

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 614 091**

51 Int. Cl.:

<b>H04N 7/173</b>	(2006.01)	<b>H04N 21/4402</b>	(2011.01)
<b>H04N 5/44</b>	(2006.01)	<b>H04N 21/61</b>	(2011.01)
<b>H04N 21/258</b>	(2011.01)		
<b>H04N 21/234</b>	(2011.01)		
<b>H04N 21/2343</b>	(2011.01)		
<b>H04N 21/431</b>	(2011.01)		
<b>H04N 21/482</b>	(2011.01)		
<b>H04N 21/436</b>	(2011.01)		
<b>H04N 21/214</b>	(2011.01)		
<b>H04N 21/2665</b>	(2011.01)		

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **20.07.2011 PCT/US2011/044688**
- 87 Fecha y número de publicación internacional: **02.02.2012 WO12015648**
- 96 Fecha de presentación y número de la solicitud europea: **20.07.2011 E 11812960 (0)**
- 97 Fecha y número de publicación de la concesión europea: **02.11.2016 EP 2599306**

54 Título: **Decodificador virtual**

30 Prioridad:

**23.11.2010 US 952882**  
**30.07.2010 US 369550 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.05.2017**

73 Titular/es:

**GUEST TEK INTERACTIVE ENTERTAINMENT LTD. (100.0%)**  
**Suite 240, 3030-3rd Avenue N.E.**  
**Calgary, AB T2A 6TZ, CA**

72 Inventor/es:

**HULSE, DAVID y**  
**DUKE, THOMAS**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 614 091 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

## Decodificador virtual

La presente invención se refiere a la virtualización de los recursos de hardware en un entorno de red y, en particular, a la virtualización de los decodificadores que proporcionan secuencia de contenido digital para una amplia variedad de dispositivos informáticos y de reproducción de medios.

Un decodificador de hardware convencional (STB) recibe un contenido digital encriptado o secuencia de medios desde una red (por ejemplo, un cable, un satélite o una red basada en IP), descifra la secuencia, la decodifica a fotogramas, y luego da formato a esos fotogramas para la transmisión y la presentación en una televisión conectada; típicamente sobre algún tipo de conexión por cable, tales como, por ejemplo, un cable coaxial estándar o, más recientemente, una interfaz multimedia de alta definición (HDMI). La mayoría de los STBs emplean un sintonizador digital para descodificar la secuencia de medios, aunque la mayoría de los televisores incluyen ahora un sintonizador digital que realiza por lo menos alguna de esta funcionalidad. Sin embargo, otras funcionalidades del STB no han sido incluidas normalmente en televisiones estándar debido a la falta de una norma única. Por ejemplo, una de las razones primarias por las que los STBs se introdujeron fue para el descifrado de contenido de alto valor, algo que no podía ser manejado por la televisión. Una vez introducido para este fin, otras funcionalidades se implementaron en la plataforma disponible, tales como, por ejemplo, interfaces de usuario de navegación avanzadas, y aplicaciones de guía de TV.

Los sistemas que entregan contenido digital para su visualización en televisiones que incluyen, por ejemplo, por cable, por satélite y los sistemas de televisión basados en IP (así como combinaciones de éstos), normalmente se basan en el despliegue de los STB con todos los televisores, lo que representa una enorme porción de los costes de capital asociados con este tipo de sistemas.

El documento US 2006/0195884 describe un sistema de distribución multimedia en el que los servidores se comunican con los clientes, como un decodificador, a través de canales de audio, de vídeo, de superposición y/o de control. La información de audio, de vídeo y/o de superposición se pueden transcodificar antes de la transmisión. Un servidor está conectado a un cliente a través de una red y por lo menos un dispositivo de almacenamiento que contiene audio, vídeo y/o información de superposición formateada de acuerdo con un primer formato de archivo multimedia incluyendo datos de interfaz de usuario para el vídeo. El cliente incluye un dispositivo de almacenamiento que almacena información indicativa del formato de audio, vídeo y/o de superposición que el cliente es capaz de decodificar y el servidor está configurado para transmitir información de audio, de vídeo, de superposición y de control a través de canales separados de audio, de vídeo y de superposición y de control. Una interfaz de usuario se puede incluir en los datos de vídeo enviados por el canal de vídeo.

Se describen los procedimientos y el aparato para proporcionar el contenido digital para cualquiera de una variedad de dispositivos que tienen capacidades de reproducción de los medios sobre una infraestructura de red usando decodificadores virtuales (VSTBs).

De acuerdo con un primer aspecto de la invención, se proporciona un sistema para proporcionar vídeo a través de una red, que comprende: una pluralidad de servidores desplegados en la red y que comprende uno o más servidores decodificadores (VSTB) configurados para crear una instancia de decodificador virtual (VSTB) y uno o más servidores de interfaz de usuario configurados para crear instancias de interfaz de usuario virtual (VUI) para generar una interfaz gráfica de usuario asociada a cada instancia de decodificador virtual (VSTB); y una pluralidad de clientes VSTB con pantallas asociadas, donde la pluralidad de servidores están configurados para gestionar la operación de una pluralidad de instancias de decodificador virtual que operan en el uno o más servidores VSTB, cada instancia VSTB correspondiendo a uno de la pluralidad de clientes VSTB en la red, estando cada instancia VSTB configurada para combinar la interfaz gráfica de usuario con una o más secuencias de vídeo, el o cada servidor VSTB estando configurado además para transmitir la interfaz de usuario gráfica combinada y una o más secuencias de vídeo al cliente VSTB correspondiente como emisión directa de medios individual que tiene un formato de trama y el codificador-decodificador soportado por el correspondiente cliente VSTB y la pantalla asociada, la pluralidad de servidores adicionales configurados para recibir la entrada del usuario que representa la interacción de un usuario asociado con cada cliente VSTB con una representación de la interfaz gráfica de usuario de cada cliente VSTB y proporcionar los datos de usuario a la instancia de VUI correspondiente a la instancia VSTB correspondiente al cliente VSTB, estando cada cliente VSTB configurado para recibir la secuencia de medios individual desde la instancia VSTB correspondiente, y para hacer que la secuencia de medios individual en la pantalla asociada, estando cada cliente VSTB configurado además para transmitir la entrada del usuario a la pluralidad de servidores, en el que para cada instancia de VUI, los uno o más servidores de interfaz de usuario están configurados para comunicar las actualizaciones de la interfaz gráfica de usuario de la instancia de VUI a la instancia VSTB correspondiente a través de una conexión primero informático de red virtual (VNC) con un primer cliente VNC dentro de la instancia VSTB correspondiente y para recibir la entrada del usuario para una particular de las interfaces gráficas de usuario en la instancia de VUI a través de una segunda conexión VNC con un segundo cliente VNC correspondiente al cliente VSTB asociado con la pantalla en la que se representa la interfaz gráfica de usuario en particular, y en el que cada instancia VSTB está configurada para procesar previamente secuencias de medios por separado las secuencias que recibe durante una sesión para empalmar las secuencias de medios

independientes en la secuencia del medio individual.

Una comprensión adicional de la naturaleza y las ventajas de la presente invención puede realizarse por referencia a las partes restantes de la memoria descriptiva y los dibujos.

La figura 1 es un diagrama de bloques simplificado de un sistema de entrega de contenido digital diseñado de acuerdo con diversas realizaciones de la invención.

Las figuras 2A - 2C son diagramas de bloques simplificados de unidades de control de televisión para su uso con diversas realizaciones de la invención.

La figura 3 es un diagrama de bloques simplificado que representa la funcionalidad de un módulo de conexión virtual (VSTB) de acuerdo con diversas realizaciones específicas de la invención.

Ahora se hará referencia en detalle a realizaciones específicas de la invención incluyendo los mejores modos contemplados por los inventores para llevar a cabo la invención. Ejemplos de estas realizaciones específicas se ilustran en los dibujos adjuntos. Aunque la invención se describe en conjunción con estas realizaciones específicas, se entenderá que no se pretende limitar la invención a las realizaciones descritas. Por el contrario, se pretende cubrir alternativas, modificaciones y equivalentes que puedan estar incluidos dentro del alcance de la invención como se define por las reivindicaciones adjuntas. En la siguiente descripción, los detalles específicos se exponen con el fin de proporcionar una comprensión completa de la presente invención. La presente invención puede ponerse en práctica sin algunos o todos estos detalles específicos. Además, las características bien conocidas pueden no haberse descrito en detalle para evitar oscurecer innecesariamente la invención.

Diversas realizaciones de la presente invención emplean decodificadores virtuales (VSTBs) que virtualizan algunas o todas las funcionalidades asociadas con un decodificador convencional (STB), eliminando así la necesidad de despliegue de dispositivos costosos y complicados para cada televisión o pantalla en un sistema de entrega de contenido digital. Mediante la virtualización o la redistribución de la funcionalidad asociada típicamente con decodificadores convencionales a otros componentes dentro de la arquitectura general del sistema, el coste del STB convencional puede ser eliminado.

Además, diversas realizaciones de la invención pueden caracterizarse por diversas ventajas. Por ejemplo, de acuerdo con realizaciones específicas, una instancia VSTB se asigna sobre una base de uno a uno a la televisión u otro dispositivo asociado a cada usuario solicitante. Como tal, una instancia de VSTB solo se requiere para cada televisión (u otro dispositivo de pantalla, informático, o de reproducción de medios) que está siendo usada activamente. Esto debe ser contrastado con STBs convencionales, que son necesarios para estar presente con cada televisión tanto si la televisión está siendo utilizada o no.

