

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 532 829**

51 Int. Cl.:

H04N 19/34 (2014.01)

H04N 21/6377 (2011.01)

H04N 21/658 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.06.1999 E 99923815 (7)**

97 Fecha y número de publicación de la concesión europea: **24.12.2014 EP 1040668**

54 Título: **Sistema de codificación de vídeo escalable**

30 Prioridad:

06.07.1998 US 110616

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

01.04.2015

73 Titular/es:

**FUNAI ELECTRIC CO., LTD. (100.0%)
7-7-1, Nakagaito, Daito
Osaka 574-0013, JP**

72 Inventor/es:

**RADHA, HAYDER;
CHEN, YINGWEI y
COHEN, ROBERT A.**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 532 829 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

Sistema de codificación de vídeo escalable

Antecedentes de la invención**Campo de la invención**

- 5 La presente invención versa sobre un sistema de codificación de vídeo escalable que codifica datos de vídeo usando tanto predicción de tramas como imágenes escalables de granularidad fina. La invención tiene utilidad particular en conexión con redes de ancho de banda variable y en sistemas de ordenadores que son capaces de acomodar diferentes tasas de bits y, por ende, imágenes de diferente calidad.

Descripción de la técnica relacionada

- 10 La codificación de vídeo escalable, en general, se refiere a técnicas de codificación que son capaces de proporcionar diferentes niveles o cantidades de datos por trama de vídeo. En la actualidad tales técnicas son usadas por los estándares punteros de codificación, tales como MPEG-2 y MPEG-4 (es decir, codificación del grupo de expertos de imágenes en movimiento) para proporcionar flexibilidad cuando se da salida a datos de vídeo codificados.

- 15 En las técnicas de codificación escalable actualmente empleadas por MPEG-2 y MPEG-4, un codificador codifica tramas de datos de vídeo y divide las tramas codificadas en una capa base (BL) y una capa de mejora (EL). Normalmente, la capa base comprende una cantidad mínima de datos requerida para decodificar los datos de vídeo codificados. La capa de mejora, por otra parte, comprende información adicional que realza (por ejemplo, mejora la calidad de) la capa base cuando es decodificada. En funcionamiento, el codificador transmite todas las tramas desde
20 la capa base a un dispositivo receptor, que puede ser un ordenador personal o similar. Sin embargo, el codificador únicamente transmite tramas desde la capa de mejora en los casos en que el dispositivo receptor tiene suficiente potencia de proceso para gestionar esas tramas adicionales y/o el medio sobre el que se transmiten las tramas tiene suficiente ancho de banda.

- 25 Las Figuras 1 y 2 muestran estructuras de escalabilidad que se usan en la actualidad en MPEG-2 y MPEG-4 para la capa base y la capa de mejora. Más específicamente, la Figura 1 muestra una estructura 1 de escalabilidad que emplea predicción de tramas en la capa base 2 para generar tramas predictivas (o P) a partir de una intratrama (o I) o de una trama P precedente. Según se muestra en la figura, la predicción de tramas también se usa en la capa de mejora para generar tramas P basadas en tramas de la capa base. La Figura 2 muestra otra estructura 3 de escalabilidad que se usa en la actualidad en MPEG-2 y MPEG-4. En la estructura de escalabilidad mostrada en la
30 Figura 2, vuelve a emplearse la predicción de tramas para determinar tramas P en la capa base. Sin embargo, a diferencia de la estructura 1 de escalabilidad, la estructura 3 de escalabilidad también usa la predicción de tramas en la capa de mejora para generar tramas bidireccionales (o B) que, en este caso, se interpolan a partir de tramas precedentes de la capa de mejora y de tramas contemporáneas de la capa base. En general, los codificadores MPEG-2 y MPEG-4 usan la predicción de tramas de la manera presentada en lo que antecede para aumentar la
35 compresión de datos y aumentar así la eficiencia de la codificación.

El documento EP 634 871 A2 describe un codificador de vídeo de múltiples capas que comprende un codificador de la capa base y un codificador de la capa de mejora, que proporcionan dos tipos de señales de vídeo codificadas que comparten un canal común de salida. Un receptor de vídeo puede usar uno o ambos de los dos tipos de señal de vídeo codificada.

- 40 El documento WO 94 11993 A1 describe un procedimiento para mejorar la interactividad de la información multimedia. Una información multimedia está separada en capas primarias y secundarias. La información de las capas primarias es mejorada de modo que mejora la calidad psicográfica percibida de la información.

- El documento WO 95 01052 A1 describe un procedimiento para incorporar futuros desarrollos en la tecnología de compresión de vídeo para migrar hacia un mejor sistema de televisión de manera compatible con los receptores. Los
45 datos de mejora pueden combinarse con datos estándar de vídeo. Los receptores estándar de HDTV utilizarán datos estándar de vídeo, ignorando los datos de mejora. Los receptores (36) más avanzados de HDTV combinarán los datos de mejora con los datos estándar.

- YINGWEI CHEN ET AL: "Request for Fine Granular Video Scalability for Media Streaming Applications", M3792, 30 de junio de 1998 (1998-06-30), XP030033064, ISSN: 0000-0311 describe un procedimiento de codificación de vídeo
50 que produce una capa base usando una técnica de codificación de predicción de tramas, generando un residuo de la diferencia entre los datos de vídeo introducidos y los datos reconstruidos de la capa base y produciendo una capa de mejora codificando los residuos usando una técnica de codificación de escalabilidad de granularidad fina.

- Otra técnica de codificación de vídeo escalable bien conocida se denomina codificación de escalabilidad de granularidad fina. La codificación de escalabilidad de granularidad fina codifica la misma imagen (por ejemplo, una
55 trama de vídeo) usando progresivamente más datos cada vez que tiene lugar la codificación. Por ejemplo, según la

muestra en la Figura 3, la imagen 4 es codificada inicialmente usando datos suficientes para producir la imagen 5. Posteriormente, se codifican datos adicionales que son suficientes para producir las imágenes mejoradas 6, 7 y 8 en sucesión.

