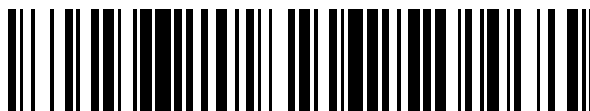


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 666 554**

51 Int. Cl.:

H04L 29/06 (2006.01)

H04L 29/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.04.2014** **E 16168880 (9)**

97 Fecha y número de publicación de la concesión europea: **07.03.2018** **EP 3094063**

54 Título: **Modificación interactiva de la distancia focal en una videoconferencia**

30 Prioridad:

26.04.2013 US 201361816662 P
17.12.2013 US 201314109121

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
07.05.2018

73 Titular/es:

INTEL IP CORPORATION (100.0%)
2200 Mission College Boulevard
Santa Clara, CA 95054, US

72 Inventor/es:

OYMAN, OZGUR

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 666 554 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

Modificación interactiva de la distancia focal en una videoconferencia

5 Campo técnico

Las formas de realización pertenecen a operaciones y comunicaciones realizadas mediante dispositivos electrónicos en redes inalámbricas. Algunas formas de realización se refieren a sesiones de conversación en vídeo que se realizan a través de redes inalámbricas.

10 Antecedentes

En aplicaciones de videoconferencia típicas, dos usuarios reciben imágenes de vídeo en directo de cada uno de ellos. En muchos casos, un usuario puede desear ampliar una característica particular que aparece en la imagen de vídeo en directo del otro usuario. Por consiguiente, existe la necesidad de modificar de manera interactiva la distancia focal en una aplicación de videoconferencia.

El documento US 2006/0215752 A1 se refiere a técnicas de procesamiento de una región de interés (ROI) para aplicaciones de videotelefonía (VT). En el documento US 2006/0215752 A1, un dispositivo receptor define información ROI para la información de vídeo transmitida por un dispositivo emisor, es decir, información de vídeo de extremo lejano, y transmite la información ROI al dispositivo emisor. Usando la información ROI transmitida por el dispositivo receptor, el dispositivo emisor aplica una codificación preferencial a una ROI en una escena de vídeo capturada y transmite toda la escena de vídeo codificada al dispositivo receptor.

El documento US 2010/0118111 A1 se refiere a un sistema para indicar operaciones de control de cámara a una parte remota durante una sesión de videoconferencia. Una interfaz está dispuesta en una parte receptora (que recibe un flujo de vídeo), que permite a la parte receptora introducir cualquier indicación deseada de control de cámara que se enviará a una parte emisora (que envía un flujo de vídeo). Las indicaciones de control de cámara recibidas por la parte emisora se renderizan y convierten en señales visuales o de audio, vibraciones, etc., que se muestran a uno o más participantes de una sesión de videoconferencia, tal como la parte emisora, invitando así a la parte emisora a que controle su cámara de manera correspondiente. Las operaciones de control de cámara incluyen visión panorámica, inclinación, modificación de la distancia focal, movimiento de la cámara en una dirección dada, etc.

35 Resumen

La invención está definida por la reivindicación independiente. Formas de realización ventajosas están sujetas a las reivindicaciones dependientes.

40 Breve descripción de los dibujos

La FIG. 1 ilustra un ejemplo de una arquitectura para una aplicación de videoconferencia.

La FIG. 2 ilustra un ejemplo de un esquema de señalización para una aplicación de videoconferencia.

La FIG. 3 ilustra un diagrama de flujo de un ejemplo de un procedimiento de funcionamiento para una aplicación de videoconferencia.

La FIG. 4 ilustra un ejemplo de un dispositivo cliente móvil en el que pueden implantarse las configuraciones y las técnicas descritas en el presente documento.

La FIG. 5 ilustra un sistema informático de ejemplo que puede usarse como una plataforma informática para los dispositivos informáticos o de interconexión descritos en el presente documento.

55 Descripción detallada

La siguiente descripción y los dibujos ilustran con suficiente claridad formas de realización específicas para permitir que los expertos en la técnica las lleven a la práctica. Otras formas de realización pueden incorporar cambios estructurales, lógicos, eléctricos, de proceso y otro tipo de cambios. Partes y características de algunas formas de realización pueden estar incluidas en, o sustituirse por, partes y características de otras formas de realización. Las formas de realización expuestas en las reivindicaciones engloban todas las equivalencias disponibles de esas reivindicaciones.

En una aplicación de videoconferencia, un usuario que ve una escena puede modificar el tamaño de la distancia focal seleccionando un área de la escena y después conmutar dinámicamente una señal de vídeo de la escena al área seleccionada de la escena. El hardware y el software de la aplicación de videoconferencia pueden limitar el

vídeo transmitido al usuario para que incluya solamente el área seleccionada de la escena. Si se transmite solamente el área seleccionada de la escena y se excluye de la transmisión el área no seleccionada de la escena, puede usarse de manera más eficaz el ancho de banda disponible de la aplicación de videoconferencia.

5 La FIG. 1 ilustra un ejemplo de una arquitectura para una aplicación de videoconferencia 100. La aplicación de videoconferencia 100 lleva a cabo una videoconferencia bidireccional entre un usuario local 102 y un usuario remoto 104.

10 El usuario local 102 ejecuta la aplicación de videoconferencia 100 en un equipo de usuario local (UE local) 106, que puede incluir un ordenador portátil, un teléfono celular u otro dispositivo adecuado. El UE local 106 incluye una cámara, que captura vídeo del usuario local 102 durante el funcionamiento, y una pantalla, que muestra vídeo del usuario remoto 104 al usuario local 102 durante el funcionamiento.

15 El usuario remoto 104 ejecuta la aplicación de videoconferencia 100 en un equipo de usuario remoto (UE remoto) 108. El UE remoto 108 también incluye una cámara, que captura vídeo del usuario remoto 104 durante el funcionamiento, y una pantalla, que muestra vídeo del usuario local 102 al usuario remoto 104 durante el funcionamiento. Tanto el UE local 106 como el UE remoto 108 incluyen un sistema de circuitos configurado para realizar una videoconferencia bidireccional a través de la aplicación de videoconferencia 100.

20 El usuario local 102 puede ampliar una característica o ubicación particular en el campo de visión de la cámara remota. En el UE local 106, el usuario local 102 puede definir una región de interés (ROI) 110 en el campo de visión del UE remoto 108. En el ejemplo de la FIG. 1, el campo de visión de la cámara remota incluye la cabeza y el torso superior del usuario remoto 104, mientras que la ROI 110, definida por el usuario local 102, incluye solamente la cabeza del usuario remoto 104. Esto es simplemente un ejemplo; el usuario local 102 puede definir también otras
25 ROI adecuadas dentro del campo de visión de la cámara remota.

Una vez que se ha definido una ROI 110, la aplicación de videoconferencia 100 puede indicar al UE remoto 108 que capture vídeo dentro de la ROI 110. En algunos ejemplos, el UE remoto 108 captura vídeo que incluye solamente la ROI 110 y excluye regiones fuera de la ROI 110, y transmite el vídeo capturado al UE local 106. En otros ejemplos,
30 el UE remoto 108 captura vídeo que incluye el campo de visión total de la cámara remota 108, pero transmite un vídeo que incluye solamente la ROI 110 y excluye regiones fuera de la ROI.

El UE local 106 recibe el vídeo desde el UE remoto 108, donde el vídeo recibido incluye regiones dentro la ROI 110 y excluye regiones fuera de la ROI 110. El UE local 106 muestra el vídeo recibido en la pantalla local.
35

En algunos ejemplos, el usuario local 102 puede definir una ROI 110 con un tamaño y una ubicación arbitrarios dentro del campo de visión de la cámara remota. En algunos ejemplos, el UE remoto 108 permanece estacionario cuando se define la ROI 110, de manera que seleccionar una ROI 110 no mueve o cambia el campo de visión de la cámara remota. En algunos ejemplos, el usuario local 102 puede seleccionar una nueva ROI 110 a voluntad. En
40 algunos ejemplos, el usuario remoto 104 también puede seleccionar una ROI análoga para disminuir la distancia focal en el usuario local 102.

