canal de sonido de audio, respectivamente. Otro ejemplo de un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada de acuerdo con una realización de la presente invención incluye:

una unidad de recepción 1001, configurada para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada y configurada además para recibir una corriente de código de audio enviada por segundo sitio de conferencia en cascada;

una unidad de descodificación 1002, configurada para descodificar la corriente de código de audio recibida por la unidad de recepción 1001;

una unidad de selección 1003, configurada para seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados, donde los datos de audio que han de ser seleccionados incluyen específicamente un resultado de la descodificación de la unidad de descodificación 1002;

una unidad de procesamiento 1004, que incluye: un módulo de identificación del sitio de conferencia 10041, configurado para, si los datos de audio que satisfacen la condición previamente configurada son datos de audio que corresponden a una pantalla de un sitio de conferencia de telepresencia, utilizar los datos de audio que corresponden a

la pantalla del sitio de conferencia de telepresencia como datos de audio que corresponden a un sitio de conferencia separado; un módulo de mezcla de sonido 10042, configurado para, si los datos de audio que satisfacen la condición previamente configurada son datos de audio que corresponden a la posición de conferencia ordinaria y la cantidad de canal de sonido del sitio de conferencia ordinaria no es un único canal de sonido, mezclar los datos de audio del sitio de conferencia ordinaria en los únicos datos de audio de canal de sonido; y un módulo de correlación 10043, configurado para, de acuerdo con una secuencia de edificación de una corriente de código de video que necesita ser enviada, ajustar una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada de acuerdo con diferentes sitios de conferencia que ocupan canales de sonido de audio diferentes, respectivamente, y utilizar los datos de audio que satisfacen la condición previamente configurada y cuya secuencia de orientación es ajustada como datos de audio del sitio de conferencia en cascada;

una unidad de codificación 1005, configurada para codificar los datos de audio del sitio de conferencia en cascada obtenidos por unidad de procesamiento 1004 para obtener una corriente de código de audio; y

una unidad de envío 1006, configurada para enviar la corriente de código de audio al primer sitio de conferencia en cascada.

Debería observarse que, en la realización de la presente invención, cuando se incluyen el módulo de identificación del sitio de conferencia 10041 y el módulo de mezcla de sonido 10042, la unidad de procesamiento 1004 también puede no incluir el módulo de correlación 10043 pero incluye los siguientes módulos: un módulo de formación de secuencia, configurado para formar una secuencia de los datos de audio que satisfacen la condición previamente configurada de acuerdo con diferentes sitios de conferencia que ocupan canales de sonido de audio diferentes, y utilizar los datos de audio secuenciados que satisfacen la condición previamente configurada como datos de audio del sitio de conferencia en cascada; y un módulo de generación, configurado para generar una información de ubicación del sitio de conferencia de audio, donde la información de ubicación del sitio de conferencia de audio es información de secuenciación de ubicación de los datos de audio que satisfacen la condición previamente configurada. Mientras tanto, la unidad de envío 1006 en el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada en la realización de la presente invención está configurada además para enviar la información de localización del sitio de conferencia de audio al primer sitio de conferencia en cascada.

En la realización de la presente invención, también pueden ser ajustados más de dos canales de audio en cascada, en particular, una corriente de código de audio es transmitida en cada canal de audio en cascada, respectivamente, y en este caso, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada incluye en la siguiente unidad: una unidad de generación, configurada para generar información de composición del canal de audio en cascada, donde la información de composición del canal de audio en cascada es información de cantidad de canales de audio en cascada establecidos por el sitio de conferencia en cascada de modo que obtenga la cantidad de canales de audio en cascada ocupados por la corriente de código de audio recibida enviada por el sitio de conferencia en cascada. Mientras tanto, la unidad de envío 1006 en el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada en la realización de la presente invención está configurada además para enviar la información de composición de canal de audio en cascada al primer sitio de conferencia en cascada.

Puede verse a partir de la realización de la presente invención que, el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada procesa los datos de audio que satisfacen la condición previamente configurada para diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada por obtener los datos de audio del sitio de conferencia en cascada, de modo que cuando el ajuste de secuencia de orientación necesita ser realizado en los datos de audio, el primer sitio de conferencia en cascada utilizado como un extremo de recepción puede ajustar directamente los datos de audio de forma individual.