5 La codificación de escalabilidad de granularidad fina tiene varias ventajas con respecto a las técnicas de predicción de tramas descritas más arriba. específicamente, dado que la codificación de escalabilidad de granularidad fina puede proporcionar una gama más amplia de imágenes mejoradas que las técnicas de predicción de tramas, generalmente se prefiere la codificación de escalabilidad de granularidad fina en entornos, tales como Internet, que tienen disponible un gran intervalo de ancho de banda. Por razones similares, también se prefiere generalmente la codificación de escalabilidad de granularidad fina cuando se abordan dispositivos receptores que tienen capacidad de procesamiento y/o ancho de banda diferentes. Es decir, dado que la codificación de escalabilidad de granularidad fina produce una amplia gama de imágenes mejoradas, es posible ajustar la imagen apropiada de forma relativamente cercana a una cantidad de ancho de banda disponible. En consecuencia, en teoría, es posible obtener la máxima cantidad de datos para una imagen para una cantidad dada de ancho de banda disponible. En la parte negativa, la codificación de escalabilidad de granularidad fina no permite el uso de la predicción de tramas. En consecuencia, requiere más datos que las técnicas de predicción de tramas descritas más arriba y, en consecuencia, degrada la eficiencia de codificación.

Así, existe la necesidad de una técnica de codificación de vídeo escalable que incorpore la eficiencia de la codificación de predicción de tramas y la precisión de la codificación de escalabilidad de granularidad fina.

La invención se define en las reivindicaciones adjuntas a la presente memoria.

20 **Breve descripción de los dibujos**

La Figura 1 representa una estructura de escalabilidad usada con una técnica de codificación escalable de tipo convencional de predicción de tramas.

La Figura 2 representa una estructura alternativa de escalabilidad usada con una técnica de codificación escalable de tipo convencional de predicción de tramas.

25 La Figura 3 representa imágenes generadas usando una técnica de codificación/decodificación de escalabilidad de granularidad fina.

La Figura 4 representa un sistema de ordenador en el que puede implementarse la presente invención.

La Figura 5 representa la arquitectura de un ordenador personal en el sistema de ordenador mostrado en la Figura 4.

30 La Figura 6 es un diagrama de bloques funcionales que muestra elementos de las realizaciones primera y segunda de la presente invención.

La Figura 7 es un diagrama de flujo que describe la técnica de codificación de escalabilidad de la presente invención.

La Figura 8 muestra una estructura de escalabilidad generada por la presente invención.

35 La Figura 9 es un diagrama de bloques de un decodificador según la presente invención.

La Figura 10 es un gráfico que representa la calidad de imagen en función de la tasa de bits para corrientes de bits de radiodifusión simultánea generadas por la segunda realización de la presente invención.

La Figura 11 es un diagrama de bloques funcionales que muestra elementos de la tercera realización de la presente invención.

40 **Descripción detallada de las realizaciones preferentes**

La Figura 4 muestra una realización representativa de un sistema 9 de ordenador en la cual puede implementarse la presente invención. Según se muestra en la Figura 4, un ordenador personal (PC) 10 incluye una conexión 11 de red para conectarse a una red, tal como una red de ancho de banda variable o Internet, y una conexión 12 de fax/módem para conectarse con otras fuentes remotas tales como una cámara de vídeo (no mostrada). El PC 10 también incluye una pantalla 14 de visualización para mostrar información (incluyendo datos de vídeo) a un usuario, un teclado 15 para introducir texto e instrucciones de usuario, un ratón 13 para situar un cursor en la pantalla 14 de visualización y para introducir instrucciones de usuario, una unidad 16 de disco para leer disquetes instalados en la misma y para escribir en los mismos, y una unidad 17 de CD-ROM para acceder a información almacenada en CD-ROM. El PC 10 también puede tener uno o más dispositivos periféricos conectados al mismo, tales como un escáner (no mostrado) para introducir imágenes de texto de documentos, imágenes gráficas o similares, y una impresora 19 para dar salida a imágenes, texto o similares.

La Figura 5 muestra la estructura interna del PC 10. Según se muestra en la Figura 5, el PC 10 incluye memoria 20, que comprende un medio legible por ordenador tal como un disco duro de ordenador. La memoria 20 almacena datos 23, aplicaciones 25, un controlador 24 de impresora y un sistema operativo 26. En realizaciones preferentes de la invención, el sistema operativo 26 es un sistema operativo de ventanas tal como Microsoft7 Windows 95, aunque la invención también puede ser usada con otros sistemas operativos. Entre las aplicaciones almacenadas en la memoria 20 figuran el codificador 21 de vídeo escalable y el decodificador 22 de vídeo escalable. El codificador 21 de vídeo escalable lleva a cabo la codificación escalable de datos de vídeo de la manera explicada con detalle más abajo, y el decodificador 22 de vídeo escalable decodifica datos de vídeo que han sido codificados de la manera

ordenada por el codificador 21 de vídeo escalable. La operación de estas aplicaciones se describe con detalle más abajo.

También están incluidos en el PC 10 una interfaz 29 de visualización, una interfaz 30 de teclado, una interfaz 31 de ratón, una interfaz 32 de la unidad de disco, una interfaz 34 de la unidad de CD-ROM, el bus 36 de ordenador, la RAM 37, un procesador 38 y una interfaz 40 de impresora. Preferentemente, el procesador 38 comprende un microprocesador o similar para ejecutar aplicaciones, como aquellas que se ha hecho notar anteriormente, de la RAM 37. Tales aplicaciones, que incluyen el codificador 21 de vídeo escalable y el decodificador 22 de vídeo escalable, pueden estar almacenadas en la memoria 20 (según se ha hecho notar anteriormente) o, alternativamente, en un disquete en la unidad 16 de disco o en un CD-ROM en la unidad 17 de CD-ROM. El procesador 38 accede a las aplicaciones (o a otros datos) almacenados en un disquete a través de la interfaz 32 de la unidad de disco y accede a las aplicaciones (o a otros datos) almacenados en un CD-ROM a través de la interfaz 34 de la unidad de CD-ROM.

La ejecución de aplicaciones y otras tareas del PC 10 pueden ser iniciadas usando el teclado 15 o el ratón 13, transmitiéndose las instrucciones de los mismos al procesador 38 a través de la interfaz 30 de teclado y de la interfaz 31 de ratón, respectivamente. Los resultados de salida de las aplicaciones que se ejecutan en el PC 10 pueden ser procesados por la interfaz 29 de visualización y luego mostrados a un usuario en el dispositivo 14 de visualización o, alternativamente, ser objeto de salida a través de la conexión 11 de red. Por ejemplo, los datos de vídeo introducidos que han sido codificados por el codificador 21 de vídeo escalable son normalmente objeto de salida a través de la conexión 11 de red. Por otro lado, los datos de vídeo codificados que se han recibido, por ejemplo, de una red de ancho de banda variable son decodificados por el decodificador 22 de vídeo escalable y luego presentados en el dispositivo 14 de visualización. Con este fin, la interfaz 29 de visualización comprende preferentemente un procesador de visualización para formar imágenes de vídeo basadas en los datos de vídeo decodificados proporcionados por el procesador 38 por el bus 36 de ordenador, y para dar salida a esas imágenes al dispositivo 14 de visualización. Los resultados de salida de otras aplicaciones, tales como programas de tratamiento de textos, que se ejecuten en el PC 10 pueden ser proporcionados a la impresora 19 a través de la interfaz 40 de impresora. El procesador 38 ejecuta el controlador 24 de impresora para realizar el formateo apropiado de tales trabajos de impresión antes de su transmisión a la impresora 19.