En el ejemplo de la FIG. 1, la aplicación de videoconferencia 100 funciona a través de servicios de telefonía multimedia basados en 3GPP, que conectan el UE local 106 y el UE remoto 108 entre sí y a la red telefónica. Esto
45 es simplemente un ejemplo; también pueden usarse otras redes adecuadas.

El UE local 106 se conecta a la red central a través de una red de acceso radioeléctrico 112, un nodo servidor de soporte de GPRS (SGSN) 114 y/o un nodo pasarela de soporte GPRS (GGSN) 116. El UE local 106 envía y recibe datos a través de una función apoderada de control de sesión de llamadas (P-CSCF) 118. La P-CSCF envía y recibe
50 datos con una función servidora de control de sesión de llamadas (S-CSCF) 120. En algunos ejemplos, la S-CSCF 120 puede enviar y recibir datos desde un servidor de aplicaciones (AS) 122, que puede proporcionar servicios complementarios, tales como interrupción/reanudación de llamadas, reenvío de llamadas, llamadas entre varios participantes, así como otros servicios. La S-CSCF 120 puede enviar y recibir datos de otras partes de la red central. El UE remoto 108 se comunica de manera similar con su propia red de acceso radioeléctrico 124, SGSN 126, GGSN
55 128, P-CSCF 130, S-CSCF 132 y AS 134. Cada S-CSCF 120, 132 puede comunicarse con una CSCF de sondeo (I-CSCF) 134. La I-CSCF 134 puede leer y escribir en un servidor de abonado local (HSS) 136 y/o una función de ubicación de abonado (SLF) 136.

El documento de la especificación 3GPP TS 26.114 describe los servicios de telefonía multimedia a través de IMS
60 (MTSI) que permite suministrar servicios de conversación multimedia avanzados y contenido a través de redes basadas en IMS. Esto incluye la especificación de la gestión e interacción de información multimedia, que incluye control multimedia, códecs multimedia, así como el transporte de datos multimedia y de control. Una llamada MTSI usa mecanismos de función de control de sesión de llamada (CSCF) para encaminar señalización de plano de control entre los UE implicados en la llamada. El terminal de UE transmisor basado en MTSI captura y graba vídeo, y
65 lo transfiere al terminal de UE receptor basado en MTSI a través de la red 3GPP, que descodifica y renderiza el vídeo. En MTSI, el protocolo de inicio de sesión (SIP) sirve como el protocolo de control de capa de aplicación para

establecer, modificar y finalizar sesiones de conversión multimedia, tales como videoconferencias, llamadas telefónicas por Internet y de otro tipo. La señalización basada en el protocolo de descripción de sesión (SDP) entre el terminal emisor y el terminal receptor permite consideraciones de ofrecimiento/respuesta en la negociación de capacidad relacionada con datos multimedia, incluidos códecs, velocidades binarias, resoluciones y otros. El transporte de medios en MTSI está basado en el protocolo de transporte en tiempo real (RTP) (especificado por IETF RFC 3550) a través de UDP/IP.

Las resoluciones de los dispositivos de captura y, por lo tanto, de los vídeos comprimidos, están aumentando rápidamente. Por ejemplo, mediante el uso de la norma de codificación de vídeo de alta eficacia (HEVC), puede transportarse y almacenarse contenido 4K como parte de un producto operativo. En la actualidad hay cámaras con una resolución 4k por 2k. Se han emitido flujos continuos de vídeo en directo con resoluciones de 8k por 4k. Es probable que las resoluciones aumenten en el futuro, en lo que respecta al número de píxeles. En la actualidad, con estos contenidos de resolución tan elevada se permiten nuevos usos del flujo continuo de vídeo, como características de modificación interactiva de la distancia focal.

Las características de modificación interactiva de la distancia focal permite la adaptación dinámica de vídeo en lo que respecta al ancho de banda, la resolución espacial, la orientación, etc. Además, las características de modificación interactiva de la distancia focal pueden conmutar dinámicamente a un área seleccionada por el usuario en el vídeo que se está transmitiendo en flujo continuo, y pueden optimizar las codificaciones con este fin.

El dispositivo y el procedimiento dados a conocer en el presente documento proponen nuevas características de gestión de datos multimedia y de gestión de sesiones, dirigidas a servicios de conversación por vídeo para mejorar las resoluciones ofrecidas al usuario final cuando se usan características de modificación interactiva de la distancia focal. En particular, para abordar el problema mencionado anteriormente, se propone la capacidad de señalar la información de la ROI del usuario desde el terminal receptor hasta el terminal emisor, por ejemplo como parte de la señalización de realimentación RTCP o de la extensión de cabecera RTP. Esta característica, denominada modificación de distancia focal basada en ROI (ROI-Zoom) puede proporcionar una mejor calidad de imagen para la región seleccionada que con una modificación gráfica simple de la distancia focal, ya que el dispositivo emisor puede usar en este caso todo el ancho de banda disponible para codificar y transmitir la ROI, lo que puede ofrecer por tanto mayores velocidades binarias y una mejor calidad al terminal receptor.

La FIG. 2 ilustra un ejemplo de un esquema de señalización para una aplicación de videoconferencia, tal como la aplicación de videoconferencia 100 de la FIG. 1. Un cliente emisor 202 envía un ofrecimiento, a través de un protocolo de descripción (SDP), a un cliente receptor 204 indicando un flujo continuo basado en ROI. El cliente receptor 204 contesta, a través de SDP, y acepta el flujo continuo basado en ROI. El cliente receptor 204 envía la información ROI al cliente emisor 202, por ejemplo en mensajes de realimentación RTCP o en una extensión de cabecera RTP para flujos de vídeo en sentido inverso. El cliente emisor 202 codifica y transmite vídeo basándose en la ROI indicada, indicando posiblemente una ROI en la extensión de cabecera RTP para flujos de vídeo en sentido directo. Esto es simplemente un ejemplo; también pueden usarse otros esquemas de señalización adecuados.

La señalización basada en SDP entre el terminal emisor y el terminal receptor permiten consideraciones de ofrecimiento/respuesta en la negociación de capacidad relacionada con datos multimedia para ROI-Zoom, por ejemplo, 1) capacidad para indicar la ROI en información de realimentación RTCP y/o mensajes de extensión de cabecera RTP y 2) capacidad de entender/procesar información ROI y codificar vídeo de manera correspondiente.

Un cliente MTSI que admite la característica ROI-Zoom puede ofrecer ROI-Zoom en SDP para todos los flujos multimedia que contienen vídeo. La característica ROI-Zoom puede ofrecerse incluyendo el atributo `a=extmap`, que indica el URN ROI-Zoom en el alcance de línea multimedia pertinente. Por ejemplo, el URN ROI-Zoom puede fijarse como:

`urn:3gpp:roi-zoom.`

Un ejemplo de una línea multimedia que incluye este URN es:

`a=extmap:7 urn:3gpp:roi-zoom`

En el ejemplo anterior de una línea multimedia, el número 7 puede sustituirse por cualquier número entre 1 y 14, ambos inclusive.

La capacidad ROI-Zoom puede admitirse de manera bidireccional o unidireccional, dependiendo de cómo los clientes negocien para admitir la característica durante las negociaciones de la capacidad SDP. En lo que respecta a terminales con capacidad asimétrica (por ejemplo, la capacidad de procesar información ROI pero no de detectar/señalar información ROI) pueden usarse los atributos *sendonly* y *recvonly*. Los terminales deben expresar su capacidad con suficiente claridad en cada sentido de manera que las señales solo se envíen en cada sentido en la medida que expresen información útil y puedan procesarse por el receptor. La característica ROI-Zoom incluye la señalización de la ROI actual del usuario receptor al emisor de manera que el UE emisor pueda codificar y transmitir

de manera óptima el vídeo capturado. Cuando la característica ROI-Zoom se negocia con éxito, puede señalizarse por el cliente MTSI. La señalización de la ROI-Zoom puede usar la notificación de realimentación RTCP o extensiones de cabecera RTP.