Además, un STB físico generalmente solo es capaz de procesar como mucho 1 o 2 secuencias de medios de entrada debido a las limitaciones del hardware STB. Un VSTB, por el contrario, se ejecuta como un procedimiento de software en un servidor con múltiples núcleos de CPU que proporcionan muchas veces más de potencia de procesamiento. Por lo tanto, un VSTB puede hacer uso de este poder de procesamiento adicional para llevar a cabo tareas que un STB convencional simplemente no puede, y nunca fue diseñado para hacer. Por ejemplo, un VSTB puede ser capaz de procesar varias secuencias de medios de entrada al mismo tiempo, organizándolos en una parrilla o patrón de mosaico, y extender su aplicación en ventanas de tamaño variable para la funcionalidad avanzada de imagen en imagen.

Una ventaja adicional al uso de VSTBs de acuerdo con realizaciones de la invención es la capacidad de acelerar el tiempo que se necesita para cambiar los canales. Algunos STB convencionales se limitan generalmente a un límite inferior fijo de alrededor de 2 segundos en el tiempo de cambio de canal, determinados por los chips descodificadores de medios específicos utilizados en la construcción del STB. En un VSTB, la tubería de decodificación de medios puede ser optimizada de tal manera que los cambios de canal pueden tener lugar en considerablemente menos tiempo (mucho menos de un segundo en algunos casos). De acuerdo con una realización específica, esta capacidad está activada, por lo menos en parte, por el hecho de que la secuencia de medios se consolida por las instancias VSTB en el extremo posterior en lugar de transmitirse en forma de emisión sin interrupciones individuales que fluyen a través de la red a un STB hardware convencional remoto.

Además, debido a la mayor potencia de procesamiento de un entorno de servidor, es posible que un VSTB procese previamente las secuencias de los medios de entrada que recibe durante una sesión determinada de modo que la reproducción de una serie de secuencias de medios separados aparece ante el espectador como una secuencia continua individual de los medios, al igual que un programa de televisión convencional. Esto está en contraste con la reproducción de la misma serie de secuencias de medios en un STB convencional, que normalmente implican retrasos cortos durante los cuales el STB cambia de una secuencia de medios a la siguiente. Para el espectador, estos retrasos aparecen como intermedios negros, en los que el contenido de los medios desaparece durante unos segundos y luego regresa con el inicio de la siguiente secuencia.

Además, dada la capacidad de algunas implementaciones VSTB para empalmar de forma continua una serie de secuencias de medios en una secuencia continua individual como se describe anteriormente, es posible insertar clips multimedia arbitrarios dentro de una serie de secuencias de medios, o para sustituir una parte de una secuencia de medios con otro clip multimedia. Esto tiene importantes aplicaciones en el ámbito de la secuencia de transmisiones televisivas en vivo, en las que puede ser deseable sustituir los anuncios genéricos en la emisión sin interrupciones entrante con anuncios localizados, quizás dirigidos más específicamente al espectador. Este tipo de aplicación es algo que un STB convencional típicamente no puede manejar debido a la falta de suficiente potencia de procesamiento para realizar las operaciones de corte y empalme y de reproducción necesarias a tiempo real.

Diversas realizaciones de la presente invención se basan en diferentes grados en las capacidades de ajuste digitales disponibles en muchos televisores estándar actuales. Algunas realizaciones se implementan para evitar tener que introducir un dispositivo de cliente o incorporar software de cliente en la televisión. De acuerdo con algunas de estas realizaciones, un ejemplo de VSTB en el lado del servidor genera una secuencia de medios (que puede incluir uno o ambos de los componentes de audio y vídeo) que puede ser decodificada por un sintonizador digital estándar en la televisión asociada. Sin embargo, como se describirá, también se contemplan realizaciones en las que los dispositivos relativamente simples y/o clientes de software se despliegan con las televisiones correspondientes para promover la compatibilidad con una amplia variedad de tipos de televisión y capacidades. En tales casos, la secuencia de medios generada fluye a través del cliente de dispositivo o software implementado en el que se procesa apropiadamente y después se transmite al dispositivo de televisión o de reproducción asociada.

También hay que señalar que, mientras que las realizaciones de la invención se describen con referencia a los televisores como el dispositivo al que los secuencia de medios se proporcionan en última instancia, se contemplan realizaciones de la invención en las que la salida de un VSTB puede ser proporcionada a una amplia gama de medios y dispositivos informáticos, tales como, por ejemplo, ordenadores de sobremesa y portátiles, equipos ultraportátiles, dispositivos portátiles y tabletas (por ejemplo, iPads), dispositivos de mano (por ejemplo, los teléfonos inteligentes como iPhone y Android), y cualquier otro tipo de dispositivo que tenga capacidades de presentación o de reproducción de medios. Además, se contemplan realizaciones de la invención que emplean una variedad de infraestructuras de red para entregar la salida de VSTB a tales dispositivos. Estas infraestructuras incluyen redes basadas en IP (tanto alámbricas como inalámbricas), por cable y redes de televisión por satélite, redes de telecomunicaciones, y diversas combinaciones de los mismos. El alcance de la presente invención no debe por lo tanto estar limitado a las realizaciones específicas descritas en el presente documento.

La figura 1 es un diagrama de nivel de bloque de una arquitectura de sistema de un sistema 100 de entrega de contenido digital implementado de acuerdo con una o más realizaciones específicas de la presente invención. Como se verá, las realizaciones de la invención pueden ser implementadas en el contexto de las redes basadas en IP cableadas e inalámbricas (por ejemplo, mediante un cable de Ethernet o de fibra óptica estándar), redes de cables coaxiales (como existe en muchos hoteles y que se han empleado por proveedores de cable), redes de telecomunicaciones (redes, por ejemplo, 3G o 4G), o combinaciones híbridas de este tipo de tecnologías. Por lo menos algunos de esta diversidad están representados en la figura 1 por los diferentes caminos entre la red 102 IP de extremo trasero y las televisiones 104 en la habitación. Se debe entender que el diagrama de la figura 1 representa realizaciones que emplean las infraestructuras basadas en IP o de cable, así como realizaciones que emplean ambos.

En este ejemplo, una infraestructura híbrida se implementa en una propiedad de un hotel, y ofrece una variedad de servicios que incluyen televisión y vídeo bajo demanda digital (VOD) (tanto en estándar y en alta definición, SD y HD), conexión a Internet, y una variedad de otros servicios de los clientes, incluyendo, por ejemplo, registro de salida a distancia, servicio de despertador, servicio de habitaciones, etc. Sin embargo, como se mencionó anteriormente, debe tenerse en cuenta que esto es simplemente un ejemplo de un contexto en el que dicha infraestructura se puede implementar de acuerdo con la invención. Otros contextos de aplicación, así como una amplia gama de servicios y conjuntos de servicios se contemplan estando dentro del alcance de la invención.

El servidor 108 VSTB (que puede ser uno o más servidores) típicamente tiene múltiples instancias VSTB ejecutándose en un momento dado, cada uno de los cuales ofrece un contenido a una televisión activa u otro dispositivo de reproducción de medios al que se ha asignado la instancia VSTB. Cada instancia VSTB está destinada a proporcionar una experiencia de usuario similar a la de un STB físico y por lo tanto también genera menús interactivos mezclados con o superpuestos en la secuencia de contenido. La forma en que se logran estas funciones de acuerdo con implementaciones específicas se describe a continuación con referencia a la figura 3. Como se mencionó anteriormente, el número de instancias de VSTB ejecutándose solo necesita corresponder al número de televisiones activas en lugar del número total de televisiones en el sistema (como se requiere con STBs convencionales en la habitación). Esto puede representar eficiencias significativas en el coste y el mantenimiento de los recursos de hardware, ya que es típico de 40 % o más de las televisiones estén inactivas en cualquier momento dado.

El servidor 110 de TV lineal (típicamente una matriz de servidores) recibe transmisiones de canal de televisión digital de cualquiera de una variedad de fuentes (por ejemplo, señales europeas DVB-C, señales de satélite de propiedad de U.S., etc.) y convierte cada canal en una secuencia de medios basada en IP que se transmite en la red 102 IP para el consumo de las instancias VSTB activas según sea necesario. El servidor 110 de TV lineal puede cifrar la

secuencia de salida, o dejar intacto el cifrado aplicado originalmente por el proveedor de contenido/locutor.

Según diversas realizaciones, servidor 110 de TV lineal emite un número de secuencias de IPTV, posiblemente cifradas, utilizando una variedad de diferentes codificadores-decodificadores y tamaños de marco. Por ejemplo, la emisión de TV en vivo se codifica comúnmente usando los estándares H.264 o MPEG-2 y puede ser de estándar (SD) o de alta definición (HD). Por otra parte, una instancia VSTB dada generará una secuencia de vídeo dirigida a un dispositivo específico, como una televisión invitada, portátil o teléfono inteligente. El dispositivo en particular tendrá un cierto tamaño de la pantalla y solo será capaz de procesar secuencias de entrada utilizando ciertos codificadores-decodificadores compatibles. Es decir, el VSTB genera una secuencia continua de salida que el dispositivo objetivo sintonizará inicialmente. A partir de entonces, la secuencia de salida proporcionará los menús interactivos y contenido al dispositivo de destino sin que el dispositivo necesite volver a sintonizar cualquier otra secuencia. Si esta secuencia siendo consumida por el dispositivo de destino contuviera cambios periódicos en el tamaño del fotograma y/o codificador-decodificador de compresión, el resultado puede en el mejor de las instancias aparecer como antiestético para el usuario desde una perspectiva estética, y en el peor de las instancias, no puede ser tratado correctamente por el dispositivo de destino lo que resulta en un mal funcionamiento del dispositivo dado que las secuencias de transmisión en vivo por lo general no presentan tales cambios y dispositivos de destino pueden no haber sido probados adecuadamente para manejar estos casos.