Primera realización

Pasando al codificador 21 de vídeo escalable, este módulo comprende etapas de procedimiento ejecutables por ordenador que codifican datos de vídeo que comprenden una o más tramas sucesivas. En breve, estas etapas de procedimiento codifican una porción de los datos de vídeo usando una técnica de codificación de predicción de tramas, generan imágenes residuales basadas en los datos de vídeo y los datos de vídeo codificados, y codifican las imágenes residuales usando una técnica de codificación de escalabilidad de granularidad fina. Las etapas envían entonces los datos de vídeo codificados y al menos una de las imágenes residuales codificadas a un receptor que, en general, puede comprender una red (de ancho de banda variable o no), un PC u otros dispositivos de soporte de vídeo conectables en red que incluyen, sin limitación, televisores digitales/convertidores de señal de televisión por cable y equipos de videoconferencia.

La Figura 6 es un diagrama de bloques que representa una fuente 42 de vídeo, una red 43 de ancho de banda variable y módulos usados para efectuar las anteriores etapas del procedimiento. La Figura 7 es un diagrama de flujo que explica la funcionalidad de los módulos mostrados en la Figura 6. Para empezar, en la etapa S701 se introducen datos de vídeo originales no codificados en el codificador de la presente invención. Estos datos de vídeo pueden ser introducidos a través de la conexión 11 de red, de la conexión 12 de fax/módem, o, según se muestra en la Figura 6, a través de una fuente de vídeo. Para los fines de la presente invención, la fuente 42 de vídeo puede comprender cualquier tipo de dispositivo de captura de vídeo, un ejemplo del cual es una cámara de vídeo digital. Según se muestra en la Figura 6, los datos de vídeo procedente de la fuente de vídeo son introducidos tanto en el codificador 44 de la BL como en el bloque 45 de cálculo de imágenes residuales. La razón de esto se hace evidente más abajo.

A continuación, la etapa S702 codifica una porción (es decir, una capa base, o BL) de los datos de vídeo originales usando una técnica estándar de codificación de predicción de tramas. La etapa S702 la lleva a cabo el codificador 44 de la BL, el cual, en realizaciones preferentes de la invención, es un codificador MPEG-1, MPEG-2 o MPEG-4. Se proporciona una visión general del estándar MPEG en "MPEG: A Video Compression Standard For Multimedia Applications", de Didier LeGall, Communications of the ACM, Vol. 34, N° 4 (abril de 1991). El codificador 44 de la BL comprime los datos de vídeo con una tasa predeterminada de bits, R_{BL} . En realizaciones preferentes de la invención, el bloque 48 de cálculo determina R_{BL} en función del ancho de banda actual de un receptor, tal como la red 43 de ancho de banda variable (o, por ejemplo, un sistema de ordenador que tenga prestaciones de procesamiento variables).

Más específicamente, el bloque 48 de cálculo mide una tasa mínima de bits (R_{MIN}), una tasa máxima de bits (R_{MAX}) y un ancho de banda disponible (R) en la actualidad de la red 43 de ancho de banda variable. El bloque 48 de cálculo fija entonces R_{BL} en un valor entre R_{MIN} y R . En la mayoría de los casos, el bloque 48 de cálculo fija R_{BL} en R_{MIN} ,

para garantizar que, incluso en sus anchos de banda más bajos, la red 43 de ancho de banda variable pueda acomodar la salida de datos de vídeo codificados por la presente invención. Esto es especialmente así en los casos en los que la codificación de la capa base tiene lugar sin conexión.

La Figura 8 muestra un ejemplo de una estructura de escalabilidad que es generada por la presente invención. Según se muestra en la Figura 8, esta estructura de escalabilidad incluye tanto una capa base (BL) como una capa de mejora (EL). La capa base 47 incluye tramas, tal como la trama 49. El codificador 44 de la BL comprime estas tramas con una tasa de bits de R_{BL} . Sin embargo, la capa de mejora 50 incluye imágenes codificadas de granularidad fina correspondientes a tramas contemporáneas de la capa base. Lo que sigue describe cómo genera la invención la capa 50 de mejora. Más específicamente, la etapa S703 genera imágenes residuales 51 basadas en los datos de vídeo originales introducidos desde la fuente 42 de vídeo y basadas en los datos de vídeo codificados (es decir, la capa base) proporcionados por el codificador 44 de la BL. En el diagrama de bloques mostrado en la Figura 6, la etapa S703 la lleva a cabo el bloque 45 de cálculo de imágenes residuales. En funcionamiento, el bloque 45 de cálculo de imágenes residuales recibe datos de vídeo codificados del codificador 44 de la BL y luego decodifica esos datos de vídeo codificados. Después se generan imágenes residuales 51 basadas en la diferencia entre píxeles en estos datos de vídeo decodificados y píxeles de los datos de vídeo originales. En general, las imágenes residuales corresponden a la diferencia entre tramas de la capa base (que comprende el mínimo número de tramas y/o la cantidad mínima de datos requeridos por un decodificador para decodificar una señal de vídeo) y las tramas de los datos de vídeo originales.

El bloque 45 de cálculo de imágenes residuales puede usar uno o más de varios procedimientos diferentes para generar imágenes residuales 51. Por ejemplo, en una realización de la invención, se realiza una simple resta píxel a píxel entre tramas de la capa base y tramas de los datos de vídeo originales. La diferencia resultante entre estos dos conjuntos de tramas (es decir, las imágenes residuales) incluye diferencias en las resoluciones de las tramas. En casos en los que la capa base no incluye tramas enteras de los datos de vídeo originales, las imágenes residuales incluyen estas tramas ausentes.