5 Cuando se usan notificaciones de realimentación RTCP, el terminal receptor incluye la información ROI actual del usuario en las notificaciones RTCP que está transmitiendo al terminal emisor. Cuando se usan extensiones de cabecera RTP, el terminal receptor incluye la información ROI actual del usuario en los paquetes RTP que está transmitiendo al terminal emisor. Estos paquetes RTP transportan flujos de vídeo en sentido inverso, que pueden usarse en comunicaciones de vídeo bidireccionales en MTSI.

10 Estas diferentes capacidades de señalización ROI pueden indicarse en el ofrecimiento SDP a través de diferentes URN. En algunos ejemplos, puede elegirse uno de estos y usar un único URN en las negociaciones de capacidad SDP. Puesto que no siempre se usa RTCP en las videollamadas y puede desactivarse (esto también forma parte de la negociación SDP), puede ser más seguro entregar la información ROI como parte de las extensiones de cabecera RTP de los flujos de vídeo de sentido inverso para permitir en mayor medida la característica ROI-Zoom.

15 Pueden usarse dos atributos para definir la ROI que va a incluirse en notificaciones de realimentación RTCP o extensiones de cabecera RTP. Un atributo es la posición, que puede especificar una esquina superior izquierda del área cubierta en el contenido original, o puede especificar otra ubicación adecuada. El otro atributo es el tamaño, que puede especificar el tamaño del área cubierta en el contenido original en valores relativos. Un ejemplo de la posición puede ser "0:0", y un ejemplo del tamaño puede ser "0.33:0.33". También pueden usarse otros valores adecuados.

20 La presencia de la característica ROI-Zoom no debe influir en las resoluciones negociadas (basadas en el atributo *imageattr* de SDP) entre el terminal emisor y el terminal receptor. La única diferencia es que el terminal emisor debería codificar solamente la ROI con la resolución negociada en lugar de toda la imagen capturada, y esto debería dar lugar a una mayor resolución global y a una mejor experiencia de usuario que haciendo que el terminal receptor disminuya la distancia focal en la ROI y descarte el resto de la imagen. Una posibilidad es que el terminal emisor pueda indicar la ROI como parte de la extensión de cabecera RTP en paquetes RTP para sus flujos de vídeo en sentido directo.

La Tabla 1 incluye un ejemplo de un ofrecimiento SDP que indica una característica ROI-Zoom.

Tabla 1

```
m=video 49154 RTP/AVP 99
a=tcap:1 RTP/AVPF
a=pcfg:1 t=1
b=AS:315
b=RS:0
b=RR:2500
a=rtmap:99 H264/90000
a=fmtp:99 packetization-mode=0 ; profile-level-id=42e00c; \
sprop-parameter-sets=JOLgDJWgUH6Af1A=, KM46gA= [0047] a=imageattr : 99 send
[x=320, y=240] [x=240, y=320] recv [x=320, y=240] [x=240, y=320]
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb : * ccm fir
a=rtcp-fb:* ccm tmmbr
a=extmap:4 urn : 3gpp : roi-zoom
```

35

La Tabla 2 incluye un ejemplo de una respuesta SDP que acepta la característica ROI-Zoom.

Tabla 2

```
m=video 49154 RTP/AVPF 99
a=acfg:1 t=1
b=AS:315
b=RS:0
```

```

b=RR:2500
a=rtpmap:99 H264/90000
a=fmtp:99 packetization-mode=0 ; profile-level-id=42e00c; \
sprop-parameter-sets=JOLgDJWgUH6Af1A=, KM46gA=
a=imageattr : 99 send [x=320, y=240] [x=20, =320] recv [x=320, y=240] [x=240, y=320]
a=rtcp-fb:* trr-int 5000
a=rtcp-fb:* nack
a=rtcp-fb:* nack pli
a=rtcp-fb:* ccm fir
a=rtcp-fb:* ccm tmmbr
a=extmap:4 urn : 3gp : roi-zoom

```

Aunque los ejemplos anteriores de conexiones de redes inalámbricas se han proporcionado haciendo referencia específica a las normas de comunicación 3GPP LTE/LTE-A, IEEE 802.11 y Bluetooth, debe entenderse que pueden usarse otros diversos protocolos y normas WWAN, WLAN y WPAN en relación con las técnicas descritas en el presente documento. Estas normas incluyen, pero no están limitadas a, otras normas de las familias de normas 3GPP (por ejemplo, HSPA+, UMTS), IEEE 802.16 (por ejemplo, 802.16p), o Bluetooth (por ejemplo, Bluetooth 4.0 o normas similares definidas por el Grupo de Interés Especial de Bluetooth). Otras configuraciones de red aplicables pueden incluirse dentro del alcance de las redes de comunicación descritas en el presente documento. Debe entenderse que las comunicaciones en tales redes de comunicaciones pueden facilitarse usando cualquier número de redes de área personal, LAN y WAN, usando cualquier combinación de medios de transmisión cableados o inalámbricos.

Las formas de realización descritas anteriormente pueden implementarse en uno de o en una combinación de hardware, firmware y software. Varios procedimientos o técnicas, o determinados aspectos o partes de los mismos, pueden adoptar la forma de código de programa (es decir, instrucciones) almacenadas en medios tangibles, tales como memoria flash, discos duros, dispositivos de almacenamiento portátiles, memoria de solo lectura (ROM), memoria de acceso aleatorio (RAM), dispositivos de memoria de semiconductor (por ejemplo, memoria de solo lectura eléctricamente programable (EPROM), memoria de solo lectura eléctricamente borrrable y programable (EEPROM)), medios de almacenamiento de disco magnético, medios de almacenamiento óptico y cualquier otro medio de almacenamiento o dispositivo de almacenamiento legible por máquina en el que, cuando el código de programa se carga en y se ejecuta por una máquina, tal como un ordenador o un dispositivo de conexión en red, la máquina pasa a ser un aparato para llevar a la práctica las diversas técnicas.

Un medio de almacenamiento legible por máquina u otro dispositivo de almacenamiento puede incluir cualquier mecanismo no transitorio para almacenar información de forma legible por una máquina (por ejemplo, un ordenador). En el caso de código de programa que se ejecuta en ordenadores programables, el dispositivo informático puede incluir un procesador, un medio de almacenamiento legible por el procesador (incluyendo elementos de memoria y/o de almacenamiento volátiles y no volátiles), al menos un dispositivo de entrada y al menos un dispositivo de salida. Uno o más programas que pueden implementar o utilizar las diversas técnicas descritas en el presente documento pueden usar una interfaz de programación de aplicaciones (API), controles reutilizables, etc. Tales programas pueden implementarse en un lenguaje de programación procedural u orientado a objetos de alto nivel para comunicarse con un sistema informático. Sin embargo, el/los programa(s) puede(n) implementarse en lenguaje máquina o ensamblador, si se desea. En cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado, y combinarse con implementaciones en hardware. La FIG. 4 ilustra un ejemplo de un dispositivo móvil 400. El dispositivo móvil 400 puede ser un equipo de usuario (UE), una estación móvil (MS), un dispositivo inalámbrico móvil, un dispositivo de comunicaciones móviles, una tableta, un microteléfono u otro tipo de dispositivo informático inalámbrico móvil. El dispositivo móvil 400 puede incluir una o más antenas 408 dentro de un alojamiento 402 que están configuradas para comunicarse con un punto activo, una estación base (BS), un NodoB evolucionado (eNodoB) u otro tipo de punto de acceso WLAN o WWAN. El dispositivo móvil 400 puede estar configurado para comunicarse usando múltiples normas de comunicaciones inalámbricas, incluidas normas seleccionadas de las definiciones de normas 3GPP LTE, WiMAX, acceso por paquetes de alta velocidad (HSPA), Bluetooth y Wi-Fi. El dispositivo móvil 400 puede comunicarse usando antenas individuales para cada norma de comunicación inalámbrica o antenas compartidas para múltiples normas de comunicación inalámbrica. El dispositivo móvil 400 puede comunicarse en una WLAN, una WPAN y/o una WWAN.