Una realización de la presente invención prevé además un sistema de conferencia en cascada que incluye:

el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada como se ha mostrado en la fig. 8;

y

el aparato para procesar sitios de conferencia en cascada en la conferencia en cascada como se ha mostrado en la fig. 9 o en la fig. 10.

Puede verse a partir de la realización de la presente invención que, debido a que la corriente de código de audio recibida por el sistema de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o canales de audio en cascada, cuando el sistema de conferencia en cascada necesita realizar el ajuste de secuencia de orientación en los datos de audio, los datos de audio que necesitan ser ajustados pueden ser ajustados directamente de formar individual, en particular, el ajuste realizado por el sistema de conferencia en cascada en una única orientación de datos de audio no influye sobre otros datos de audio, y la correspondencia uno a uno entre una orientación de imagen y una orientación de voz de cada sitio de conferencia en la conferencia en cascada puede ser implementada, mejorando de este modo la experiencia del usuario de un participante.



Debido a que el contenido tal como de procesos de ejecución e intercambio de información entre todos los módulos en el aparato y sistema anteriores está basado en un mismo concepto que las realizaciones del método de la presente invención, para el contenido específico, puede hacerse referencia a las descripciones en las realizaciones del método de la presente invención, lo que no se ha descrito aquí de nuevo reiteradamente.

- 5 Los expertos en la técnica deberían comprender que todas o una parte de las operaciones de los métodos en las realizaciones anteriores pueden ser implementadas por un programa que instruye un hardware relevante. El programa puede ser almacenado en un medio de almacenamiento legible por ordenador, y el medio de almacenamiento mencionado con anterioridad puede ser una memoria sólo de lectura, un disco magnético, o un disco compacto, y así sucesivamente.
- 10 El método, el aparato, del sistema para procesar los sitios de conferencia en cascada en la conferencia en cascada prevista en la presente invención son introducidos de forma detallada en lo anterior.

Expertos en la técnica pueden hacer variaciones en términos de maneras de implementación específica y marcos de aplicación de acuerdo con las ideas de las realizaciones de la presente invención.

15

**REIVINDICACIONES**

1. Un método para procesar sitios de conferencia en cascada en una conferencia en cascada, que comprende:

5 recibir (201) una corriente de código de audio enviada por un sitio de conferencia en cascada, en la que la corriente de código de audio enviada por el sitio de conferencia en cascada, es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada;

recibir (202) una corriente de código de audio enviada por un sitio de conferencia que no está en cascada;

10 seleccionar (203) datos de audio que satisfacen una condición previamente configurada a partir de datos de audio que han de ser seleccionados, en que los datos de audio que han de ser seleccionados comprenden: la corriente de código de audio recibida enviada por el sitio de conferencia en cascada y la corriente de código de audio recibida enviada por el sitio de conferencia que no está en cascada; y

ajustar (204) una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada;

en que la operación de

15 ajustar una secuencia de orientación de los datos de audio que satisface la condición previamente configurada comprende:

20 si sólo se presenta una pantalla de una fuente de video del sitio de conferencia que no está en cascada en una de las múltiples pantallas o en una imagen en una ubicación en una imagen múltiple, una secuencia de orientación de salida de datos de audio que corresponde a todas las pantallas de la fuente de video es una secuencia de orientación de presentación de la pantalla presentada en múltiples pantallas o una secuencia de orientación de la pantalla presentada en la imagen en la ubicación en la imagen múltiple;

25 si se presentan más de dos de las múltiples pantallas de la fuente de video del sitio de conferencia que no está en cascada, una secuencia de orientación de salida de datos de audio que corresponde a las más de dos pantallas presentadas de la fuente de video corresponde a una secuencia de orientación de las más de dos pantallas presentadas de la fuente de video una a una, y una secuencia de orientación de salida de datos de audio que corresponde a pantallas no presentadas de las múltiples pantallas de la fuente de video es mantenida igual que la de una de todas las pantallas presentadas de la fuente de video;

30 si se presenta de forma simultánea una pantalla de la fuente de video del sitio de conferencia que no está en cascada, en múltiples imágenes o en una imagen múltiple y una pantalla independiente, las prioridades de secuencias de orientación de salida de datos de audio que corresponde a la fuente de video desde la más alta a la más baja en una secuencia son: una orientación de la pantalla independiente, una orientación de una pantalla con una gran imagen secundaria, y una orientación de una pantalla presentada en prioridades intermedia, izquierda, y derecha.