En otra realización de la invención, el bloque 45 de cálculo de imágenes residuales genera imágenes residuales 51 filtrando en primer lugar los datos de vídeo decodificados y luego determinando una diferencia entre estos datos de vídeo filtrados y los datos de vídeo originales. Esta técnica tiene la ventaja de eliminar ruido no deseado y similares de los datos de vídeo decodificados causados, por ejemplo, por los procesos de codificación y decodificación. En realizaciones preferentes de la invención, se usa un filtro de desbloqueo para filtrar los datos de vídeo decodificados, aunque la invención no esté limitada al uso de este tipo de filtro.

En otra realización adicional de la invención, el bloque 45 de cálculo de imágenes residuales genera imágenes residuales 51 filtrando tanto el vídeo decodificado como los datos de vídeo originales, y determinando luego una diferencia entre ambos tipos de datos filtrados. En esta realización puede aplicarse el mismo tipo de filtro (por ejemplo, un filtro de desbloqueo) tanto a los datos de vídeo originales como a los datos de vídeo decodificados. Alternativamente, pueden aplicarse tipos de filtros diferentes a los datos de vídeo originales y a los datos de vídeo decodificados.

En general, cuando se usa filtrado para generar imágenes residuales 51, un decodificador para recibir datos de vídeo que se hayan codificado según la presente invención debería estar sincronizado con el tipo de filtrado usado por los mismos, lo que quiere decir que debería aplicarse sustancialmente el mismo tipo de filtrado en el decodificador para compensar los efectos del filtrado. Por ejemplo, si las imágenes residuales 51 están codificadas basándose en datos de vídeo filtrados decodificados, debería aplicarse ese mismo filtrado a las imágenes residuales durante la decodificación de las mismas.

Volviendo a la Figura 7, después de la etapa S703, el procesamiento prosigue a la etapa S704. La etapa S704 codifica las imágenes residuales usando una técnica embebida de codificación de escalabilidad de granularidad fina, según se muestra en la capa de mejora de la estructura de escalabilidad de la Figura 8. En la realización de la invención mostrada en la Figura 6, esta etapa la lleva a cabo el codificador 54 de la EL escalable de granularidad fina. El codificador 54 de la EL codifica imágenes residuales 51 con una tasa de bits de $R_{MAX}-R_{BL}$ (es decir, la diferente en el ancho de banda de la capa base y el máximo ancho de banda de la red 43) usando una técnica de codificación de granularidad fina. En este punto se hace notar que, dado que se usa una técnica de escalado de granularidad fina para codificar tramas para la capa de mejora, no se emplea en la misma la predicción de tramas.

Según se muestra en la Figura 6, los valores para R_{MAX} y R_{BL} son proporcionados al codificador 54 de la EL por el bloque 48 de cálculo. El codificador 54 de la EL puede usar varias técnicas de codificación de granularidad fina bien conocidas. Ejemplos de estas incluyen una técnica embebida de transformada discreta del coseno (DCT) y una técnica escalable de búsqueda de coincidencias (MP). Sin embargo, las realizaciones preferentes de la invención usan una de la familia de transformadas ondulatorias (por ejemplo, transformadas ondulatorias de árbol nulo) para efectuar la codificación de la capa de mejora. Por ejemplo, la realización preferente de la invención usa la técnica de codificación de imágenes fijas proporcionada en MPEG-4 para llevar a cabo la codificación de escalabilidad de granularidad fina. Este enfoque codifica imágenes en conjunto usando transformadas ondulatorias.

Con independencia del tipo de codificación de escalabilidad de granularidad fina que use el codificador 54 de la EL, produce una corriente de bits de la EL que tiene una tasa de bits de $R_{MAX}-R_{BL}$. Esta corriente de bits de la EL comprende varias imágenes escalables embebidas de granularidad fina, lo que quiere decir que la corriente de bits comprende una imagen basta inicial y una o más mejoras de la misma. por ejemplo, la corriente de bits de la EL puede incluir una imagen basta que comprenda un número predeterminado de bits (por ejemplo, los 100 primeros bits) en la corriente de bits; una imagen de mejora que comprenda la imagen basta y el siguiente número predeterminado de bits (por ejemplo, los siguientes 100 bits) en la corriente de bits; una imagen adicional de mejora que comprenda la imagen basta, la imagen de mejora, y el siguiente número predeterminado de bits (por ejemplo, los siguientes 100 bits) en la corriente de bits; etcétera. Se denomina granularidad de la imagen al número de bits usados para mejorar estas imágenes (100 bits en este ejemplo).

En este punto se hace notar que la presente invención no está limitada al uso de una granularidad de 100 bits, ni siquiera al uso del mismo número de bits para mejorar la imagen. De hecho, la granularidad usada por la invención puede variar y, en realizaciones preferentes, puede bajar hasta el nivel de byte o incluso hasta el nivel de un solo bit, en el que se usan bits individuales para mejorar una imagen.

Según se muestra en la Figura 6, se proporciona la corriente de bits de la EL al controlador 55 de la tasa de vídeo escalable en tiempo real, que lleva a cabo, en tiempo real, las etapas S705 y S706 mostradas en la Figura 7. En la etapa S705, el controlador 55 recibe R_{BL} , R_{MAX} y R del bloque 48 de cálculo, y luego selecciona, para cada trama de la capa base, una o más de las imágenes residuales codificadas de la capa 50 de mejora (véase la Figura 8) en función de estos valores. En particular, el controlador 55 selecciona una o varias imágenes de la capa de mejora que tengan un ancho de banda que sustancialmente corresponda a $R-R_{BL}$, es decir, la diferencia entre el ancho de banda real de la red 43 y el ancho de banda de la capa base. El controlador 55 selecciona estas imágenes transmitiendo imágenes de la corriente de bits de la EL (por ejemplo, una imagen basta y/o mejoras de imágenes) que tengan un ancho de banda que corresponda a $R-R_{BL}$, y bloqueando la transmisión de aquellas mejoras de imagen que queden fuera de ese intervalo. Al implementar la invención usando una granularidad relativamente fina, tal como una granularidad de un solo bit, la invención es capaz de llenar sustancialmente todo el ancho de banda entre R y R_{BL} . En estos casos, la invención es capaz de proporcionar sustancialmente la cantidad máxima de datos de vídeo para la cantidad dada de ancho de banda disponible. Por supuesto, en casos en los que el receptor pueda gestionar únicamente imágenes codificadas de la capa base, el controlador 55 no transmitirá ninguna imagen escalable de granularidad fina de la capa de mejora.