La FIG. 4 también ilustra un micrófono 420 y uno o más altavoces 412 que pueden usarse para la entrada y salida de audio desde el dispositivo móvil 400. Una pantalla de visualización 404 puede ser una pantalla de cristal líquido (LCD) u otro tipo de pantalla de visualización, tal como una pantalla de diodos orgánicos de emisión de luz (OLED). La pantalla de visualización 404 puede estar configurada como una pantalla táctil. La pantalla táctil puede usar una tecnología capacitiva, resistiva u otro tipo de tecnología de pantalla táctil. Un procesador de aplicaciones 414 y un

procesador de gráficos 418 pueden estar acoplados a una memoria interna 416 para proporcionar capacidades de procesamiento y visualización. Un puerto de memoria no volátil 410 también puede usarse para proporcionar opciones de entrada/salida de datos a un usuario. El puerto de memoria no volátil 410 también puede usarse para ampliar las capacidades de memoria del dispositivo móvil 400. Un teclado 406 puede estar integrado en el dispositivo móvil 400 o conectado de manera inalámbrica al dispositivo móvil 400 para proporcionar una entrada de usuario adicional. También puede proporcionarse un teclado virtual usando la pantalla táctil. Una cámara 422 situada en el lado delantero (pantalla de visualización) o en el lado trasero del dispositivo móvil 400 también puede estar integrada en el alojamiento 402 del dispositivo móvil 400.

La FIG. 5 es un diagrama de bloques que ilustra una máquina de sistema informático 500 de ejemplo en la que puede ejecutarse una cualquiera o más de las metodologías descritas en el presente documento. La máquina de sistema informático 500 puede realizarse como el UE local 106, el UE remoto 108 o cualquier otra plataforma informática descrita o mencionada en el presente documento. En formas de realización alternativas, la máquina funciona como un dispositivo independiente o puede conectarse (por ejemplo, a través de una red) a otras máquinas. En una implantación conectada en red, la máquina puede funcionar utilizando la capacidad de un servidor o una máquina cliente en entornos de red cliente-servidor, o puede actuar como una máquina homóloga en entornos de red de igual a igual (o distribuida). La máquina puede ser un ordenador personal (PC) que puede ser portátil, o no, (por ejemplo, un *notebook* o un *netbook*), una tableta, un codificador (STB), una consola de juegos, un asistente digital personal (PDA), un teléfono móvil o teléfono inteligente, una aplicación web, un encaminador de red, un conmutador o puente, o cualquier máquina capaz de ejecutar instrucciones (secuenciales o de otro tipo) que especifican acciones que serán realizadas por tal máquina. Además, aunque solo se ilustra una única máquina, cabe señalar que el término "máquina" incluye cualquier colección de máquinas que ejecuten de manera individual o conjunta un conjunto (o varios conjuntos) de instrucciones para llevar a cabo una cualquiera o más de las metodologías descritas en el presente documento.

La máquina de sistema informático 500 de ejemplo incluye un procesador 502 (por ejemplo, una unidad de procesamiento central (CPU), una unidad de procesamiento de gráficos (GPU) o ambas), una memoria principal 504 y una memoria estática 506, que se comunican entre sí a través de una interconexión 508 (por ejemplo, un enlace, un bus, etc.). La máquina de sistema informático 500 puede incluir además una unidad de visualización de vídeo 510, un dispositivo de entrada alfanumérico 512 (por ejemplo, un teclado) y un dispositivo de navegación de interfaz de usuario (UI) 514 (por ejemplo, un ratón). En una forma de realización, la unidad de visualización de vídeo 510, el dispositivo de entrada 512 y el dispositivo de navegación UI 514 son una pantalla táctil. La máquina de sistema informático 500 puede incluir además un dispositivo de almacenamiento 516 (por ejemplo, una unidad de disco), un dispositivo de generación de señales 518 (por ejemplo, un altavoz), un controlador de salida 532, un controlador de gestión de energía 534 y un dispositivo de interfaz de red 520 (que puede incluir o comunicarse de manera operativa con una o más antenas 530, transceptores u otro hardware de comunicaciones inalámbricas), y uno o más sensores 528, tal como un sensor de sistema de posicionamiento global (GPS), una brújula, un sensor de ubicación, un acelerómetro u otro sensor.

El dispositivo de almacenamiento 516 incluye un medio legible por máquina 522 en el que hay almacenado uno o más conjuntos de estructuras de datos e instrucciones 524 (por ejemplo, software) que representan o son utilizadas por una cualquiera o más de las metodologías o funciones descritas en el presente documento. Las instrucciones 524 también pueden residir, completamente o al menos parcialmente, en la memoria principal 504, la memoria estática 506 y/o en el procesador 502 durante su ejecución mediante la máquina de sistema informático 500, donde la memoria principal 504, la memoria estática 506 y el procesador 502 constituyen además un medio legible por máquina.

Aunque el medio legible por máquina 522 se ilustra en una forma de realización de ejemplo como un único medio, el término "medio legible por máquina" puede incluir un único medio o múltiples medios (por ejemplo, una base de datos centralizada o distribuida, y/o memorias caché y servidores asociados) que almacenen la una o más instrucciones 524. Debe tenerse en cuenta también que el término "medio legible por máquina" incluye cualquier medio tangible que pueda almacenar, codificar o llevar instrucciones para su ejecución mediante la máquina y que hagan que la máquina realice una cualquiera o más de las metodologías de la presente divulgación, o que pueda almacenar, codificar o llevar estructuras de datos utilizadas por o asociadas a tales instrucciones.

Las instrucciones 524 pueden transmitirse o recibirse además a través de una red de comunicaciones 526 usando un medio de transmisión a través del dispositivo de interfaz de red 520 utilizando uno cualquiera de una pluralidad de protocolos de transferencia ampliamente conocidos (por ejemplo, HTTP). Debe tenerse en cuenta que el término "medio de transmisión" incluye cualquier medio intangible que sea capaz de almacenar, codificar o llevar instrucciones para su ejecución mediante la máquina, e incluye señales de comunicación digitales o analógicas u otro medio intangible para facilitar la comunicación de tal software.

Debe entenderse que las unidades o capacidades funcionales descritas en esta memoria descriptiva pueden hacerse mencionado o etiquetado como componentes o módulos, con el fin de enfatizar en particular su independencia de implantación. Por ejemplo, un componente o módulo puede implementarse como un circuito de hardware que comprende circuitos o matrices de puertas personalizadas de integración a muy gran escala (VLSI),

semiconductores comerciales tales como chips lógicos, transistores u otros componentes discretos. Un componente o módulo también puede implementarse en dispositivos de hardware programables, tales como matrices de puertas programables *in situ*, lógica de matrices programable, dispositivos lógicos programables o similares. Los componentes o módulos también pueden implementarse en software para su ejecución mediante varios tipos de procesadores. Un componente o módulo identificado de código ejecutable puede comprender, por ejemplo, uno o más bloques físicos o lógicos de instrucciones de ordenador, que pueden organizarse, por ejemplo, como un objeto, procedimiento o función. Sin embargo, los ejecutables de un componente o módulo identificado no tienen que estar situados físicamente juntos, sino que pueden comprender instrucciones diferentes almacenadas en diferentes ubicaciones en las que, cuando se unen de manera lógica entre sí, comprenden el componente o módulo y consiguen la finalidad prevista del componente o módulo.