35 2. El método para procesar sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 1, en el que el sitio de conferencia que no está en cascada comprende: un sitio de conferencia ordinaria y/o un sitio de conferencia de telepresencia.

40 3. El método para procesar los sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 1, en el que cuando la corriente de código de audio enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio, el método comprende además: recibir información de composición de canal de audio en cascada enviada por el sitio de conferencia en cascada, en el que la información de composición de canal de audio en cascada es información de cantidad de canales de audio en cascada establecida por el sitio de conferencia en cascada, de modo que se obtiene la cantidad de canales de audio en cascada ocupados por la corriente de código de audio recibida por el sitio de conferencia en cascada.

45 4. El método para procesar los sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 1, en el que la condición previamente configurada es reservar datos de audio de varias partes con las voces máximas en el sitio de conferencia en cascada y en el sitio de conferencia que no está en cascada.

5. Un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada, que comprende:

50 una unidad de recepción (801), configurada para recibir una corriente de código de audio enviada por un sitio de conferencia en cascada, en que la corriente de código de audio enviada por el sitio de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de los canales de audio en cascada, y

la unidad de recepción está configurada además para recibir una corriente de código de audio enviada por un

sitio de conferencia que no está en cascada;

5 una unidad de selección (802), configurada para seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados, en que los datos de audio que han de ser seleccionados comprenden: la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada; y

una unidad de ajuste de secuencia (803), configurada para ajustar una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada;

en que el ajuste de

10 una secuencia de orientación de los datos de audio que satisfacen la condición previamente configurada comprende:

15 si sólo se presenta una pantalla de una fuente de video del sitio de conferencia que no está en cascada en una de las múltiples pantallas o en una imagen en una ubicación en una imagen múltiple, una secuencia de orientación de salida de datos de audio que corresponde a todas las pantallas de la fuente de video es una secuencia de orientación de presentación de la pantalla presentada en múltiples pantallas o una secuencia de orientación de la pantalla presentada en la imagen en la ubicación en la imagen múltiple;

20 si se presentan más de dos de las múltiples pantallas de la fuente de video del sitio de conferencia que no está en cascada, una secuencia de orientación de salida de datos de audio que corresponde a las más de dos pantallas presentadas de la fuente de video corresponde a una secuencia de orientación de las más de dos pantallas presentadas de la fuente de video una a una, y una secuencia de orientación de salida de datos de audio que corresponde a pantallas no presentadas de las múltiples pantallas de la fuente de video es mantenida igual que la de una de todas las pantallas presentadas de la fuente de video; y

25 si se presenta de forma simultánea una pantalla de la fuente de video del sitio de conferencia que no está en cascada, en múltiples imágenes o en una imagen múltiple y una pantalla independiente, las prioridades de secuencias de orientación de salida de datos de audio que corresponde a la fuente de video desde la más alta a la más baja en una secuencia son: una orientación de la pantalla independiente, una orientación de una pantalla con una gran imagen secundaria, y una orientación de una pantalla presentada en las prioridades intermedia, izquierda, y derecha.

30 6. El aparato para procesar los sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 5, en que cuando la corriente de código de audio enviada por el sitios de conferencia en cascada es enviada a diferentes sitios de conferencia a través de diferente canales de audio en cascada,

la unidad de recepción (801) está configurada además para recibir una información de composición de canal de audio en cascada recibida enviada por el sitio de conferencia en cascada, en que la información de composición de canal de audio en cascada es información de cantidad de canales de audio en cascada establecida por el sitio de conferencia en cascada, de modo que se obtiene la cantidad de canales de audio en cascada ocupados por la corriente de código de audio recibida.

35 7. El aparato para procesar los sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 5, que comprende además:

una unidad de decodificación (902), configurada para decodificar la corriente de código de audio enviada por el sitio de conferencia en cascada y la corriente de código de audio enviada por el sitio de conferencia que no está en cascada.