Sin embargo, suponiendo que hayan de transmitirse estas imágenes, una vez que el controlador 55 haya seleccionado las imágenes escalables de granularidad fina apropiadas (es decir, imágenes residuales codificadas), el procesamiento prosigue a la etapa S706. En la etapa S706, el controlador 55 da salida a la capa base y a las imágenes escalables de granularidad fina seleccionadas en la etapa S705. Según se muestra en la Figura 6, las imágenes son enviadas a la red 43 de ancho de banda variable como una corriente de BL y una corriente de EL.

Un decodificador, un diagrama de bloques funcionales del cual se muestra en la Figura 9, recibe entonces estas corrientes de bits y decodifica los datos que hay en las mismas. El decodificador 57 puede comprender un PC, tal como el mostrado en la Figura 4 o, alternativamente, cualquiera de los otros receptores mencionados más arriba. Según se muestra en la figura, el decodificador 57 incluye un módulo decodificador 58 de vídeo escalable que es ejecutado por un procesador en el mismo. Este módulo decodificador de vídeo escalable comprende un módulo 59 de decodificación de la EL escalable de granularidad fina para decodificar datos en la corriente de bits de la EL y un módulo 60 de decodificación de la BL de predicción de tramas para decodificar tramas en la corriente de bits de la BL. En realizaciones preferentes de la presente invención, el módulo 60 de decodificación de la BL comprende un módulo de decodificación MPEG-1, MPEG-2 o MPEG-4. Debido a la granularidad fina de la corriente de bits de la EL, el decodificador de la EL puede decodificar cualquier porción apropiada de la corriente de bits de la EL limitado, por ejemplo, por restricciones de procesamiento del decodificador o similares. Una vez que los respectivos módulos de decodificación han decodificado las corrientes de datos de vídeo, el bloque 61 de procesamiento añade y, si es necesario, reordena tramas de las mismas. Estas tramas pueden ser mostradas entonces a un usuario.

Segunda realización

La segunda realización de la presente invención genera una estructura de escalabilidad como la mostrada en la Figura 8 para cada una de varias corrientes de bits de radiodifusión simultánea. Brevemente, en la segunda realización de la presente invención, el codificador 21 de vídeo escalable incluye etapas de procedimiento ejecutables por ordenador para codificar una porción (por ejemplo, la capa base) de datos de vídeo introducidos con varias tasas diferentes de bits para producir múltiples versiones de datos de vídeo codificados, para generar varias imágenes residuales para cada versión de los datos de vídeo codificados, para codificar las varias imágenes residuales para cada versión de los datos de vídeo codificados usando una técnica de codificación de escalabilidad de granularidad fina, y luego para dar salida a una versión (por ejemplo, una capa base) de los datos de vídeo codificados junto con una o más imágenes residuales codificadas al efecto.

Más específicamente, en esta realización de la invención, el codificador 44 de la BL codifica la capa base con varias tasas diferentes de bits R_{B1} , R_{B2} , R_{B3} ... R_{BN} , siendo $R_{MIN} < R_{B1} < R_{B2} < R_{B3} \dots < R_{BN} < R_{MAX}$.

Para cada una de estas corrientes codificadas resultantes de radiodifusión simultánea, el bloque 45 de cálculo de imágenes residuales genera imágenes residuales de la manera descrita más arriba. Posteriormente, el codificador 54 de la EL genera las correspondientes imágenes codificadas de granularidad fina para cada conjunto de imágenes residuales. Estas imágenes codificadas de granularidad fina tienen tasas de bits de R_{E1} , R_{E2} , R_{E3} ... R_{EN} , que se determinan sustancialmente de la misma manera que las de la corriente de bits de la EL de la primera realización. En un caso en el que la tasa máxima de bits de la EL para una corriente de bits particular de la BL es fijada como la tasa mínima de bits de la siguiente corriente de bits de la BL de radiodifusión simultánea, se aplican las ecuaciones siguientes:

$$R_{E1} = R_{B2} - R_{B1}$$

$$R_{E2} = R_{B3} - R_{B2}$$

$$R_{E(N-1)} = R_{BN} - R_{B(N-1)}$$

$$R_N = R_{MAX} - R_{BN}$$

La Figura 10 es un ejemplo de un gráfico de la calidad de la imagen en función de la tasa de bits que explica el caso correspondiente a las ecuaciones (2). Más específicamente, según se muestra en la Figura 10, la invención selecciona inicialmente una estructura de escalabilidad que tiene una capa base con una tasa de bits R_{B1} (que, en este caso, es R_{MIN}). La invención monitoriza entonces los parámetros de la red 43 de ancho de banda variable por medio del bloque 48 de cálculo, y determina periódicamente al efecto un nuevo ancho de banda R . A medida que el ancho de banda de la red 43 de ancho de banda variable aumenta con el tiempo, el controlador 55 selecciona imágenes residuales codificadas de granularidad fina progresivamente más detalladas para cada trama de la estructura seleccionada de escalabilidad/capa base, y envía esas imágenes al receptor. El receptor proporciona entonces esas imágenes a un dispositivo de visualización, tal como el dispositivo 14 de visualización de más arriba, llevando con ello al aumento progresivo en la calidad de la imagen mostrada por la línea 64 en la Figura 10. Sin embargo, usando la estructura de escalabilidad para R_{B1} , solo es posible proporcionar un aumento limitado en la calidad de la imagen, según se muestra con la línea discontinua 65 en la Figura 10.

En consecuencia, una vez que el ancho de banda R de la red 43 de ancho de banda variable alcanza un nivel predeterminado (que puede estar fijado de antemano en el controlador 55), se selecciona la estructura de escalabilidad para la tasa de bits R_{B2} . Como ocurrió más arriba, la invención sigue entonces monitorizando la red 43 de ancho de banda variable mediante el bloque 48 de cálculo, y recalculando el ancho de banda de la misma en el transcurso del tiempo. A medida que aumenta el ancho de banda de la red 43 de ancho de banda variable, el controlador 55 selecciona imágenes residuales codificadas de granularidad fina progresivamente más detalladas para cada trama de la estructura seleccionada de escalabilidad/capa base, y envía esas imágenes al receptor. El receptor proporciona entonces esas imágenes a un dispositivo de visualización, tal como el dispositivo 14 de visualización de más arriba, llevando con ello al aumento progresivo en la calidad de la imagen mostrada por la línea 66 en la Figura 10. Se lleva a cabo un procedimiento similar a este hasta R_{MAX} .