Por lo tanto, la instancia de VSTB correspondiente será configurada para generar su salida para que coincida con el tamaño de la pantalla del dispositivo de destino y utilizar un codificador-descodificador soportado por el dispositivo de destino. En consecuencia, cualquier secuencia de vídeo que se usa como entrada para la instancia de VSTB tendrá que ser convertida para que coincida con los de tamaño de la pantalla de destino y la salida del codificador-descodificador. Se contemplan realizaciones de la invención en las que la instancia de VSTB (o una funcionalidad separada en el servidor 108 de VSTB o en otra parte) realiza esta conversión secuencia de entrada. Sin embargo, debido a que las secuencias de entrada son un recurso de red común, puede ser más eficiente en algunas implementaciones tener la secuencia procesada previamente apropiadamente antes de que se consuman mediante un VSTB dado.

De acuerdo con una clase particular de realizaciones, el procesamiento previo de la secuencia es manejada por el remodelador 111 IPTV (que puede ser uno o más servidores), que está diseñado para procesar un número de secuencias en paralelo. El remodelador 111 IPTV consume la salida de las secuencias del servidor 110 de TV lineal y convierte cada secuencia a un formato estándar usando un codificador-descodificador y un fotograma de tamaño especificado. La conversión de secuencia de esta naturaleza se conoce comúnmente como "transcodificación" y generalmente implica las siguientes etapas. En primer lugar, la secuencia de entrada se descifra, si es necesario. La secuencia de entrada entonces se demultiplexa y se descodifica en una secuencia de fotogramas de vídeo no comprimidos y muestras de audio utilizando codificadores-decodificadores de entrada adecuados. Los fotogramas de vídeo sin comprimir se escalan con el tamaño requerido para la secuencia de salida. Los fotogramas de vídeo no comprimidos y muestras de audio se codifican utilizando codificadores-decodificadores de salida adecuados, y las sub-secuencias se multiplexan usando un formato contenedor adecuado, tal como una secuencia de transporte MPEG-2. La secuencia de salida puede entonces ser encriptada si se desea.

Como ejemplo, se considera una secuencia de entrada codificada utilizando MPEG-2 utilizando un tamaño de fotograma de vídeo de 640x480. Una secuencia de este tipo sería decodificada a una secuencia de fotogramas de vídeo formateado YUV de tamaño de 640 x 480 píxeles y una secuencia de muestras de audio formateadas PCM. Cualesquiera otras subsecuencias que contienen información tal como subtítulos o datos privados no son decodificadas y simplemente se pasan a través de la salida. Se considera entonces que los fotogramas de vídeo se escalan a un tamaño de 1280 x 768 de alta definición listos para la salida. Estos fotogramas escalados y las muestras de audio sin comprimir son entonces codificados con un codificador-descodificador H.264 para el vídeo y un codificador-descodificador AC3 para el audio. Todas las subsecuencias son entonces multiplexadas en una secuencia de transporte MPEG-2 para la salida. Debe tenerse en cuenta que cualquiera de una amplia variedad de protocolos de secuencia de transporte de vídeo y mecanismos adecuados se pueden emplear sin apartarse de la presente invención.

Un problema frecuentemente citado con los procedimientos de transcodificación de vídeo es una llamada pérdida generacional de calidad, en el que la naturaleza con pérdida de los codificadores-decodificadores involucrados puede resultar en la pérdida de información adicional cuando se vuelve a codificar la secuencia, por lo tanto, degradar visiblemente la calidad de imagen. Con una cierta cantidad de cuidado durante las etapas de descodificación y codificación, esta pérdida puede evitarse en algunos casos mediante el registro de la información relativa al nivel de cuantificación de cada macrobloque y otros parámetros relevantes que afectan el nivel de capacidad de disipación. Cuando es el momento de volver a codificar la secuencia, los parámetros registrados pueden ser útiles en la determinación de cómo codificar cada macrobloque de salida sin inducir una pérdida adicional. La eficacia de esta técnica depende de la cantidad de escalado involucrada durante la transcodificación y en la similitud de los codificadores-decodificadores de entrada y de salida. Más información sobre esta técnica se puede encontrar en Real-Time Transcoding of MPEG-2 Video Bit Streams de P.N. Tudor y O.H. Werner BBC R&D - IBC'97, Amsterdam.

El servidor 112 VOD (de nuevo típicamente una matriz de servidores) alojan "bajo demanda" el contenido de vídeo

(por ejemplo, largometrajes) y hace que estos contenidos estén disponibles para su transmisión a los VSTBs asignados a los usuarios solicitantes. El servidor 112 VOD activa la funcionalidad de control de vídeo tal como reproducción, pausa, avance rápido, rebobinado, saltar hacia adelante, saltar hacia atrás, etc. Ejemplos de los servidores de VOD que se pueden emplear con sistemas implementados de acuerdo con realizaciones específicas se proporcionan por BitBand Technologies Ltd. de Netanya, Israel. En contraste con la secuencia de medios proporcionados por el servidor 110 de TV lineal, las secuencias de contenido de medios proporcionados por servidor 112 VOD son codificados previamente para tener un tamaño de fotograma y el codificador-decodificador estándar. Además, como con las secuencias de medios de TV lineal, este contenido puede ser encriptado previamente según sea necesario, por ejemplo, por el propietario del contenido.

Como se puede apreciar con referencia a la descripción anterior, el remodelador 111 IPTV y el servidor 112 VOD generan colectivamente un conjunto de secuencias de medios digitales (opcionalmente encriptadas) que comparten un tamaño común de fotograma, el codificador-decodificador, y el esquema de cifrado. Estas secuencias se ponen a disposición para el consumo en la red 102 IP mediante cualquier instancia VSTB que opera el servidor 108 VSTB que pueden así requerirlo. De acuerdo con diversas realizaciones, estas secuencias no están destinadas para el consumo de cualquier otra parte del sistema, incluyendo las televisiones o dispositivos de reproducción multimedia. Por lo tanto, la única manera de que una secuencia se vea en un dispositivo de reproducción en el sistema es para la secuencia que pase como entrada a una instancia VSTB asignada a ese dispositivo. Cabe señalar, sin embargo, que, aunque se describen algunas realizaciones en términos de codificadores-decodificadores comunes y tamaños de fotograma y suponiendo dispositivos de destino homogéneos, se contemplan otras realizaciones que sirven una mezcla heterogénea de los dispositivos de destino (por ejemplo, cualquier tipo de televisión, un dispositivo de computación, dispositivo de telecomunicaciones, y/o cualquier otro tipo de dispositivo de reproducción de medios). Es decir, de acuerdo con algunas realizaciones, los diversos componentes del subsistema de contenido están configurados para soportar tal mezcla heterogénea de dispositivos de destino mediante la configuración previa adecuada del contenido y poniéndolo a disposición en la red IP para el consumo de las instancias VSTB. Por lo tanto, hacer referencia a realizaciones específicas en las que los codificadores-decodificadores y tamaños de fotograma son comunes y/o un conjunto homogéneo de los dispositivos de destino está siendo servido no debe utilizarse para limitar el alcance de la invención.

De acuerdo con una clase específica de realizaciones, el servidor 116 UI (que puede ser varios servidores) proporciona interfaces de usuario virtuales (VUIs) a las instancias VSTB en servidor 108 VSTB. Tales VUIs pueden incluir, por ejemplo, menús interactivos o guías para navegar y seleccionar opciones de contenidos y servicios disponibles en el sistema. Para estas realizaciones, existe una correspondencia uno a uno entre las instancias VUI y VSTB de tal manera que cada vez que se crea una nueva instancia VSTB, una instancia de VUI correspondiente se crea también. Cada instancia VSTB adquirirá una serie de fotogramas UI que representan su correspondiente interfaz de usuario virtual y las superpondrá sobre la secuencia de medios que está generando tal como se describe a continuación con referencia a la figura 3.

De acuerdo con realizaciones específicas, estas interfaces de usuario virtuales se hacen accesibles a instancias VSTB a través de sesiones de Virtual Network Computing (VNC) con clientes VNC que forman parte de las instancias VSTB. El servidor 116 UI proporciona un servidor VNC de alojamiento de cada instancia de VUI lo que permite a cada instancia VSTB conectarse utilizando su cliente VNC interno. De acuerdo con realizaciones más específicas, cada instancia de VUI ejecuta un reproductor de Adobe Flash que convierte la VUI en respuesta a entradas de usuario recibidas desde otra parte de la red.

De acuerdo con una realización más específica, el cliente VNC dentro de cada instancia VSTB se conecta al servidor de VUI solo para recibir las actualizaciones gráficas a su instancia de VUI correspondiente. Este cliente VNC no proporciona ninguna entrada de usuario a la VUI porque tal entrada del usuario no está disponible para el VSTB. En cambio, se utiliza un segundo cliente VNC solo de entrada para proporcionar los eventos de entrada de usuario necesarios para operar con éxito la interfaz de usuario virtual. Estos clientes VNC solo de entrada, específicos para cada instancia de VUI, son parte del servidor 106 administrador de sesión, que se describe a continuación. Por lo tanto, se utilizará un ejemplo de VUI dado en la interfaz de usuario del servidor 116 exportará su interfaz de usuario a través de un servidor de VNC y dos conexiones de cliente VNC separadas, el primero suministrando la representación gráfica de la VUI a la instancia VSTB correspondiente, y la segunda suministrando eventos de entrada de usuario que accionan la interfaz de usuario.

Según algunas realizaciones, el servidor 116 de UI puede ser implementado de una manera similar a la del servidor de UI virtual que se describe en la solicitud de patente US n.º. 12/688.343, publicada como publicación US 2011/0179106 de VIRTUAL USER INTERFACE, presentada el 15 de enero de 2010.