40 8. Un sistema de conferencia en cascada, que comprende:

el aparato para procesar los sitios de conferencia en cascada en la conferencia en cascada según la reivindicación 5 ó 7;

y

un aparato para procesar sitios de conferencia en cascada en una conferencia en cascada, que comprende:

45 una unidad de recepción (901), configurada para recibir una corriente de código de audio enviada por un sitio de conferencia que no está en cascada;

50 una unidad de selección (903), configurada para seleccionar datos de audio que satisfacen una condición previamente configurada de los datos de audio que han de ser seleccionados, en que los datos de audio que han de ser seleccionados comprenden al menos la corriente de código de audio enviada por el sitio de conferencia que no está en cascada;.

una unidad de procesamiento (904), configurada para procesar los datos de audio que satisfacen la condición previamente configurada para diferentes sitios de conferencia a través de diferentes canales de sonido de audio o de canales de audio en cascada para obtener datos de audio del sitio de conferencia en cascada, de modo que un primer sitio de conferencia en cascada puede identificar los datos de audio del sitio de conferencia en cascada;

5 una unidad de codificación (905), configurada para codificar los datos de audio del sitio de conferencia en cascada para obtener una corriente de código de audio; y

una unidad de envío (906), configurada para enviar la corriente de código de audio al primer sitio de conferencia en cascada.