De hecho, un componente o módulo de código ejecutable puede ser una única instrucción, o muchas instrucciones, y puede estar distribuido a través de varios segmentos de código diferentes, entre diferentes programas y a través de varios dispositivos de memoria. Asimismo, los datos de funcionamiento pueden identificarse e ilustrarse en el presente documento dentro de componentes o módulos, y pueden adoptar cualquier forma adecuada y organizarse dentro de cualquier tipo adecuado de estructura de datos. Los datos de funcionamiento pueden recopilarse como un único conjunto de datos o pueden estar distribuidos en diferentes ubicaciones, incluidos diferentes dispositivos de almacenamiento, y pueden existir, al menos parcialmente, como señales electrónicas en un sistema o red. Los componentes o módulos pueden ser pasivos o activos, incluidos agentes que pueden hacerse funcionar para llevar a cabo las funciones deseadas.

Ejemplos adicionales de las formas de realización de procedimiento, sistema y dispositivo descritas en el presente documento incluyen las siguientes configuraciones no limitativas. Cada uno de los siguientes ejemplos no limitativos pueden existir por sí solos o pueden combinarse en cualquier permutación o combinación con uno cualquiera o más de los otros ejemplos proporcionados a continuación o a lo largo de la presente divulgación.

El ejemplo 1 incluye el contenido objeto de un procedimiento llevado a cabo por un equipo de usuario local (UE local) para modificar la distancia focal en una aplicación de videoconferencia que proporciona una comunicación de vídeo bidireccional entre el UE local y un equipo de usuario remoto (UE remoto), incluyendo el procedimiento: definir, en el UE local, una región de interés (ROI) definida por el usuario dentro del campo de visión en el UE remoto; indicar al UE remoto que capture vídeo dentro de la ROI; recibir, en el UE local, el vídeo capturado desde el UE remoto, donde el vídeo recibido incluye regiones dentro de la ROI y excluye regiones fuera de la ROI; y mostrar, en el UE local, el vídeo recibido.

En el ejemplo 2, el contenido del ejemplo 1 puede incluir, opcionalmente, el procedimiento del ejemplo 1, donde indicar al UE remoto que capture vídeo dentro de la ROI incluye indicar la ROI en notificaciones de realimentación RTCP; y enviar las notificaciones de realimentación RTCP al UE remoto.

En el ejemplo 3, el contenido del ejemplo 1 puede incluir, opcionalmente, indicar al UE remoto que capture vídeo dentro de la ROI, lo que incluye capturar vídeo local en el UE local; incluir la ROI en al menos una extensión de cabecera RTP en el vídeo local capturado; y enviar al UE remoto el vídeo local capturado, con la al menos una extensión de cabecera RTP incluida.

En el ejemplo 4, el contenido del ejemplo 3 puede incluir, opcionalmente, que la al menos una extensión de cabecera RTP incluya una posición y un tamaño, donde la posición y el tamaño definen la ROI.

En el ejemplo 5, el contenido del ejemplo 4 puede incluir, opcionalmente, que la posición incluya dos valores de posición que indican una posición de esquina de la ROI dentro del campo de visión en el UE remoto.

En el ejemplo 6, el contenido del ejemplo 5 puede incluir, opcionalmente, que los dos valores de posición se normalicen para indicar la posición relativa en dos dimensiones.

En el ejemplo 7, el contenido de un ejemplo o de cualquier combinación de los ejemplos 4 a 6 puede incluir, opcionalmente, que el tamaño incluya dos valores de tamaño que indican el tamaño de la ROI dentro del campo de visión en el UE remoto.

En el ejemplo 8, el contenido del ejemplo 7 puede incluir, opcionalmente, que los dos valores de tamaño se normalicen para indicar el tamaño relativo en dos dimensiones.

En el ejemplo 9, el contenido de un ejemplo o de cualquier combinación de los ejemplos 1 a 8 puede incluir, opcionalmente, que la ROI incluya un número variable de píxeles; y donde el vídeo recibido tiene una tasa de actualización variable que es inversamente proporcional al número de píxeles de la ROI.

En el ejemplo 10, el contenido de un ejemplo o de cualquier combinación de los ejemplos 1 a 9 puede incluir, opcionalmente, que la ROI incluya un número fijo de píxeles; y donde el vídeo recibido tiene una tasa de actualización constante.

El ejemplo 11 incluye el contenido objeto de un equipo de usuario (UE), que incluye un sistema de circuitos configurado para llevar a cabo una videoconferencia bidireccional entre el UE y un equipo de usuario remoto (UE remoto), estando dispuesto el sistema de circuitos para: definir, en el UE, una región de interés (ROI) definida por el usuario dentro del campo de visión en el UE remoto; indicar al UE remoto que capture vídeo dentro de la ROI; recibir, en el UE, el vídeo capturado desde el UE remoto, donde el vídeo recibido incluye regiones dentro de la ROI y excluye regiones fuera de la ROI; y mostrar, en el UE, el vídeo recibido.

En el ejemplo 12, el contenido del ejemplo 11 puede incluir, opcionalmente, el procedimiento del ejemplo 1, donde indicar al UE remoto que capture vídeo dentro de la ROI incluye indicar la ROI en notificaciones de realimentación RTCP; y enviar las notificaciones de realimentación RTCP al UE remoto.

En el ejemplo 13, el contenido del ejemplo 11 puede incluir, opcionalmente, indicar al UE remoto que capture vídeo dentro de la ROI, lo que incluye capturar vídeo local en el UE local; incluir la ROI en al menos una extensión de cabecera RTP en el vídeo local capturado; y enviar al UE remoto el vídeo local capturado, con la al menos una extensión de cabecera RTP incluida.

En el ejemplo 14, el contenido del ejemplo 13 puede incluir, opcionalmente, que la al menos una extensión de cabecera RTP incluya una posición y un tamaño, donde la posición y el tamaño definen la ROI.

En el ejemplo 15, el contenido del ejemplo 14 puede incluir, opcionalmente, que la posición incluya dos valores de posición que indican una posición de esquina de la ROI dentro del campo de visión en el UE remoto.

En el ejemplo 16, el contenido del ejemplo 15 puede incluir, opcionalmente, que los dos valores de posición se normalicen para indicar la posición relativa en dos dimensiones.

En el ejemplo 17, el contenido de un ejemplo o de cualquier combinación de los ejemplos 14 a 16 puede incluir, opcionalmente, que el tamaño incluya dos valores de tamaño que indican el tamaño de la ROI dentro del campo de visión en el UE remoto.

En el ejemplo 18, el contenido del ejemplo 17 puede incluir, opcionalmente, que los dos valores de tamaño se normalicen para indicar el tamaño relativo en dos dimensiones.

En el ejemplo 19, el contenido de un ejemplo o de cualquier combinación de los ejemplos 11 a 18 puede incluir, opcionalmente, que la ROI incluya un número variable de píxeles; y donde el vídeo recibido tiene una tasa de actualización variable que es inversamente proporcional al número de píxeles de la ROI.

En el ejemplo 20, el contenido de un ejemplo o de cualquier combinación de los ejemplos 11 a 19 puede incluir, opcionalmente, que la ROI incluya un número fijo de píxeles; y donde el vídeo recibido tiene una tasa de actualización constante.

El ejemplo 21 incluye el contenido objeto de un procedimiento llevado a cabo por un equipo de usuario local (UE local) para modificar la distancia focal en una aplicación de videoconferencia que proporciona una comunicación de vídeo bidireccional entre el UE local y un equipo de usuario remoto (UE remoto), incluyendo el procedimiento definir, en el UE local, una región de interés (ROI) definida por el usuario dentro del campo de visión en el UE remoto; capturar vídeo local en el UE local; incluir la ROI en al menos una extensión de cabecera RTP en el vídeo local capturado; y enviar el vídeo local capturado, con la al menos una extensión de cabecera RTP incluida, al UE remoto, incluyendo la al menos una extensión de cabecera RTP una posición y un tamaño, definiendo la posición y el tamaño la ROI, incluyendo la posición dos valores de posición que indican una posición de esquina de la ROI dentro del campo de visión en el UE remoto, incluyendo el tamaño dos valores de tamaño que indican el tamaño de la ROI dentro del campo de visión en el UE remoto; indicar al UE remoto que capture vídeo dentro de la ROI; recibir, en el UE local, el vídeo capturado desde el UE remoto, donde el vídeo recibido incluye regiones dentro de la ROI y excluye regiones fuera de la ROI; y mostrar, en el UE local, el vídeo recibido.