55

Un servidor 106 administrador de sesión (que puede ser varios servidores) recibe los datos de entrada del usuario en representación de la interacción del usuario con las televisiones, dispositivos informáticos, o dispositivos de reproducción de medios (ya sea desde el dispositivo, o por medio de unidades de control de televisión, por ejemplo, 120A-C, tal como se describe a continuación) y proporciona los datos a la instancia de VUI correspondiente en el servidor 116 de UI. Como se ha indicado anteriormente, la entrada del usuario que llega desde una televisión dada o en otro dispositivo de reproducción de medios es recopilada por el servidor 106 administrador de sesión y se envía a la correspondiente instancia VUI en el servidor 116 de UI mediante una conexión de cliente VNC. Esta conexión es especial en que no se solicitan actualizaciones gráficas desde el servidor VNC como normalmente sería el caso. En su lugar, solo los eventos de entrada se reenvían a la instancia de VUI. Esto separa efectivamente los cambios gráficos a partir de la entrada del usuario y permite que el procesamiento de estos aspectos se lleve a cabo en diferentes servidores dentro del sistema.

De acuerdo con algunas realizaciones, el servidor 106 administrador de sesión también se encarga de la responsabilidad de la gestión de los diferentes recursos asignados a cada usuario activo. Por ejemplo, cuando la televisión de un usuario está conectada, el gestor de sesión responde mediante la coordinación de la asignación de recursos en otros servidores, tales como el servidor 108 de VSTB, el servidor 116 de UI, y, posiblemente, Edge QAM Muxer 118 si la televisión está conectada a través de una infraestructura de cable coaxial. Cuando la sesión del usuario se haya completado, el gestor de sesión responde liberando los recursos asignados durante la sesión. De esta manera, el servidor 106 administrador de sesión realiza la función de coordinación de recursos y, por tanto, es fundamental para la operación del sistema y la interconexión y la interoperación de sus diversas partes constituyentes.

De acuerdo con realizaciones en las que por lo menos parte de la infraestructura de red emplea cable coaxial, el Edge QAM Muxer 118 opera como un convertidor de medios que retransmite las secuencias de medios digitales recibidos desde el dominio de IP (por ejemplo, red 102 IP) utilizando modulación de amplitud de cuadratura (QAM) en el dominio coaxial (por ejemplo, planta 119 de cableado del hotel). De acuerdo con una aplicación específica, Edge QAM Muxer 118 opera en una plataforma Linux y convierte múltiples MPEG-2 (SD o HD) o secuencias de transporte de programa individual H.264 (SD o HD) recibidas a través de IP (IP-TV) a varias secuencias de transporte de programa múltiple que se transmiten a través de cable coaxial de acuerdo con el estándar DVB-C usando cualquiera de una variedad de esquemas de modulación QAM, por ejemplo, 64 QAM, 256-QAM, etc. Cabe señalar que otros esquemas de modulación (por ejemplo, desplazamiento de fase en cuadratura (QPSK)) y estándares, (por ejemplo, estándar televisión digital terrestre (DVB-T), estándar "Advanced Television Systems Committee" (ATSC), etc.) pueden emplearse alternativamente. Según algunas realizaciones, Edge QAM Muxer puede ser implementado como se describe en la solicitud de patente U.S. n.º 12/463.993 publicada como publicación de patente US 2010/0251316 de COAX AND IP HYBRID DIGITAL TV AND VOD SYSTEM presentada el 11 de mayo de 2009. Cabe señalar que, de acuerdo con algunas realizaciones, para cada televisión conectada a través de un cable coaxial, el administrador 106 de sesión destinará una porción dedicada de ancho de banda dentro de una salida de secuencia de transporte de programa múltiple desde Edge QAM Muxer 118 para la transmisión de la secuencia de los medios generados por la instancia VSTB asignada a la televisión.

Las descripciones anteriores han subrayado realizaciones mediante las cuales una instancia VSTB, con la ayuda de otros recursos asociados, genera una secuencia de medios destinada al consumo directo por una televisión de destino específico u otro dispositivo de reproducción de medios.

Esta secuencia generada está formada por la composición de cero o más secuencias de medios de entrada que representan la TV lineal de difusión o contenido de VOD y una capa de interfaz gráfica de usuario. La composición de estos elementos de entrada resulta en una sola secuencia de salida de los medios que proporciona toda la experiencia interactiva para el usuario. Por lo tanto, la televisión necesita simplemente sintonizar esta secuencia generada cuando se activa primero al inicio de la sesión del usuario. A partir de entonces, la secuencia generada seguirá reflejando las elecciones del usuario cuando se utiliza el sistema de televisión interactiva. Si el usuario desea ver un canal de televisión o VOD activo lineal diferente, las entradas a la instancia VSTB cambiarán y la secuencia generada reflejará estos cambios. Sin embargo, la televisión no necesita volver a sintonizar a una secuencia diferente en cualquier momento, mientras que la sesión del usuario está activa. También hay que señalar que la secuencia generada se suministra a la televisión de destino ya sea a través Edge QAM Muxer 118 o como una secuencia IPTV directa, dependiendo de los tipos de entradas que la televisión está equipada para gestionar. Por ejemplo, muchas televisiones aceptarán una entrada coaxial digital, mientras que otras televisiones pueden ser capaces de recibir entradas de IPTV directamente.

Aunque no es necesario para todas las realizaciones, las diversas realizaciones de la invención pueden incluir unidades de control de la TV (por ejemplo, 120A-C) asociadas a las televisiones (por ejemplo, 104A-C) que generalmente no incluyen el tipo de capacidades de procesamiento de los medios típicos de un STB convencional. Ejemplos de dichas unidades de control de la TV se muestran en los diagramas de bloques de las figuras 2A - 2C. De acuerdo con algunas de estas realizaciones, la unidad de control de la televisión puede proporcionar un servicio de comunicación para que los servidores en el otro lado de la red puedan controlar los dispositivos conectados a la unidad remota usando, por ejemplo, IP como protocolo. En el caso de una red Ethernet, la facilidad de comunicación puede incluir un chip interruptor de Ethernet, que recibe las comunicaciones Ethernet entrando en la habitación de la

red. Véase, por ejemplo, interruptor de chips de Ethernet 222 y 242 de las figuras 2B y 2C. En el caso de una red coaxial, la instalación de comunicación puede ser, por ejemplo, un mecanismo de IP-sobre-coaxial tal como, por ejemplo, el estándar HPNA (anteriormente el "Home Phoneline Networking Alliance"). Véase, por ejemplo, el conjunto de chips HPNA 202 de la figura 2A. El usuario puede conectar su propio dispositivo informático, por ejemplo, un ordenador portátil u otro dispositivo informático, a la red a través de estos recursos y los conectores Ethernet (HSIA) disponibles.

Puertos USB 2.0 también se proporcionan con algunas realizaciones para la conexión de varios tipos de dispositivos de los huéspedes para una variedad de propósitos, incluyendo, por ejemplo, la carga de los dispositivos, o los dispositivos de interfaz con el sistema de alguna manera. Por ejemplo, un dispositivo huésped podría ser una cámara que, cuando se conecta al sistema, podría cargar sus imágenes al sistema para su visualización en la televisión conectada. En otro ejemplo, un usuario puede conectar un dispositivo de memoria (por ejemplo, una unidad USB) que, cuando está conectado al sistema, haría que los archivos almacenados en la unidad estuvieran disponibles en el sistema. En otro ejemplo, un usuario puede conectar un reproductor de música (por ejemplo, un iPod o un dispositivo similar) que, cuando se conecta al sistema, haría que los archivos de música del usuario estuvieran disponibles para su reproducción en el sistema. Cualquiera de una variedad de otros ejemplos será evidente para los expertos en la técnica.

Cada unidad de control de la TV incluye recursos de procesamiento (por ejemplo, un microcontrolador con firmware) para la administración de la operación de la unidad, y para ejecutar cualquier protocolo necesario para la comunicación con cualquier recurso o dispositivo conectado, por ejemplo, interactuar con los servidores remotos (por ejemplo, a través de una intervención facilidad de comunicación), recibir y responder a los comandos basados en IP, así como las aportaciones de la televisión u otros recursos conectados, etc. Véanse, por ejemplo, los microcontroladores 204, 224, y 244 de las figuras 2A - 2C. De acuerdo con realizaciones específicas, estos recursos de procesamiento ejecutan un protocolo en serie para comunicarse con y controlar la televisión (por ejemplo, a través de los puertos RJ12 que se muestran en cada una de las figuras) según sea necesario durante la operación del sistema (por ejemplo, para encender o apagar la televisión, sintonizar un canal en particular, etc.).

Para realizaciones implementadas en las redes de cable coaxial, una parte del espectro de RF presente en el cable se utiliza para la transmisión bidireccional de paquetes de datos IP (por ejemplo, datos HPNA), mientras que otra parte del espectro se utiliza para la transmisión unidireccional de contenido digital hacia la televisión. En tales realizaciones, la unidad de control de la TV incluye un mecanismo de división (por ejemplo, el divisor 205 de la figura 2A) con filtros asociados en cada una de las rutas de salida (no mostrados), que separan los datos IP a partir de los datos de los medios digitales. Los datos IP se interpretan mediante el firmware del microcontrolador, mientras que los datos de medios digitales se pasan al sintonizador digital en la televisión.

Para realizaciones de Ethernet (por ejemplo, las figuras 2B y 2C), la unidad de control de la TV solo recibe paquetes de datos IP que incluyen datos de control y los paquetes de contenido que contienen los datos de medios digitales. La unidad de control de la TV reconoce qué paquetes son y los procesa en consecuencia. Si la televisión incluye una capacidad de entrada de IP, los paquetes de contenido simplemente se pueden pasar a través de la televisión (como se muestra en la figura 2B). Si no, la unidad de control de la TV puede convertir estos paquetes de Ethernet a un formato de cable coaxial (por ejemplo, a través de DAC 245 QAM y controlador 247 CATV de la figura 2C) de manera que la salida puede ser consumida directamente por un puerto de entrada de la televisión.