10

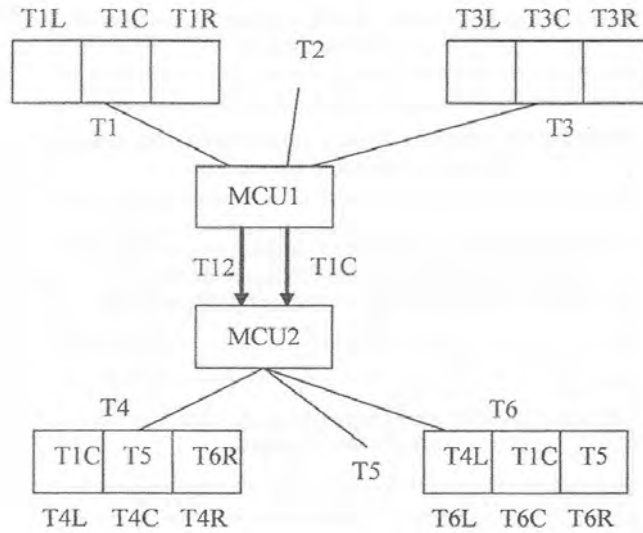


FIG. 1

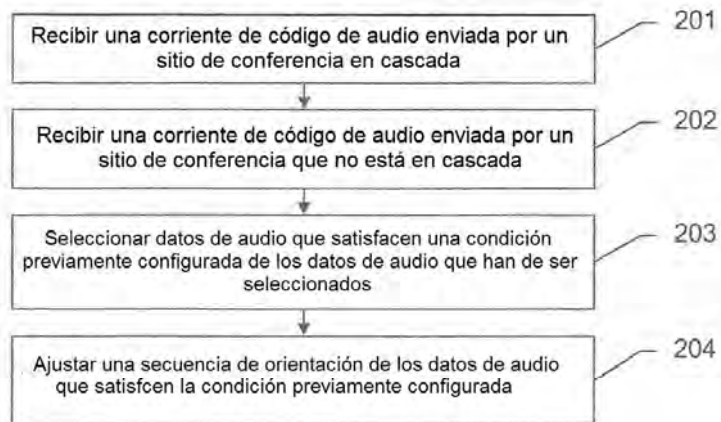


FIG. 2

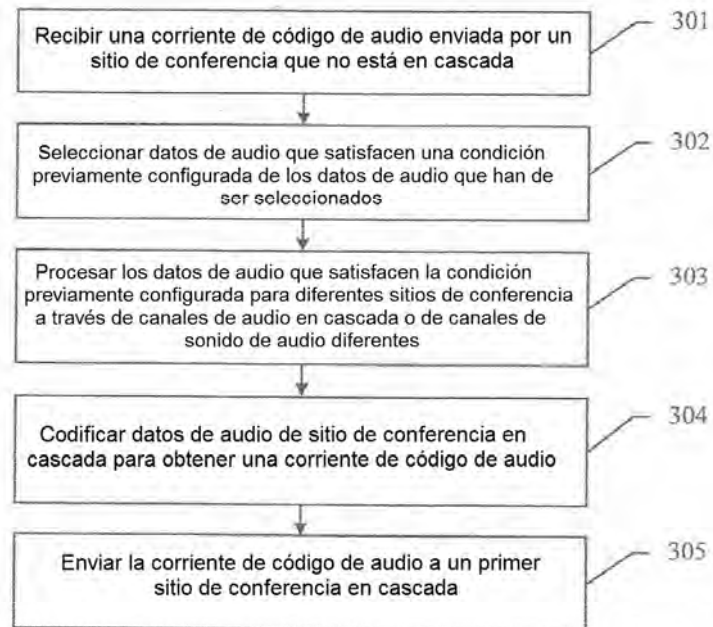


FIG. 3