REIVINDICACIONES

1. Un aparato para usarse en un equipo de usuario (108, 204) basado en MTSI (servicios de telefonía multimedia a través de IMS) que puede configurarse para señalar una región de interés de vídeo y funcionar como un emisor MTSI (108, 204), comprendiendo el aparato:

una memoria (416, 504); y
un sistema de circuitos de procesamiento (414, 418, 502) configurado para:

señalar información de región de interés de vídeo para una primera región de interés (110) de un receptor MTSI (106, 202) en paquetes de protocolo en tiempo real, RTP, donde los paquetes RTP incluyen al menos un comando de modificación de distancia focal para capturar la primera región de interés (110), siendo la primera región de interés (110) una región de interés solicitada (110) del receptor MTSI (106, 202);

descodificar paquetes de datos útiles RTP recibidos desde el receptor MTSI (106, 202), comprendiendo los paquetes de datos útiles RTP vídeo correspondiente a la primera región de interés (110);

recibir señalización en notificaciones de realimentación de protocolo de control de transporte en tiempo real, RTCP, desde el receptor MTSI (106, 202) que solicita una segunda región de interés (110), donde la segunda región de interés (110) se proporciona durante una negociación de capacidad de protocolo de descripción de sesión, SDP, siendo la segunda región de interés (110) una región de interés predefinida (110) del emisor MTSI (108, 204); y

codificar vídeo correspondiente a la segunda región de interés (110) en paquetes de datos útiles RTP para su transmisión al receptor MTSI (106, 202).

2. El aparato según la reivindicación 1, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para codificar vídeo correspondiente a la segunda región de interés (110) de una imagen capturada para su transmisión al receptor MTSI (106, 202) y evitar la codificación de vídeo fuera de la segunda región de interés (110) de la imagen capturada.

3. El aparato según la reivindicación 1, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para señalar la información de región de interés de vídeo para la primera región de interés (110) del receptor MTSI (106, 202) en paquetes RTP cuando funciona en modo de control de cámara remota, donde los paquetes RTP incluyen al menos el comando de modificación de distancia focal para capturar la región de interés solicitada.

4. El aparato según la reivindicación 1, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para señalar la información de región de interés de vídeo para la primera región de interés (110) del receptor MTSI (106, 202) en paquetes RTP cuando funciona en modo de control de cámara remota, donde los paquetes RTP incluyen al menos el comando de modificación de distancia focal para capturar la región de interés solicitada sin cambiar el campo de visión del receptor MTSI (106, 202).

5. El aparato según la reivindicación 1, en el que el emisor MTSI (108, 204) está configurado para recibir señalización en notificaciones de realimentación RTCP desde el receptor MTSI (106, 202) que solicita la segunda región de interés (110) del emisor MTSI (108, 204).

6. El aparato según la reivindicación 5, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para codificar vídeo correspondiente a la segunda región de interés (110) en paquetes de datos útiles RTP para su transmisión al receptor MTSI (106, 202) sin cambiar el campo de visión del emisor MTSI (108, 204).

7. El aparato según la reivindicación 1, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado además para:

llevar a cabo una negociación de capacidad SDP con el receptor MTSI (106, 202);
obtener un conjunto de segundas regiones de interés (110) a partir del receptor MTSI (106, 202) durante la negociación de capacidad SDP; y
señalar una de las segundas regiones de interés (110) en notificaciones de realimentación RTCP para el receptor MTSI (106, 202) cuando funciona en un modo de segunda región de interés (110).

8. El aparato según la reivindicación 1, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado además para:

llevar a cabo una negociación de capacidad SDP con el receptor MTSI (106, 202);
proporcionar un conjunto de segundas regiones de interés (110) al receptor MTSI (106, 202) durante la negociación de capacidad SDP; y

recibir señalización para una de las segundas regiones de interés (110) en notificaciones de realimentación RTCP desde el receptor MTSI (106, 202) cuando el receptor MTSI (106, 202) está funcionando en un modo de segunda región de interés (110).

- 5 9. El aparato según la reivindicación 8, en el que para señalar la información de región de interés de vídeo para el receptor MTSI (106, 202), el sistema de circuitos de procesamiento (414, 418, 502) está configurado para codificar la información de región de interés de vídeo en paquetes RTP para su transmisión al receptor MTSI (106, 202).
- 10 10. El aparato según la reivindicación 9, que comprende además una cámara para capturar vídeo correspondiente a la segunda región de interés (110).
11. El aparato según la reivindicación 10, que comprende además una pantalla para mostrar los paquetes de datos útiles RTP descodificados y recibidos desde el receptor MTSI (106, 202).
- 15 12. Un medio de almacenamiento legible por ordenador no transitorio que almacena instrucciones para su ejecución por el sistema de circuitos de procesamiento (414, 418, 502) de un equipo de usuario (108, 204) basado en MTSI (servicios de telefonía multimedia a través de IMS), que pueden configurarse para señalar una región de interés de vídeo y funcionar como un emisor MTSI (108, 204), donde las instrucciones configuran el equipo de usuario (108, 204) para que:
20 lleve a cabo una negociación de capacidad de protocolo de descripción de sesión, SDP, con un receptor MTSI (106, 202);
codifique la información de región de interés de vídeo para señalar una primera región de interés (110) en paquetes de protocolo en tiempo real, RTP, donde los paquetes RTP incluyen al menos el comando de
25 modificación de distancia focal para capturar la primera región de interés (110), siendo la primera región de interés (110) una región de interés solicitada (110) del receptor MTSI (106, 202);
descodifique paquetes de datos útiles RTP recibidos desde el receptor MTSI (106, 202), comprendiendo los paquetes de datos útiles RTP vídeo correspondiente a la primera región de interés (110);
30 reciba señalización en notificaciones de realimentación de protocolo de control de transporte en tiempo real, RTCP, desde el receptor MTSI (106, 202) que solicita una segunda región de interés (110), donde la segunda región de interés (110) se proporciona durante la negociación de capacidad SDP, siendo la segunda región de interés (110) una región de interés predefinida (110) del emisor MTSI (108, 204); y
codifique vídeo correspondiente a la segunda región de interés (110) en paquetes de datos útiles RTP para su transmisión al receptor MTSI (106, 202).
- 35 13. El medio de almacenamiento legible por ordenador no transitorio según la reivindicación 12, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para codificar vídeo correspondiente a la segunda región de interés (110) de una imagen capturada para su transmisión al receptor MTSI (106, 202) y evitar la codificación de vídeo fuera de la segunda región de interés (110) de la imagen capturada.
- 40 14. El medio de almacenamiento legible por ordenador no transitorio según la reivindicación 12, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para señalar la información de región de interés de vídeo para la primera región de interés (110) del receptor MTSI (106, 202) en paquetes RTP cuando funciona en modo de control de cámara remota, donde los paquetes RTP incluyen al menos el comando de modificación de
45 distancia focal para capturar la región de interés solicitada.
15. El medio de almacenamiento legible por ordenador no transitorio según la reivindicación 12, en el que el sistema de circuitos de procesamiento (414, 418, 502) está configurado para señalar la información de región de interés de vídeo para la primera región de interés (110) del receptor MTSI (106, 202) en paquetes RTP cuando funciona en
50 modo de control de cámara remota, donde los paquetes RTP incluyen al menos el comando de modificación de distancia focal para capturar la región de interés solicitada sin cambiar el campo de visión del receptor MTSI (106, 202).