De acuerdo con algunas realizaciones, se proporciona un mando a distancia de infrarrojos estándar para que el usuario interactúe con la televisión. Los comandos desde el mando a distancia sean recibidos por la televisión y se transmiten a la unidad de control de la TV a través del mecanismo de paso de infrarrojos de la televisión, normalmente a través de la interfaz serie a la unidad de control de la TV. Según otras realizaciones, la unidad de control de la TV incluye un receptor de control remoto inalámbrico (por ejemplo, los receptores 206, 226, y 246 de las figuras 2A - 2C) de manera que un dispositivo de control no direccional (por ejemplo, un dispositivo de control RF) puede ser empleado para interactuar con el sistema. Además, y de acuerdo con realizaciones específicas, los receptores de RF en las unidades de control de la TV pueden trabajar en conjunción con uno o más dispositivos de RF inalámbricos, incluyendo, por ejemplo, un teclado, un ratón u otro control dispositivo de cursor, etc., para permitir que dichos dispositivos interactúen con el sistema.

En términos generales, la unidad de control de la TV recibe la entrada del usuario desde los dispositivos de control tal como se describe arriba y, dependiendo de la naturaleza de la entrada, puede utilizarla ya sea para controlar la televisión, proporcionar información apropiada a los servidores de servicios de fondo, o ambos. Por ejemplo, si el usuario presiona el botón de "encendido" en el mando a distancia, al recibir este evento de entrada la unidad de control de la TV ordena que la televisión se encienda, y alerta al servidor 106 administrador de sesión del evento para que una instancia VSTB y otros recursos puedan ser instanciados y asignados a la televisión que se acaba de encender. En otro ejemplo, si el usuario pulsa la flecha "arriba" en su mando a distancia, este evento puede ser simplemente reenviado al servidor administrador de sesión de modo que puede efectuarse una actualización correspondiente en la interfaz de usuario virtual con la que el usuario está interactuando. En última instancia, la secuencia de vídeo que se genera por el VSTB asociado reflejará los cambios y el resultado de las acciones del usuario serán visibles en la televisión correspondiente. La manera en que esto se consigue se describe con mayor detalle a continuación.



Varias implementaciones de las unidades de control de la televisión permiten una única solución general del sistema que puede trabajar con una variedad de diferentes televisiones con diferentes capacidades incorporadas, así como diferentes infraestructuras de red; es simplemente una cuestión de la implementación de la unidad de control de la TV apropiada para cada situación específica. Por otra parte, las televisiones que tienen suficientes capacidades incorporadas en combinación con un tipo específico de la red (por ejemplo, basada en IP o el cable coaxial) pueden obviar la necesidad de la unidad de control de la TV en algunas realizaciones, ya sea parcial o totalmente. Para realizaciones sin la unidad de control de la TV, los comandos recibidos desde dispositivos de control en las habitaciones (ya sea un control remoto, teclado, ratón u otro dispositivo) pueden ser transmitidos al servidor 106 administrador de sesión directamente a través de la conexión IP de televisión incorporada en la televisión (por ejemplo, un puerto Ethernet).

La figura 3 es un diagrama de bloques que representa la operación básica de una implementación de software de una instancia de un VSTB de acuerdo con una realización específica de la invención. De acuerdo con diversas realizaciones de la invención, las funcionalidades básicas de un hardware STB convencional se replican ya sea en el ejemplo VSTB o en otros lugares en la arquitectura general del sistema, por ejemplo, en los servidores de apoyo (por ejemplo, el remodelador IPTV), en unidades de control asociadas con los dispositivos de reproducción, en los dispositivos de reproducción, etc. La realización ilustrada en la figura 3 muestra varias funcionalidades tales como ser replicado en la instancia VSTB. Sin embargo, como se ha descrito anteriormente, por lo menos algunas de estas funcionalidades pueden ser descargadas a otros recursos para reducir la carga computacional en el servidor VSTB. Como tal, el alcance de la invención no debe limitarse con referencia a esta figura o su descripción.

Con referencia ahora a la figura 3, la instancia 300 VSTB recibe una emisión 302 sin interrupción de entrada (por ejemplo, una secuencia de medios digitales) desde uno o más de los servidores en el sistema (por ejemplo, servidor 110 de TV lineal, servidor 112 VOD, o remodelador 111 IPTV) sobre una red (por ejemplo, red 102IP), y descifra (304), demultiplexa (306), y decodifica (308 y 309) la secuencia a secuencias de fotogramas de vídeo en bruto (por ejemplo, fotogramas YUV) y tramas de audio, que posteriormente da forma para la salida a la televisión a la que ha sido asignada. Esto puede implicar, por ejemplo, el desentrelazado y escalado (312) al tamaño de fotograma apropiado. Como se mencionó anteriormente y de acuerdo con algunas realizaciones, el uso de contenido VOD codificado previamente desde el servidor VOD y formado previamente de salida de contenido de TV lineal desde el remodelador IPTV puede descargar gran parte de esta carga de cálculo desde la instancia VSTB. También hay que señalar que una secuencia de entrada se muestra en la figura 3 con fines explicativos, pero en realidad, puede haber varias secuencias de entrada siendo todas procesadas en paralelo de acuerdo a las etapas que se muestran en la figura 3. Las secuencias decodificadas entonces serían compuestas junto con los fotogramas de UI (314).

Una interfaz gráfica de usuario (GUI) 314 puede ser superpuesta o compuesta con la secuencia de vídeo decodificada y procesada (316). Esto puede implicar la mezcla pixel por pixel para lograr una interfaz parcial o semitransparente sobre el vídeo subyacente. También puede implicar escalar el vídeo para que aparezca dentro de una ventana en la capa de UI. La etapa de composición también puede implicar varias secuencias de vídeo como se mencionó anteriormente. Los fotogramas de audio decodificados también pueden someterse a la mezcla con otras fuentes de audio (por ejemplo, sonidos emitidos durante la interacción con la superposición de interfaz gráfica de usuario). Los fotogramas de vídeo y de audio resultantes se vuelven a codificar (318, 319), se vuelven a multiplexar (320), se cifran (por ejemplo, utilizando un mecanismo de administración de derechos digitales (DRM) compatible con la televisión asociada) (322), y se transmite como emisión 324 sin interrupción de salida a través de la infraestructura de red de intervención a la televisión correspondiente u otro dispositivo de reproducción para su visualización.

En varias ocasiones, la UI puede no ser visible (por ejemplo, debido a que el usuario está viendo un canal de televisión lineal y no requiere una UI visible). En esos momentos, la etapa 316 de superposición/composición puede ser omitida, ya sea parcial o totalmente. En estas circunstancias, el VSTB puede no tener necesidad de decodificar las subsecuencias de audio y de vídeo y puede pasar las subsecuencias desmultiplexadas a través del conducto 326 solo Remux en lugar de a través de los respectivos conductos de decodificación (308-318). El conducto solo Remux simplemente remultiplexa la subsecuencia de entrada a la secuencia de salida, lo que requiere mucha menos potencia de cálculo. Por otra parte, si la UI es visible y/o requiere composición, el conducto solo Remux no se utiliza debido a que la secuencia de entrada debe ser decodificada a fotogramas de vídeo en bruto para facilitar las etapas de superposición/composición.

Un STB en la habitación convencional transmitiría típicamente su salida a la televisión conectada en un formato no comprimido a través de, por ejemplo, un enlace de HDMI. Sin embargo, debido a que el VSTB no está en la habitación con la televisión, esto no es una opción. En lugar de ello, de acuerdo con diversas realizaciones de la invención, la secuencia de salida generada por el VSTB se transmite a la televisión a través de la red como una secuencia de transporte MPEG-2 de unidifusión cifrada. Es decir, puesto que la secuencia de salida generada puede contener contenido de alto valor, puede ser necesario cifrar la secuencia de salida para la transmisión sobre la red que interviene a la televisión para proteger el contenido de alto valor de la piratería. Cuando se utiliza un STB convencional, esta etapa normalmente se realiza utilizando HDCP sobre el enlace de HDMI a la televisión. Sin embargo, se requiere una solución diferente cuando se utiliza una arquitectura VSTB. Más específicamente, si el cifrado es necesario, de acuerdo con una implementación específica, puede requerirse que la televisión soporte un mecanismo adecuado de gestión de derechos digitales (DRM) internamente de modo que una secuencia cifrada

mediante ese mecanismo puede ser consumida y es descifrada directamente por la televisión. Por ejemplo, un mecanismo de DRM común en el contexto de la hospitalidad es el sistema DRM Pro: Idiom®, una tecnología que puede ser licencia de Zenith Electronics LLC de Lincolnshire, Illinois. Cualquiera que sea el mecanismo DRM soportado, la secuencia de salida generada por una instancia VSTB dada está cifrada de acuerdo con que el mecanismo antes de la transmisión por la red intermedia.

Haciendo de nuevo referencia a la figura 1, la secuencia de salida de unidifusión es transportada por la red 102 de extremo posterior de IP para su posterior transmisión a la televisión correspondiente sobre el equilibrio de la infraestructura de red. Esto puede ser, por ejemplo, una infraestructura de cable coaxial (por ejemplo, planta 119 de cable de hotel), en cuyo caso se proporciona la secuencia de vídeo a Edge QAM Muxer 118. Alternativamente, la secuencia de vídeo puede permanecer en el dominio de IP (cableada o inalámbrica) si el dispositivo de reproducción o la televisión (o la unidad de control de la TV asociada) está conectada a la infraestructura de red IP.

Haciendo referencia de nuevo a la figura 3, del fotograma 314 de UI es efectivamente un cliente del servidor de UI en la que se construye una imagen de la GUI en un momento dado (por ejemplo, un servidor 116 de UI). Cuando fotograma de UI 314 tiene fotogramas de GUI que incluyen los píxeles no transparentes, esta información se inyecta en la canalización de procesamiento de vídeo de superposición/composición con los fotogramas de vídeo.