En virtud del procedimiento anterior, esta realización de la invención es capaz de usar corrientes de bits de radiodifusión simultánea para proporcionar un aumento general de la calidad de imagen sin grandes saltos en los puntos de transición R_{B2} y R_{B3} . Es decir, los sistemas convencionales que usan corrientes de bits de radiodifusión simultánea para aumentar la cantidad de la imagen tienen un gran salto en cada punto de transición entre dos corrientes de bits de radiodifusión simultánea. Esto da como resultado una transición abrupta en la imagen mostrada. En cambio, dado que la presente invención usa imágenes de granularidad fina entre los puntos de transición, la invención es capaz de proporcionar una transición gradual entre corrientes de bits, junto con un aumento continuo en la calidad de la imagen en el transcurso del tiempo.

Por supuesto, ocurre lo contrario de lo anterior para redes de ancho de banda variable que tienen un ancho de banda decreciente. Es decir, para un ancho de banda de receptor que disminuya de B_3 a B_2 , siendo $B_3 > B_2$, la invención selecciona una primera capa base y selecciona sucesivamente imágenes residuales codificadas de granularidad fina correspondientes a cada trama de la primera capa base que estén codificadas con tasas de bits sucesivamente menores. A medida que el ancho de banda disminuye de B_2 a B_1 , siendo $B_2 > B_1$, la invención selecciona una segunda capa base y selecciona sucesivamente imágenes residuales codificadas de granularidad fina correspondientes a cada trama de la segunda capa base que estén codificadas con tasas de bits sucesivamente menores. Esto da como resultado una disminución relativamente sin altibajos en la calidad de imagen para redes de ancho de banda variable que no tienen anchos de banda ni continuamente crecientes ni continuamente decrecientes, sino que, más bien, tienen anchos de banda fluctuantes u oscilantes. Tal es el caso de sistemas de ordenador o similares, que tienen prestaciones de procesamiento cambiantes.

En este punto se hace notar que aunque se han descrito las primeras dos realizaciones de la presente invención con respecto a una red de ancho de banda variable, estas realizaciones pueden ser usadas fuera de un contexto de red. Es decir, en vez de medir el ancho de banda de la red, la invención puede medir las prestaciones de procesamiento de un dispositivo receptor (por ejemplo, un PC) y variar entonces la codificación en consecuencia.

5 Tercera realización

La Figura 11 representa una tercera realización de la presente invención. En breve, esta realización es un procedimiento y un correspondiente aparato y etapas de procedimiento para codificar datos de vídeo y para multidifundir datos de vídeo codificados a varios receptores. En esta realización, el codificador 21 de vídeo escalable codifica una primera porción de los datos de vídeo (por ejemplo, la capa base) usando una técnica de codificación de predicción de tramas para producir una primera corriente de bits (por ejemplo, la corriente de bits de la BL), y luego codifica una segunda porción de los datos de vídeo (por ejemplo, la capa de mejora) usando una técnica de codificación de escalabilidad de granularidad fina para producir una segunda corriente de bits (por ejemplo, la corriente de bits de la EL). Posteriormente, se envía la primera corriente de bits a uno o más de los varios receptores, y se divide la segunda corriente de bits en dos o más subcorrientes. Después, estas dos o más subcorrientes son también enviadas a los varios receptores.

Según se muestra en la Figura 11, la tercera realización de la invención incluye la fuente 70 de vídeo, el codificador 71 de la BL, el bloque 72 de cálculo de imágenes residuales y el codificador 73 de la EL. Estas características son idénticas a las descritas más arriba con respecto a la primera realización. En consecuencia, en la presente memoria se omiten las descripciones detalladas de las mismas en aras de la brevedad. Según se muestra en la Figura 11, la tercera realización también incluye el controlador 74 de la tasa de multidifusión y el bloque 75 de cálculo. Siguen las descripciones detalladas de estas características de la invención.

El bloque 75 de cálculo es similar al bloque 48 de cálculo descrito más arriba, por cuanto determina R_{MIN} , R_{MAX} y R_{BL} . Sin embargo, en esta realización, R_{MIN} comprende el mínimo ancho de banda entre varios receptores (por ejemplo, varios PC) en la red 76 y R_{MAX} comprende el máximo ancho de banda entre los varios receptores de la red 76. Como anteriormente, el bloque 75 de cálculo fija R_{BL} en un valor entre R_{MIN} y R_{MAX} , y habitualmente igual a R_{MIN} para garantizar que incluso el receptor con el menor ancho de banda pueda procesar los datos de vídeo codificados producidos por la presente invención. Según se muestra en la Figura 11, en esta realización de la invención, el bloque 75 de cálculo también determina los anchos de banda $R_1, R_2 \dots R_N$ para las correspondientes categorías de receptores 1, 2...N (no mostrados) de la red 76. Esto puede hacerse monitorizando el tráfico de red destinado a estos receptores y procedente de los mismos y/o emitiendo consultas de estado a los respectivos receptores. Posteriormente, estos valores para $R_1, R_2 \dots R_N$ son proporcionados al controlador 74 de la tasa de multidifusión.

El controlador 74 de la tasa de multidifusión usa $R_1, R_2 \dots R_N$ para dividir la corriente de bits de la EL en subcorrientes que oscilan entre 0 bits y R_N bits. Es decir, según se muestra en la Figura 11, el controlador 74 de la tasa de multidifusión divide la corriente de bits de la EL en subcorrientes que tienen anchos de banda de:

35 $0 \rightarrow R_1$
 $R_1 \rightarrow R_2$
 $R_{N-1} \rightarrow R_N$,
siendo R_N menor o igual que $R_{MAX}-R_{BL}$. Cada una de estas subcorrientes corresponde a imágenes residuales codificadas de granularidad fina embebidas. Específicamente, la corriente de bits de 0 a R_1 comprende una imagen basta; la subcorriente de R_1 a R_2 comprende una mejora a la imagen basta; etcétera. Las subcorrientes descritas en la expresión (3) anterior son después enviadas a receptores de la red 76, junto con la corriente de bits de la BL. Estos receptores aceptarán entonces la corriente de bits de la BL y una, algunas, todas o ninguna de estas subcorrientes, dependiendo de las prestaciones de procesamiento del receptor y/o de la red. Entonces pueden usarse decodificadores, tales como el mostrado en la Figura 9, de estos receptores para decodificar las corrientes de bits.