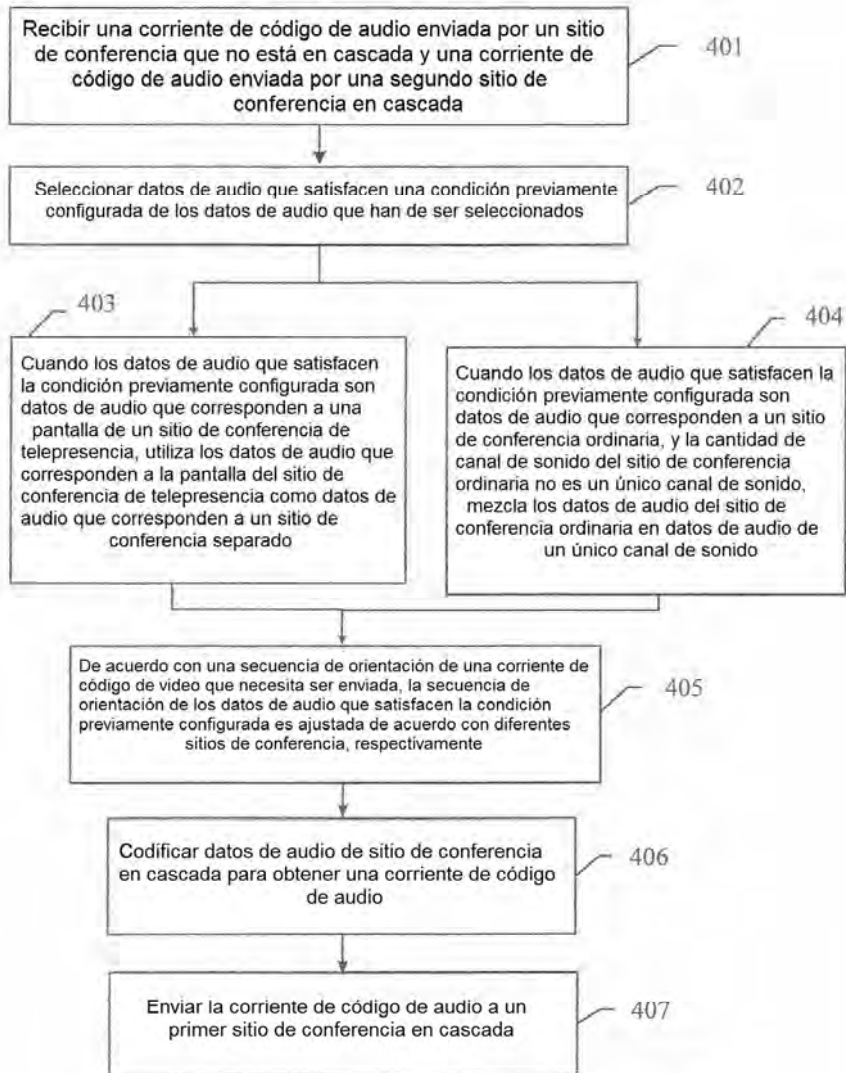


FIG. 4

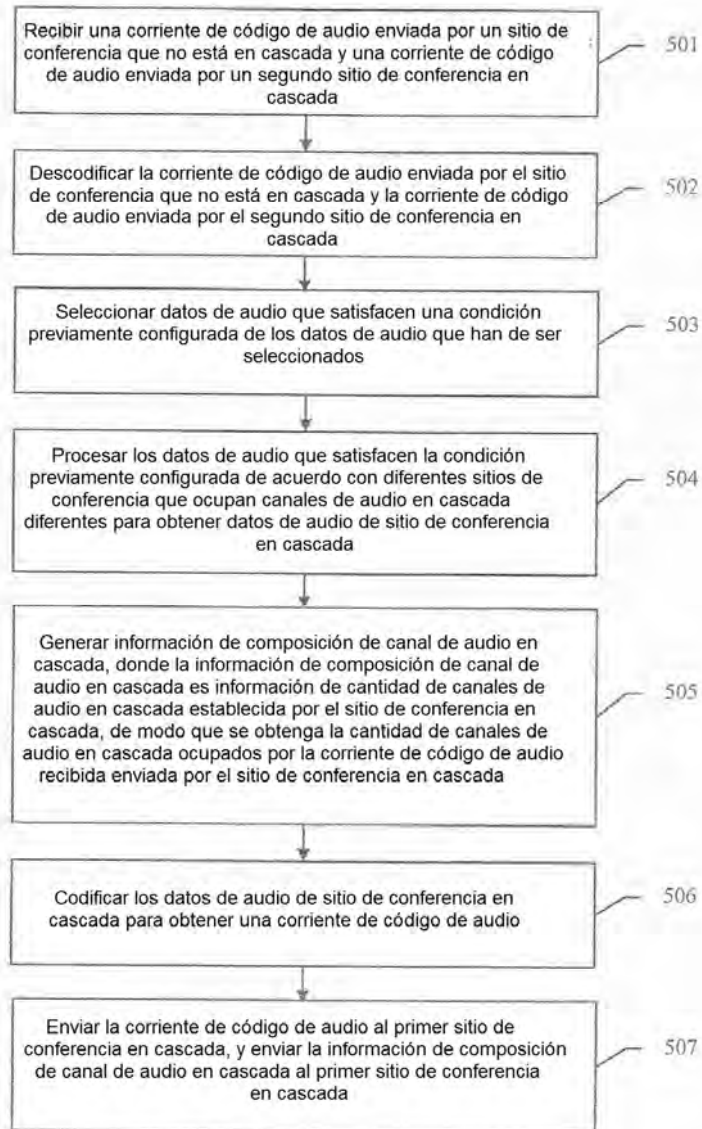


FIG. 5



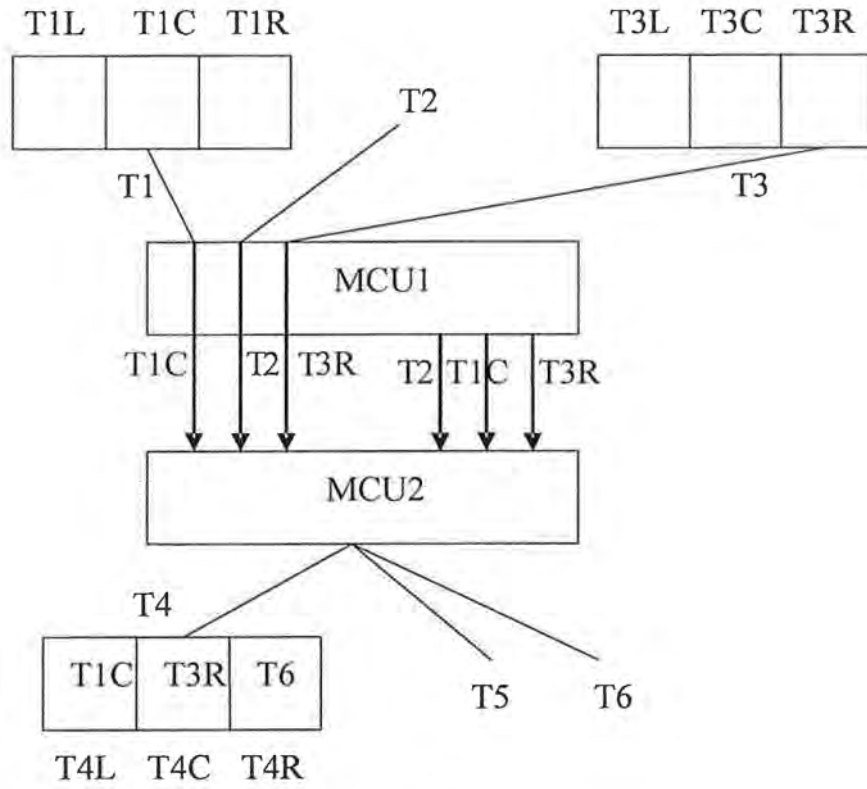


FIG. 6