De acuerdo con diversas realizaciones, una vez que se crea y se asigna una instancia de VSTB debe generar y cifrar de forma continua su secuencia de salida de vídeo para su presentación en el dispositivo de reproducción a el que se asigna. Inicialmente, y en otras ocasiones cuando puede que no haya secuencia de entrada o una interfaz gráfica de usuario disponible, los fotogramas de vídeo negros son sustituidos y estos se utilizan eficazmente para componerse con una capa de memoria intermedia de fotograma gráfico vacío, lo que resulta en un fotograma compuesto negro. De lo contrario, la pantalla del dispositivo de reproducción aparecería como si no se proporcionaba ninguna entrada, por ejemplo, una pantalla de “no hay señal” en una televisión.

Según diversas realizaciones, los componentes de la arquitectura de un VSTB están configurados con un mínimo de memoria intermedia para evitar efectos de latencia no deseados que son perceptibles para el usuario. Es decir, es deseable evitar la situación en la que el usuario es capaz de percibir cualquier retraso en la interacción con la GUI en la televisión con respecto a la experiencia STB convencional en la habitación. Por lo tanto, las realizaciones de la invención se aplican para asegurar que grandes cantidades de datos no necesitan ser amortiguados antes de que se procesen los datos. Es decir, amortiguar grandes cantidades de datos conduce a una inyección de latencia, mientras que la memoria intermedia se llena y se traduce directamente en un retraso significativo en la interacción con la GUI. Además, es deseable proporcionar suficiente potencia de cálculo para ser capaz de generar, procesar y transmitir fotogramas de vídeo a una velocidad suficientemente alta, por ejemplo, 25 o 30 fotogramas por segundo. De acuerdo con realizaciones específicas, las capacidades de procesamiento paralelo de servidores con múltiples núcleos de procesamiento permiten que el procesamiento de los fotogramas de vídeo consecutivos se solape en el tiempo para cumplir con mayor eficacia estas limitaciones, y ser capaces de ejecutar un número suficiente de VSTBs por servidor para que este enfoque económicamente viable.

De acuerdo con varias realizaciones, otras funciones convencionalmente manejadas por los STBs de hardware se replican por varios componentes del sistema. De acuerdo con una realización específica, funciones específicas se manejan de la siguiente manera.

La recepción y procesamiento de eventos de entrada generados por los mandos a distancia, teclados, ratones u otros dispositivos de entrada del usuario se realiza por el firmware de una unidad de control de la TV o directamente por la televisión en conjunción con un servidor administrador de sesión.

La comunicación con los distintos servidores para controlar la transmisión y recepción de la emisión(es) sin interrupción de medios de entrada se realiza por la instancia de software VSTB.

La comunicación con la televisión para controlar la funcionalidad básica, como el nivel de volumen y estado de la alimentación se realiza por el firmware de una unidad de control de la TV, típicamente usando una interfaz de comunicaciones serie para la TV, aunque son posibles otras formas de conexión de control. En algunos casos, las televisiones pueden conectarse directamente a una red IP para fines de control y en tales casos, esta funcionalidad puede ser realizada directamente por la televisión (es decir, los mensajes de control enviados por otros componentes del sistema son recibidos directamente por la televisión a través de una conexión IP) y una unidad de control de la TV puede no ser necesaria.

La interconexión con otros dispositivos conectados, como los ordenadores portátiles de los usuarios, dispositivos USB, etc., se lleva a cabo por el firmware de una unidad de control de la TV, que ofrece un control directo sobre los dispositivos conectados. Estos dispositivos pueden entonces ser accesibles a otros componentes del sistema a través de un protocolo de red adecuado.

La comunicación con los distintos servidores para facilitar el control del STB y los dispositivos conectados, incluyendo la televisión se lleva a cabo por el firmware de una unidad de control de la TV junto con el hardware de comunicaciones previsto en la unidad (por ejemplo, un módem o HPNA conmutador Ethernet). Un protocolo adecuado se define permitiendo que otros componentes del sistema controlen el comportamiento de la unidad de

control televisión y otros dispositivos conectados, incluyendo la televisión. Este protocolo también puede permitir que la unidad de control de la televisión informe a los demás componentes del sistema de la ocurrencia de varios eventos asincrónicos tales como la recepción de la entrada del usuario, o la conexión de un nuevo dispositivo, o un cambio en el estado de alimentación de una televisión.

5 La composición de píxeles de interfaz de usuario con la secuencia de vídeo decodificada, después volver a codificar el vídeo compuesto, puede ser computacionalmente intensivas. Por lo tanto, de acuerdo con algunas realizaciones, esta composición se puede realizar directamente en la secuencia de vídeo codificado usando técnicas de edición de vídeo de dominio comprimido tal como, por ejemplo, las técnicas descritas en Compressed Domain Video Editing, Ragip Kurceren and Fehmi Chebil, ICASSP 2006 (IEEE ©2006). Esto evitaría la necesidad de codificación y  
10 decodificación completas.

Si bien la invención se ha mostrado y descrito particularmente con referencia a realizaciones específicas de la misma, se entenderá por los expertos en la materia que los cambios en la forma y detalles de las realizaciones descritas pueden ser hechos sin apartarse del alcance de la invención. Por ejemplo, las funcionalidades descritas en el presente documento pueden implementarse utilizando instrucciones de programa de ordenador almacenado en la  
15 memoria física, por ejemplo, cualquier tipo de memoria RAM, ROM, u otro medio de almacenamiento legible por ordenador, asociado con o accesible a los dispositivos informáticos particulares que realizan las funciones descritas. Cualquiera de las diferentes herramientas informáticas y de programación y lenguajes conocidos por los expertos en la materia que pueden ser adaptadas para tales fines están dentro del alcance de la presente invención. Alternativamente, por lo menos parte de la funcionalidad descrita puede implementarse utilizando otros tipos de  
20 hardware, tales como, por ejemplo, dispositivos lógicos programables, matrices de puertas programables en campo, circuitos integrados específicos de la aplicación, y similares. Una vez más, las alternativas adecuadas serán evidentes para los expertos en la materia.

Por último, aunque diversas ventajas, aspectos y objetos de la presente invención se han discutidos aquí con referencia a varias realizaciones, se entenderá que el alcance de la invención no debe estar limitado por referencia a  
25 tales ventajas, aspectos y objetos. Más bien, el alcance de la invención debe determinarse con referencia a las reivindicaciones adjuntas.

**REIVINDICACIONES**

1. Un sistema (100) para proporcionar vídeo a través de una red (102), que comprende:

una pluralidad de servidores desplegados en la red y que comprenden uno o más servidores decodificadores virtuales (VSTB) (108) configurados para crear instancias de decodificador virtual (VSTB) y uno o más servidores de interfaz de usuario configurados para crear instancias de interfaz de usuario virtual (VUI) para generar una interfaz gráfica de usuario asociada a cada instancia de decodificador virtual (VSTB); y una pluralidad de clientes VSTB con pantallas (104A, 104B, 104C) asociadas, en el que la pluralidad de servidores está configurada para gestionar la operación de una pluralidad de instancias de decodificador virtual que operan en el uno o más servidores (108) VSTB, correspondiendo cada instancia (300) VSTB a una de la pluralidad de clientes VSTB en la red, estando configurada cada instancia (300) VSTB para combinar la interfaz (314) gráfica de usuario con una o más secuencias de vídeo (302), estando además configurado el o cada servidor VSTB para transmitir la interfaz gráfica de usuario combinada y una o más secuencias de vídeo al cliente VSTB correspondiente como una única secuencia de medios (324) que tiene un formato de fotograma y el codificador-descodificador soportado por el cliente VSTB correspondiente y la pantalla asociada, estando además configurada la pluralidad de servidores para recibir la entrada del usuario que representa la interacción de un usuario asociado con cada cliente VSTB con una representación de la interfaz gráfica de usuario de cada cliente VSTB y proporcionar la entrada del usuario a la instancia de VUI correspondiente a la instancia VSTB correspondiente al cliente VSTB, estando configurado cada cliente VSTB para recibir la única secuencia de medios de la instancia VSTB correspondiente, y para reproducir la única secuencia de medios en la pantalla asociada, estando configurado además cada cliente VSTB para transmitir la entrada del usuario a la pluralidad de servidores, en el que, para cada instancia de VUI, el uno o más servidores de interfaz de usuario (116) están configurados para comunicar las actualizaciones de la interfaz gráfica de usuario desde la instancia de VUI a la instancia VSTB correspondiente a través de una primera conexión informática en red virtual (VNC) con un primer cliente VNC dentro de la instancia VSTB correspondiente y recibir la entrada del usuario para una particular de las interfaces gráficas de usuario en la instancia de VUI a través de una segunda conexión VNC con un segundo cliente VNC correspondiente al cliente VSTB asociado con la pantalla sobre la que se está reproduciendo la interfaz gráfica de usuario en particular, y en el que cada instancia VSTB está configurada para procesar previamente secuencias de medios por separado que recibe durante una sesión para empalmar las secuencias de medios separadas en la única secuencia de medios.

2. El sistema de la reivindicación 1, en el que la pluralidad de servidores comprende además uno o más servidores (111) de transcodificación configurados para recibir por lo menos algunas de las secuencias de vídeo y proporcionar versiones transcodificadas de las secuencias de vídeo recibidas a las instancias VSTB.

3. El sistema de la reivindicación 1, en el que la pluralidad de servidores comprende además uno o más servidores (110, 112) de medios desde los que se originan por lo menos algunas de las secuencias de vídeo.

4. El sistema de la reivindicación 1, en el que la pluralidad de servidores comprende además un servidor (106) administrador de sesión configurado para coordinar la asignación de recursos en el servidor VSTB y el servidor de interfaz de usuario cuando una de la pluralidad de pantallas está encendida y para liberar cualquier recurso asignado durante una sesión cuando la sesión de un usuario está completa.