Por supuesto, los expertos en la técnica se darán cuenta de que también es posible combinar las realizaciones segunda y tercera de la invención para producir un codificador que multidifunda subcorrientes para varias corrientes de bits de la BL de radiodifusión simultánea. Además, aunque se ha descrito esta realización con respecto a receptores en red, se hace notar que la realización también puede ser usada con receptores no conectados en red. La invención también puede ser usada para proporcionar datos codificados a varias redes conectadas de ancho de banda variable, por ejemplo a un solo PC o similar, mediante varias conexiones de red.

De forma similar, aunque las tres realizaciones de la invención descritas en la presente memoria se implementan preferentemente como código de ordenador, todos los componentes mostrados en las Figuras 6 y 11, o algunos de ellos, pueden implementarse usando elementos diferenciados de soporte físico y/o circuitos lógicos. Lo mismo se aplica para el decodificador mostrado en la Figura 9. Así, por ejemplo, los bloques 48 y 75 de cálculo pueden comprender una estación de trabajo, un PC u otro dispositivo controlado por un operario para introducir y seleccionar parámetros requeridos de control e instrucción. Por último, aunque las técnicas de codificación y decodificación de la presente invención han sido descritas en un entorno de PC, estas técnicas pueden ser usadas en cualquier tipo de

dispositivo de vídeo, incluyendo, sin limitación, televisores digitales/convertidores de señal de televisión por cable, equipos de videoconferencia y similares.

5 En este sentido, la presente invención ha sido descrita con respecto a realizaciones ilustrativas particulares. Ha de entenderse que la invención no está limitada a las realizaciones anteriormente descritas ni a modificaciones de la misma, y que las personas con un dominio normal de la técnica pueden realizar diversos cambios y modificaciones sin apartarse del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un procedimiento de codificación de datos de vídeo que comprenden una o más tramas, comprendiendo el procedimiento:
- 5 una primera etapa (44, S702) de codificación para producir una capa base que comprende datos (47) de vídeo codificados codificando una porción de los datos de vídeo usando una técnica de codificación de predicción de tramas;
- una etapa (45, S703) de generación para generar datos (51) de imágenes residuales correspondientes a las diferencias entre tramas en la capa base y las tramas de los datos de vídeo;
- 10 una segunda etapa (54, S704) de codificación para producir una capa de mejora que comprende datos codificados de imágenes residuales codificando los datos (51) de imágenes residuales; y
- una etapa de salida, en la que
- la segunda etapa de codificación usa una técnica de codificación de escalabilidad de granularidad fina, mediante la cual los datos codificados de imágenes residuales para una trama de los datos de vídeo comprenden varias imágenes residuales codificadas que incluyen una imagen basta inicial representada por un número predeterminado de bits y una imagen de mejora representada por el número predeterminado de bits para la imagen basta inicial y el siguiente número predeterminado de bits; y
- 15 la etapa (55, S706) de salida comprende dar salida a los datos (47) de vídeo codificados de la capa base y una o más de las imágenes residuales codificadas (50) de la capa de mejora a un receptor (43) para cada trama de los datos de vídeo originales,
- 20 **caracterizado porque:**
- la primera etapa (44, S702) de codificación comprende codificar la porción de los datos de vídeo con varias tasas de bits diferentes para producir múltiples versiones de los datos de vídeo codificados;
- la etapa (45, S703) de generación comprende generar datos (51) de imágenes residuales para cada versión de los datos de vídeo codificados;
- 25 la segunda etapa (54, S704) de codificación comprende codificar los datos (51) de imágenes residuales para cada versión de los datos de vídeo codificados usando una técnica de codificación de escalabilidad de granularidad fina; y
- la etapa (55, S706) de salida comprende dar salida a una versión seleccionada de los datos de vídeo codificados junto con al menos una imagen residual codificada correspondiente al efecto basada en un intervalo del ancho de banda disponible.
- 30
2. Un procedimiento según la reivindicación 1 en el que los datos codificados de imágenes residuales para una trama de los datos de vídeo comprenden, además, una imagen adicional de mejora representada por el número predeterminado de bits para la imagen basta inicial, el siguiente número predeterminado de bits para la imagen de mejora, y el siguiente número predeterminado de bits.
- 35
3. Un procedimiento según la reivindicación 1 que, además, comprende las etapas de:
- determinar (48) un ancho de banda del receptor (43); y
- seleccionar (55, S705) a cuál de las imágenes residuales codificadas (51) dar salida en la etapa (55, S706) de salida en función del ancho de banda del receptor (43).
- 40
4. Un procedimiento según la reivindicación 3 en el que la etapa (55, S705) de selección selecciona, para cada trama de los datos (47) de vídeo codificados, una imagen residual codificada que tiene la mayor tasa de bits que puede ser acomodada por el ancho de banda del receptor (43).
5. Un procedimiento según la reivindicación 4 en el que la etapa (55, S705) de selección se lleva a cabo en tiempo real por un controlador (55) de la tasa de vídeo escalable en tiempo real.
- 45
6. Un procedimiento según la reivindicación 1 en el que la primera etapa (44, S702) de codificación codifica la porción de los datos de vídeo usando uno de MPEG-1, MPEG-2 y MPEG-4.
7. Un procedimiento según la reivindicación 1 en el que la etapa (45, S703) de generación comprende las etapas de:
- 50 decodificar los datos (47) de vídeo codificados para producir datos de vídeo decodificados; y
- determinar los datos (51) de imágenes residuales determinando una diferencia entre píxeles en los datos de vídeo y píxeles en los datos de vídeo decodificados.
8. Un procedimiento según la reivindicación 7 en el que la etapa (45, S703) de generación comprende las etapas de:
- decodificar los datos (47) de vídeo codificados para producir datos de vídeo decodificados;
- filtrar los datos de vídeo decodificados para producir datos de vídeo filtrados; y

el primer medio codificador (44) es operable para codificar la porción de los datos de vídeo con varias tasas de bits diferentes para producir múltiples versiones de los datos de vídeo codificados; el medio generador (45) es operable para generar datos (51) de imágenes residuales para cada versión de los datos de vídeo codificados;

5 el segundo medio codificador (54) es operable para codificar los datos (51) de imágenes residuales para cada versión de los datos de vídeo codificados usando una técnica de codificación de escalabilidad de granularidad fina; y

10 el medio (55) de salida es operable para dar salida a una versión seleccionada de los datos de vídeo codificados junto con al menos una imagen residual codificada correspondiente al efecto basada en un intervalo del ancho de banda disponible.

16. Un sistema de red que comprende:

un aparato para codificar datos de vídeo según se reivindica en la reivindicación 15 y un aparato para decodificar los datos (47) de vídeo codificados objeto de salida de la capa base y las imágenes residuales codificadas (50) objeto de salida de la capa de mejora.