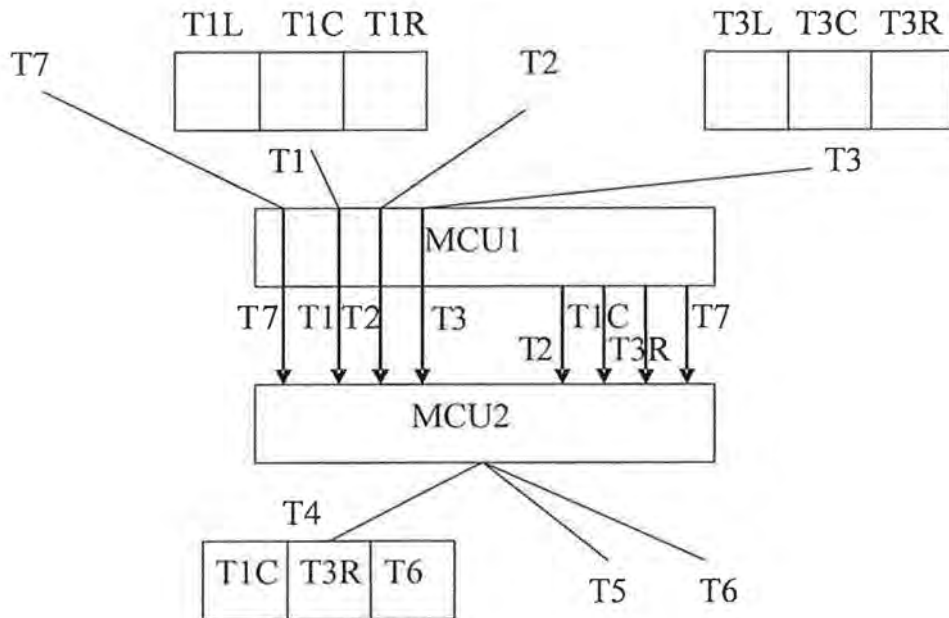


FIG. 7



FIG. 8

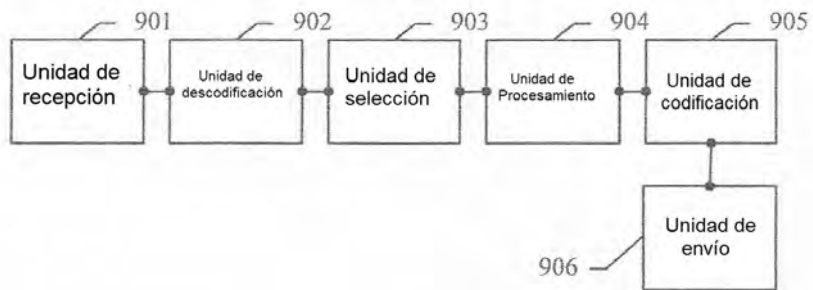


FIG. 9

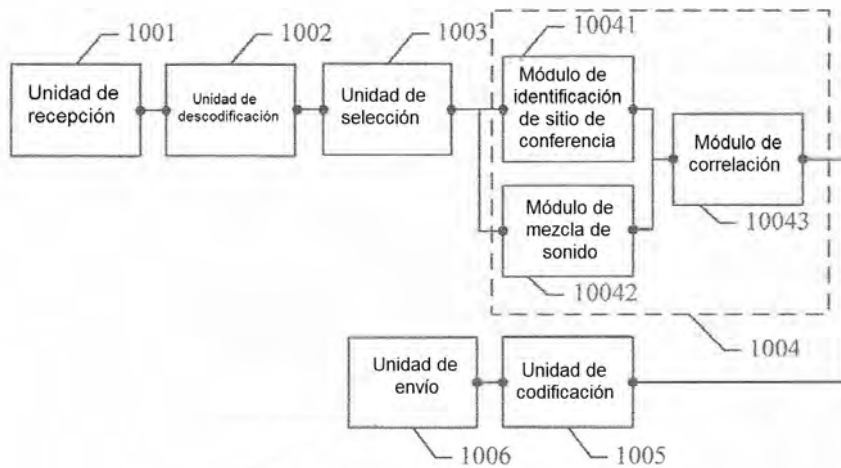


FIG. 10