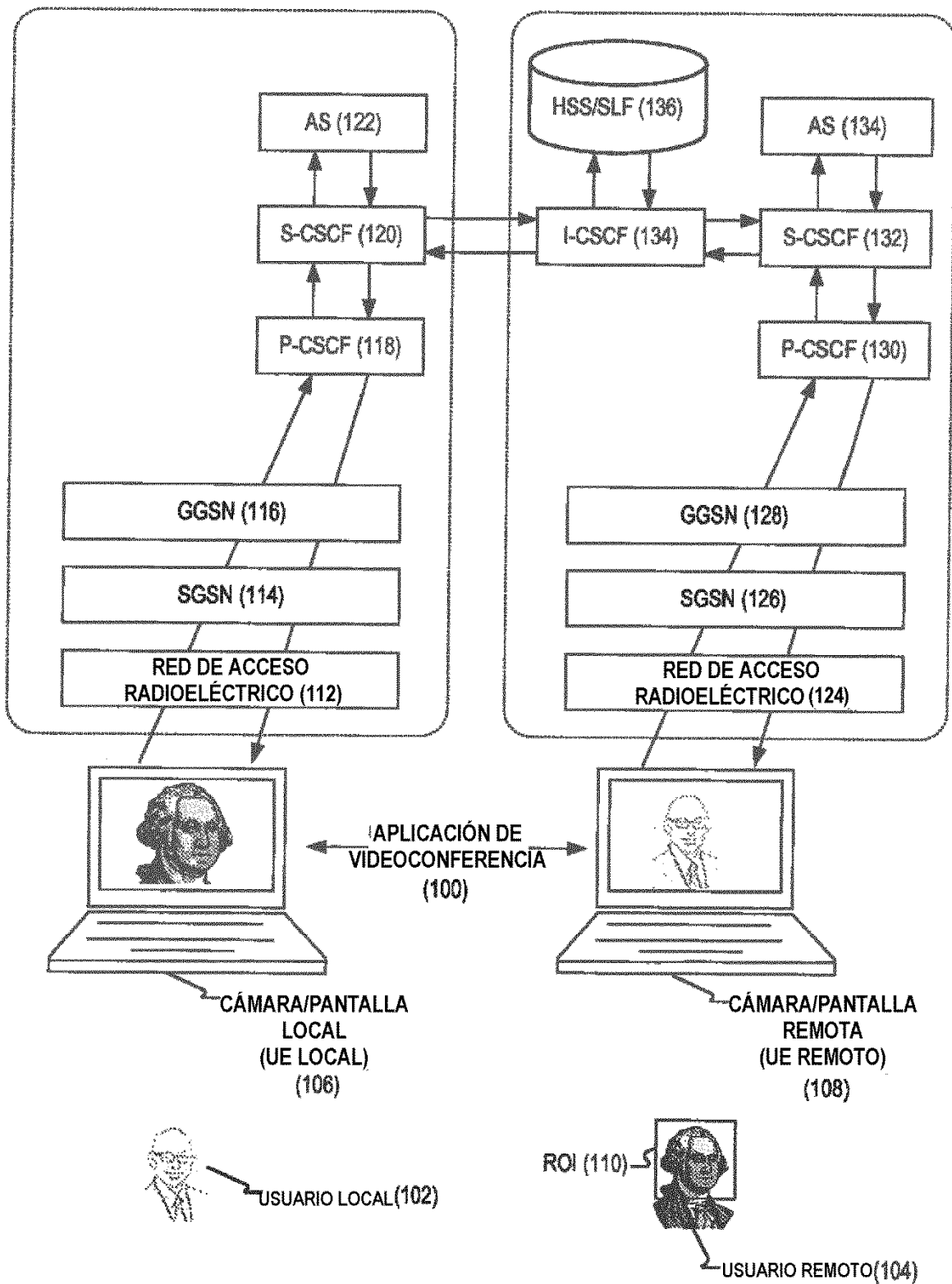


FIG. 1

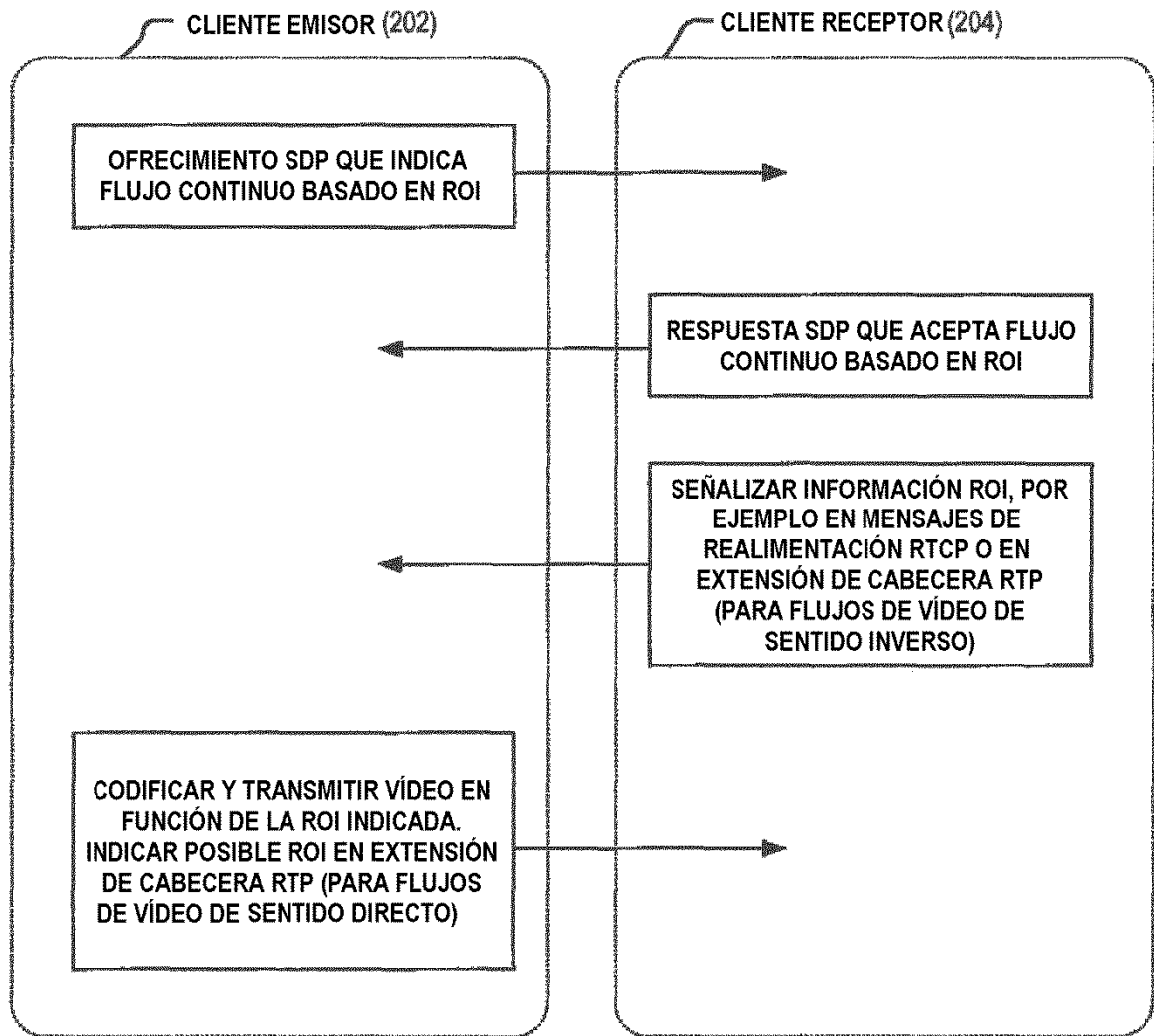


FIG. 2

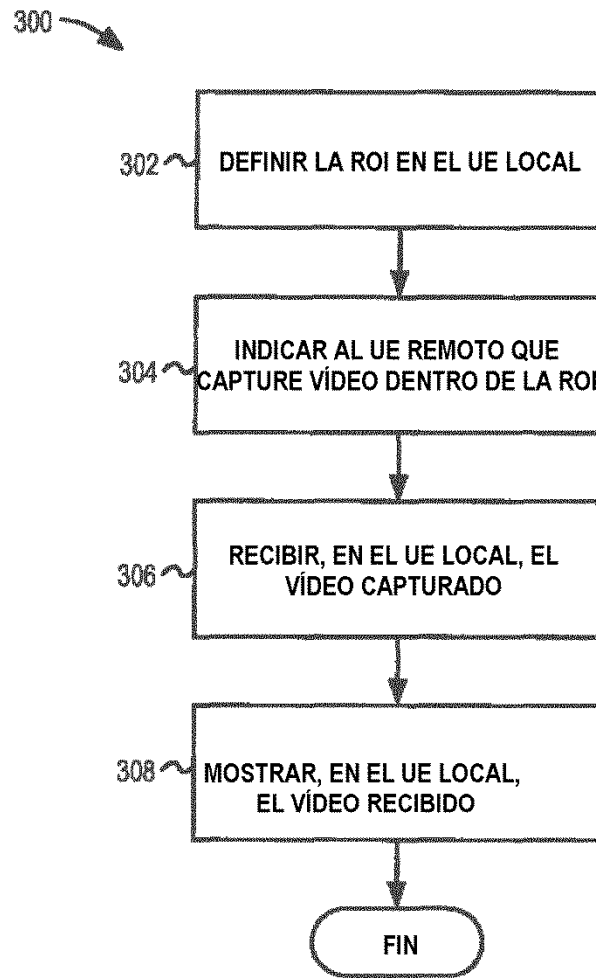