5. El sistema de la reivindicación 1, en el que la pluralidad de servidores comprende además un servidor (106) administrador de sesión configurado para coordinar los recursos dentro del sistema y en el que el servidor administrador de sesión recibe la entrada del usuario y transmite la entrada del usuario al servidor (116) de interfaz de usuario usando el segundo cliente VNC.

6. El sistema de cualquier reivindicación anterior, en el que la red es una red de hotel.

7. El sistema de la reivindicación 1, en el que por lo menos algunos de los clientes VSTB comprenden una unidad (120A, 120B, 120C) de control separada de un dispositivo de reproducción correspondiente a la pantalla (104a, 104B, 104C) asociada, comprendiendo la unidad de control una primera interfaz configurada para recibir la única secuencia de medios desde el uno o más servidores VSTB, una segunda interfaz configurada para transmitir la única secuencia de medios al dispositivo de reproducción, y los recursos (204, 224, 244) de procesamiento para la comunicación con el dispositivo de reproducción y el uno o más servidores.

8. El sistema de la reivindicación 7, en el que la unidad (120A, 120B, 120C) de control comprende además una tercera interfaz configurada para interactuar con y controlar la operación del dispositivo de reproducción.

9. El sistema de la reivindicación 1, en el que las pantallas asociadas con la pluralidad de clientes VSTB corresponden a cualesquiera de una televisión, un ordenador de sobremesa, un ordenador portátil, un dispositivo portátil, una tableta, un dispositivo de mano, un equipo ultraportátil o un teléfono inteligente.

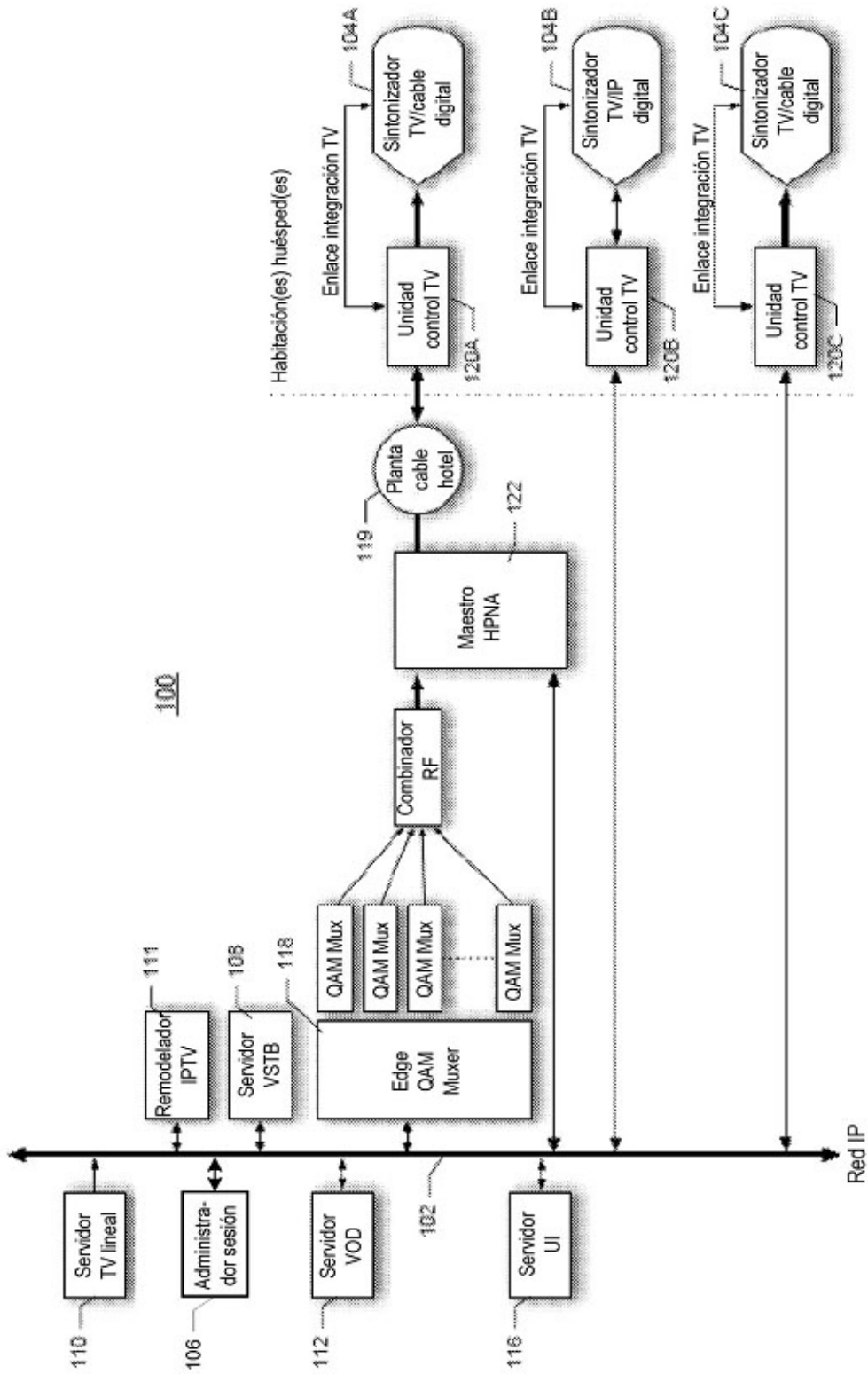


FIG. 1

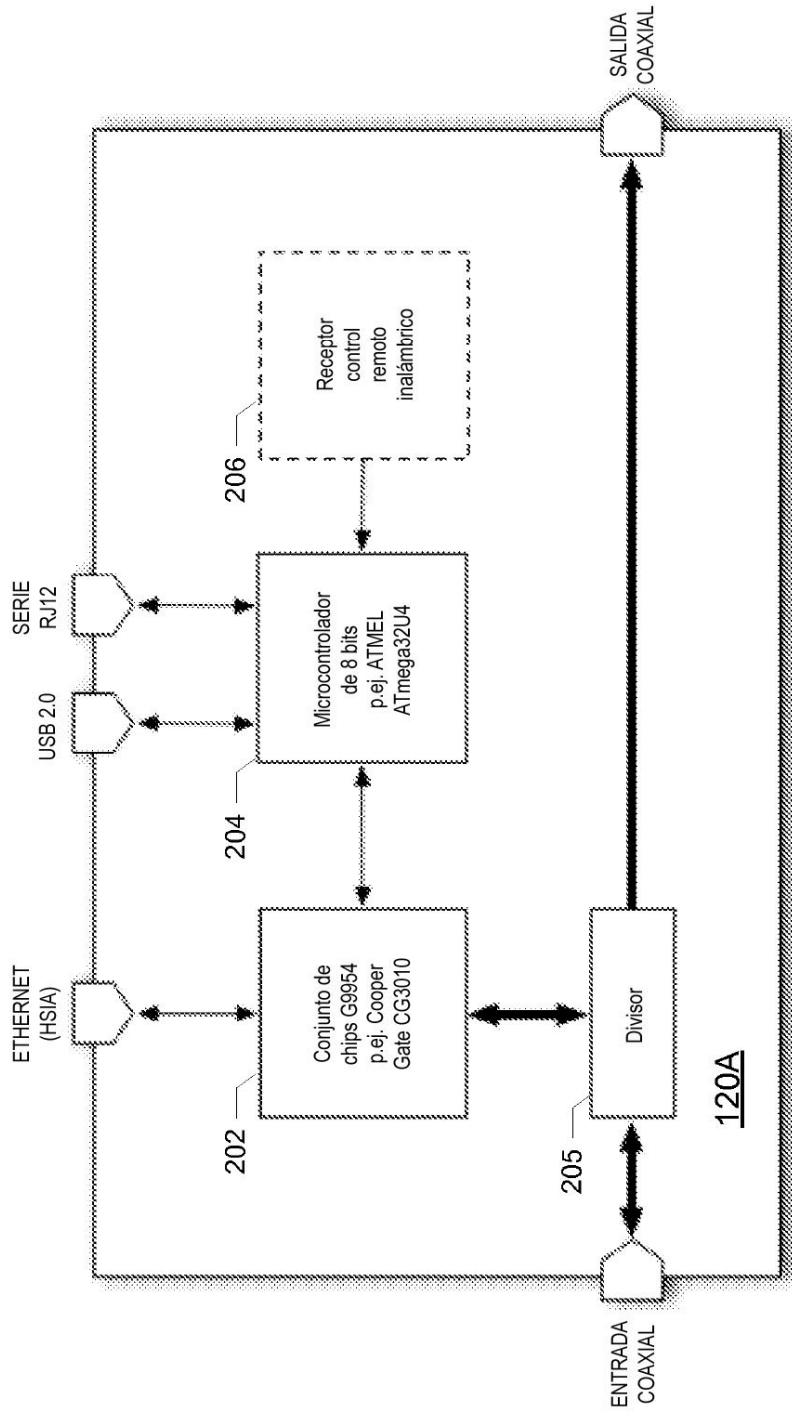
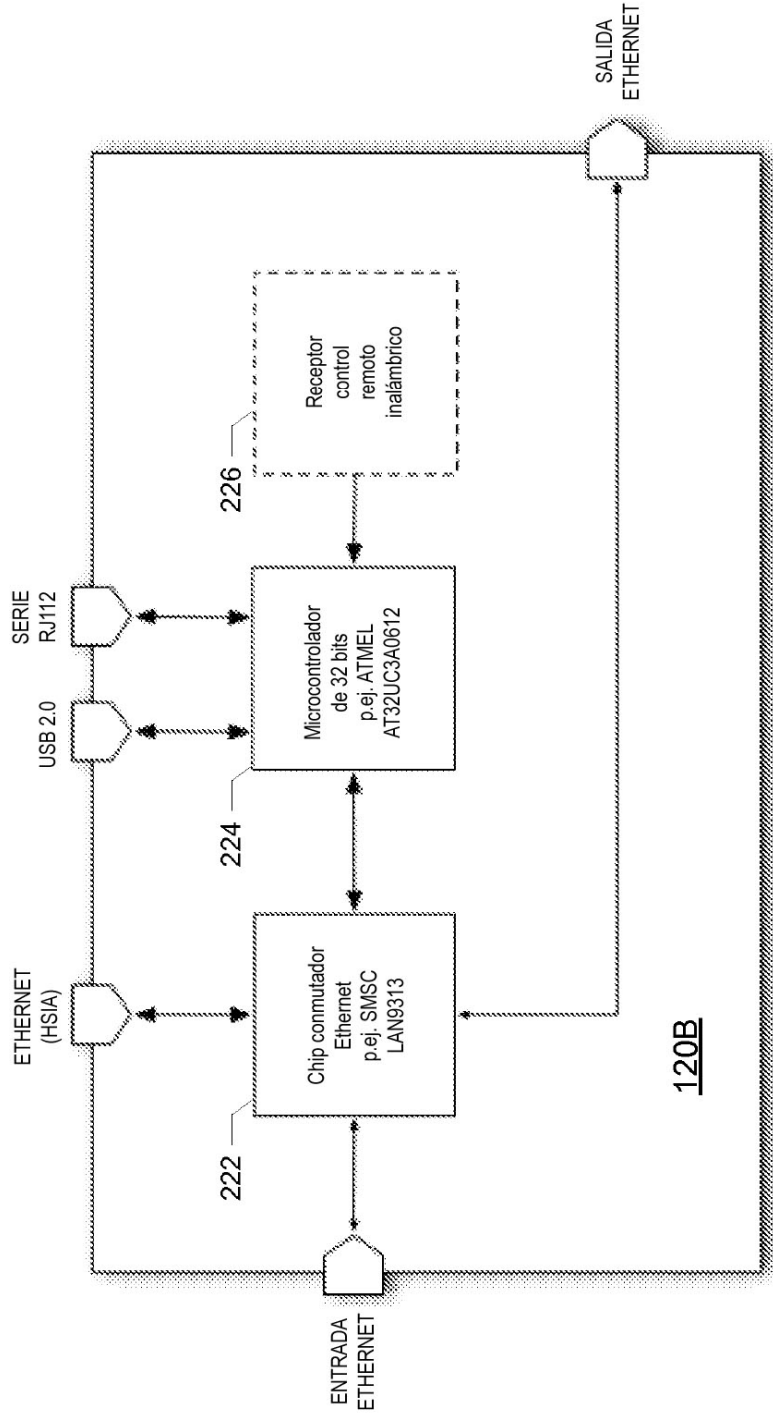