FIG. 3

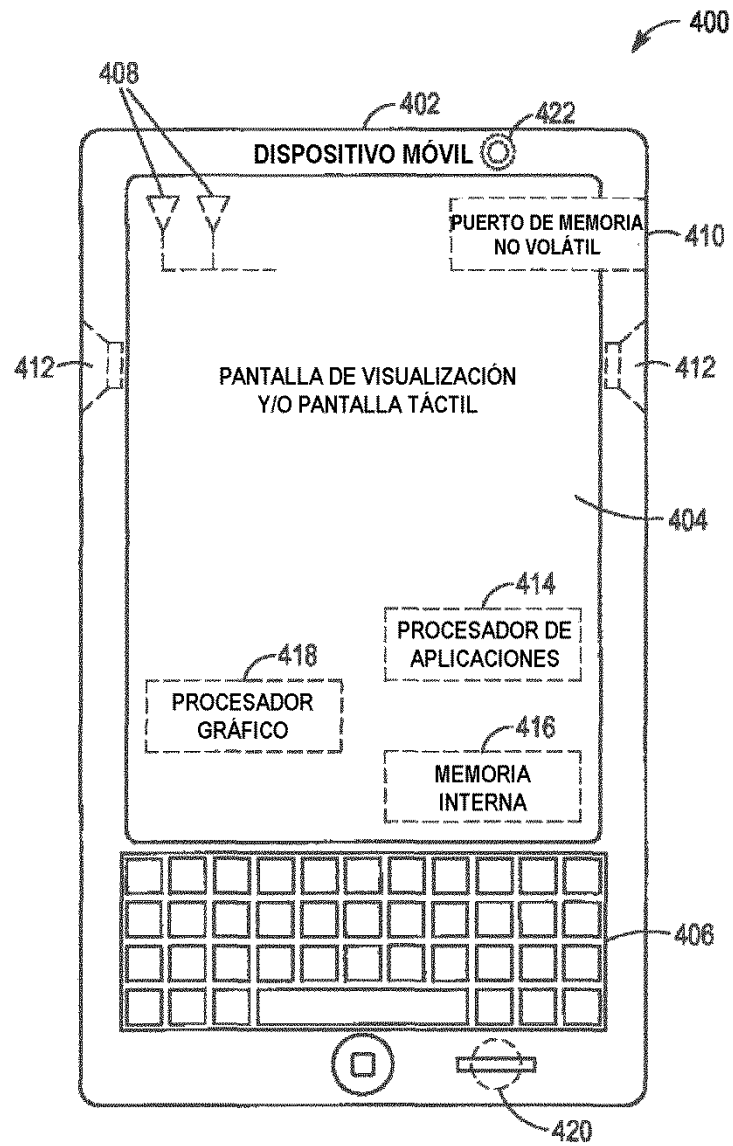


FIG. 4

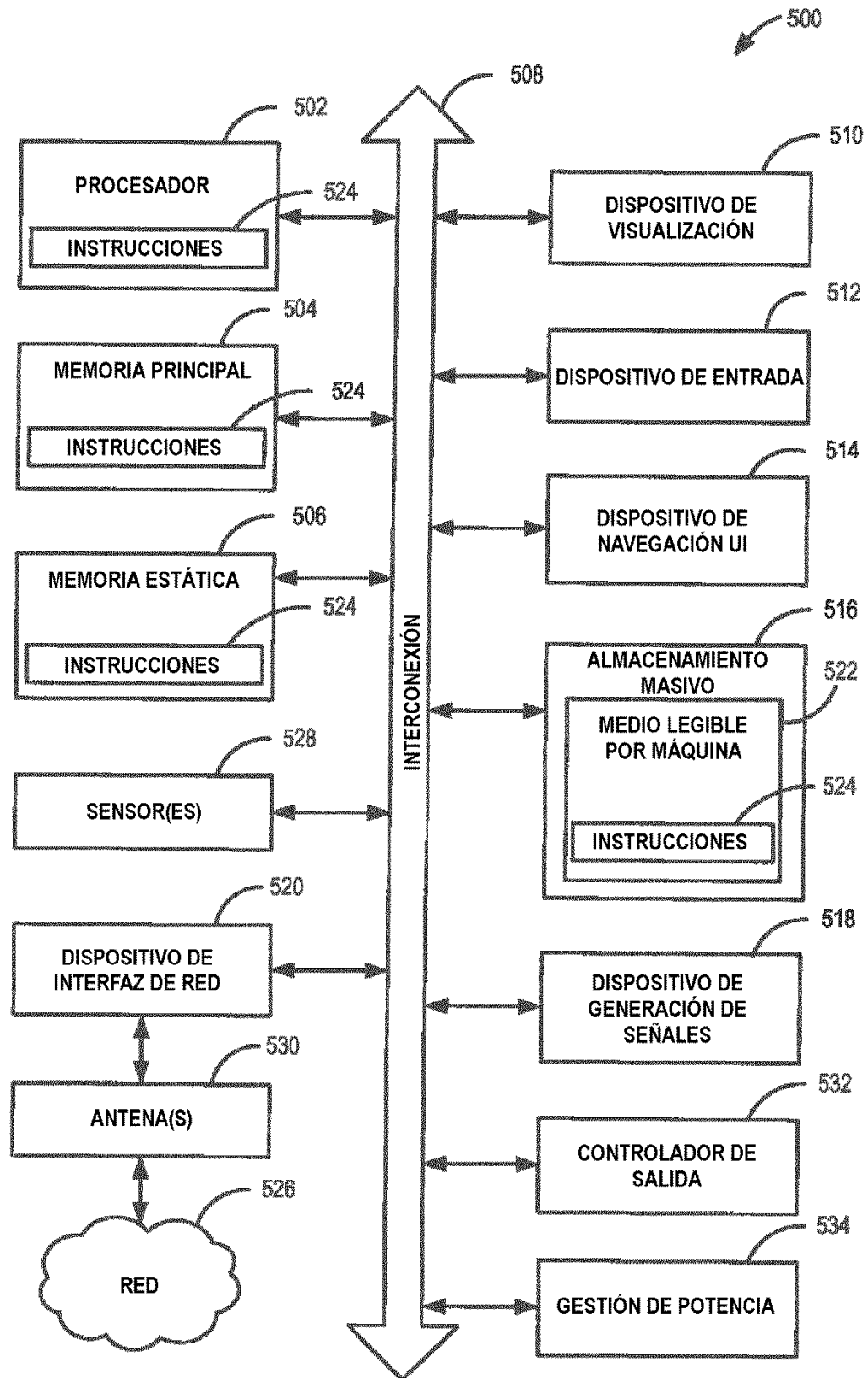


FIG. 5