FIG. 2A



**FIG. 2B**

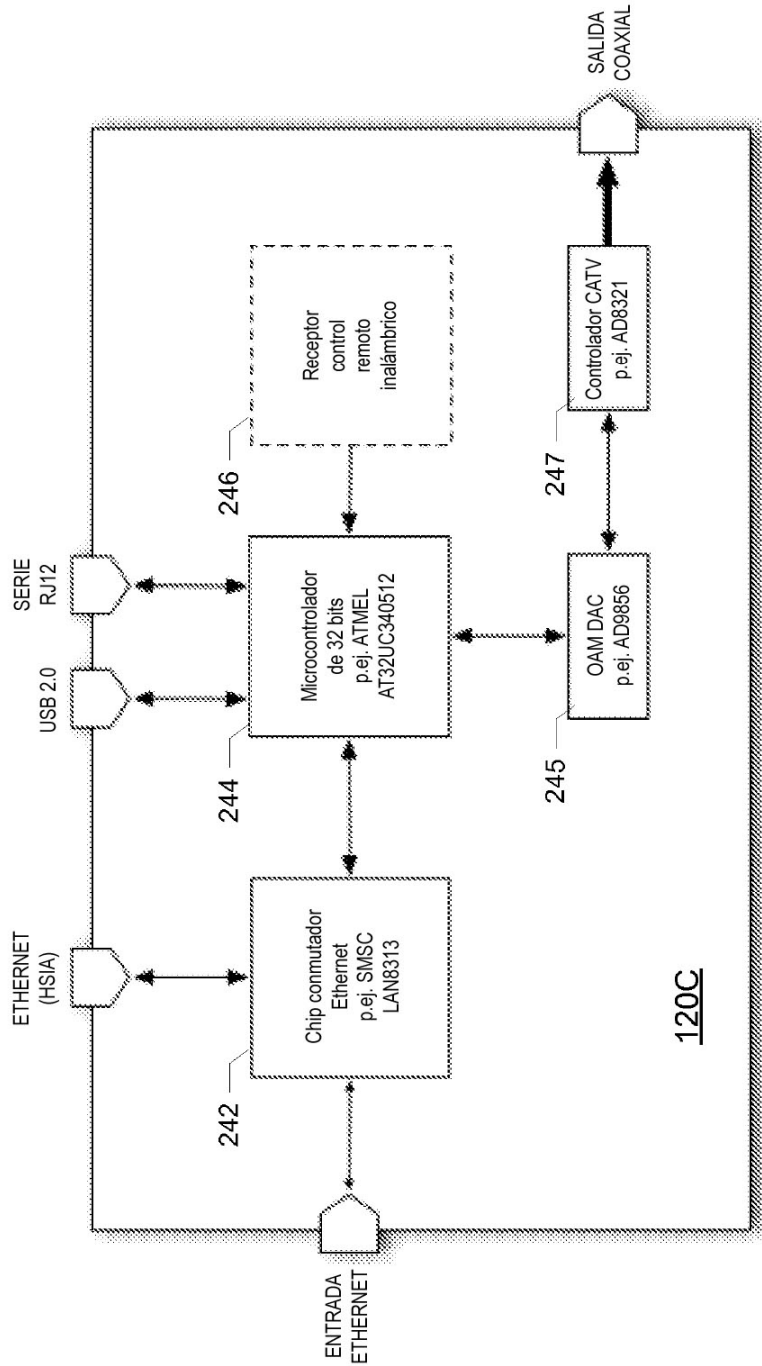


FIG. 2C



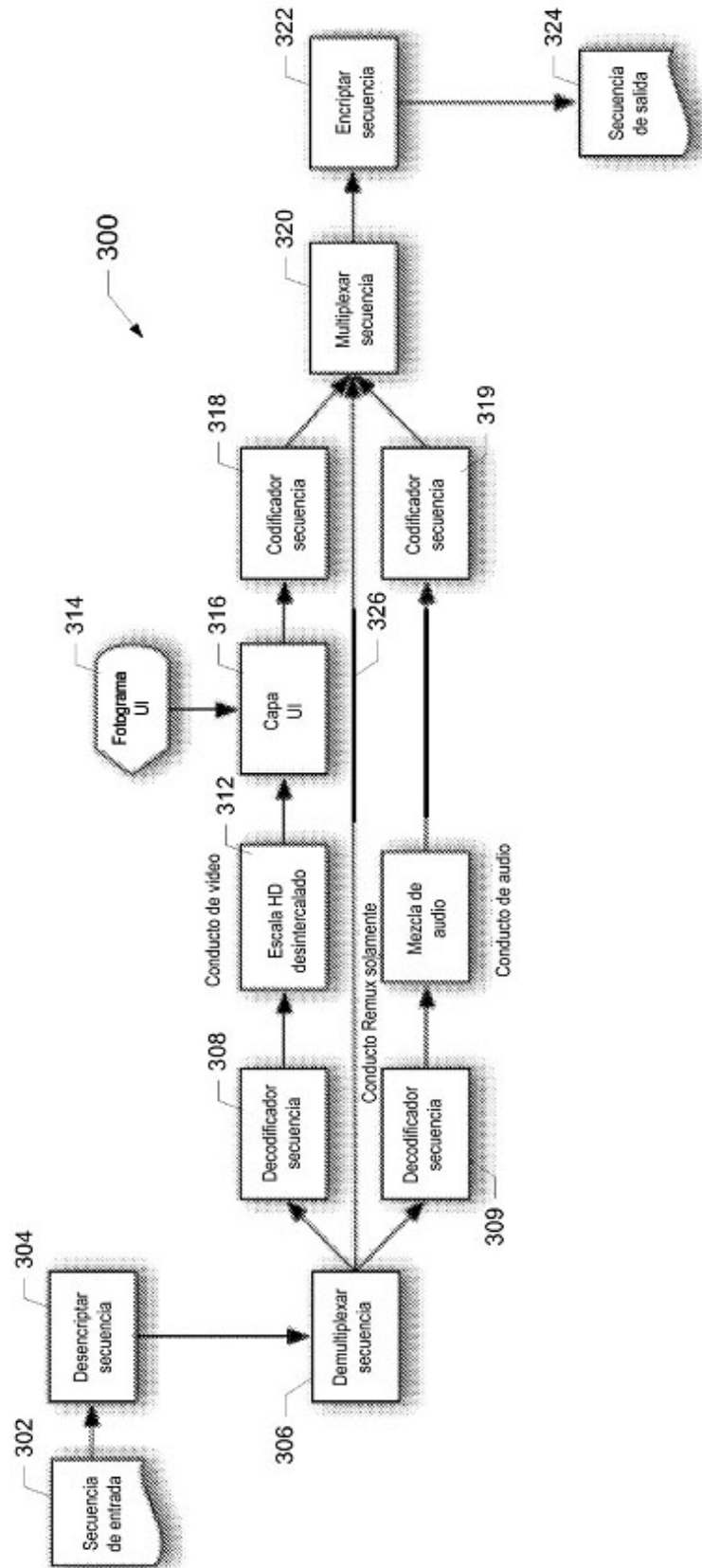


FIG. 3