15

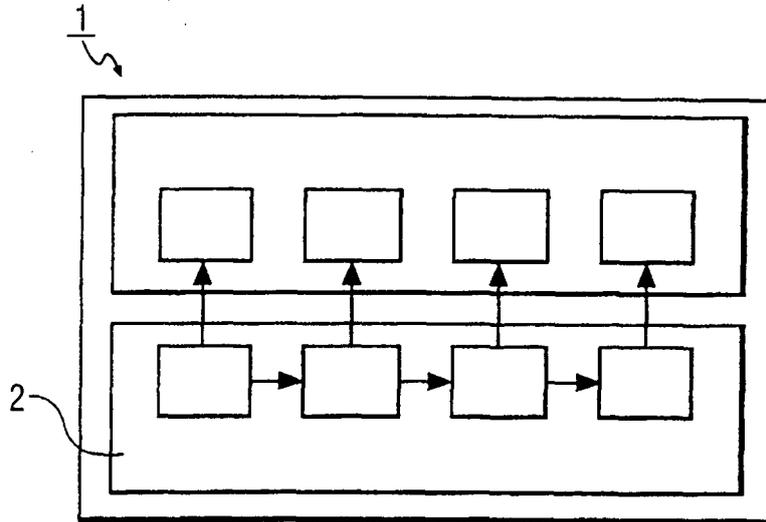


FIG. 1

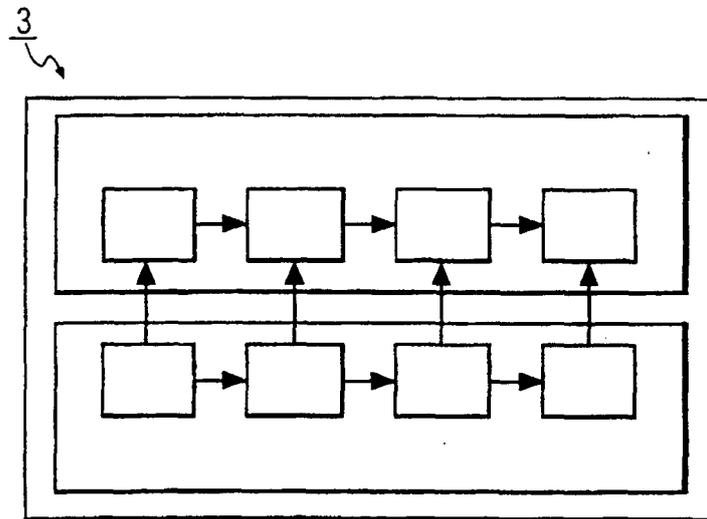


FIG. 2

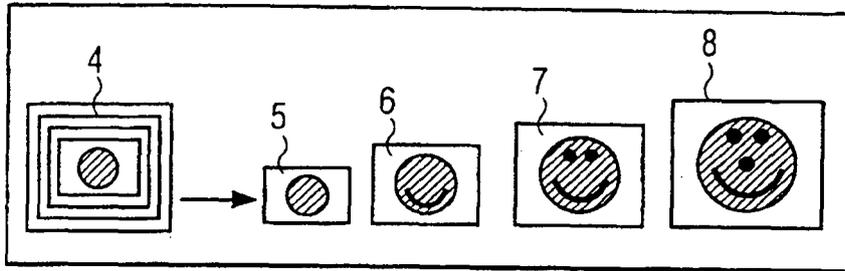


FIG. 3

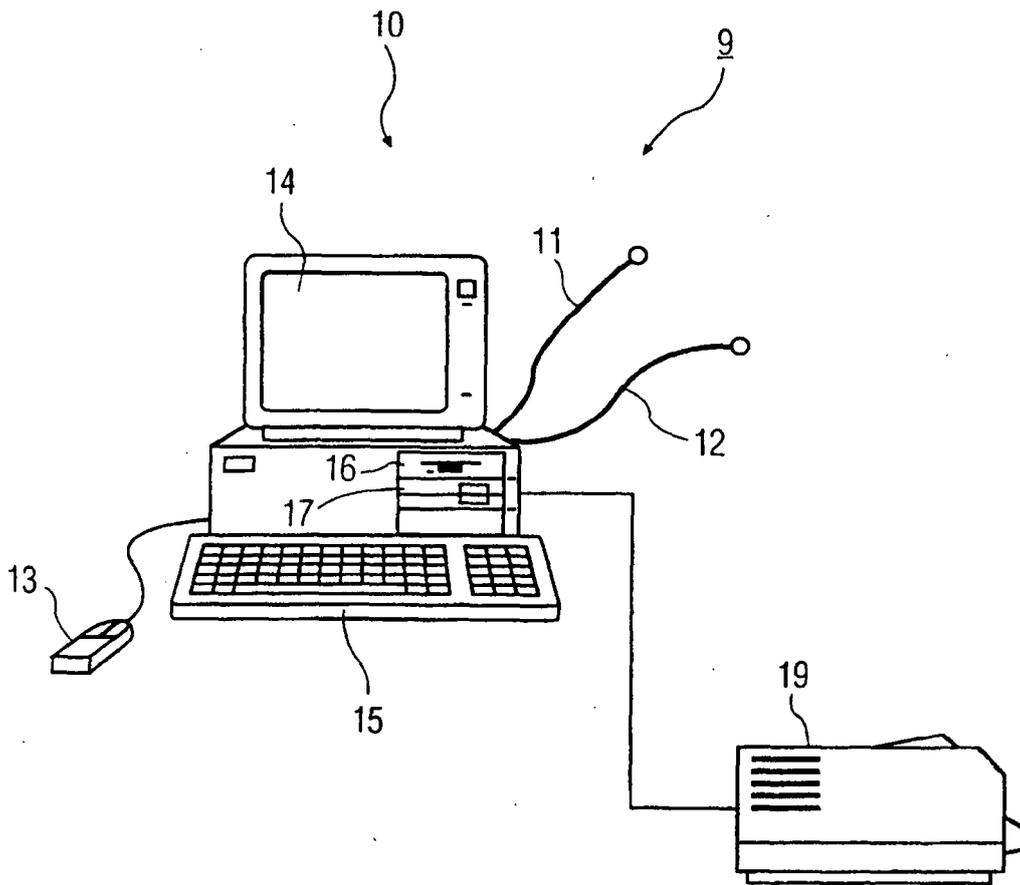


FIG. 4

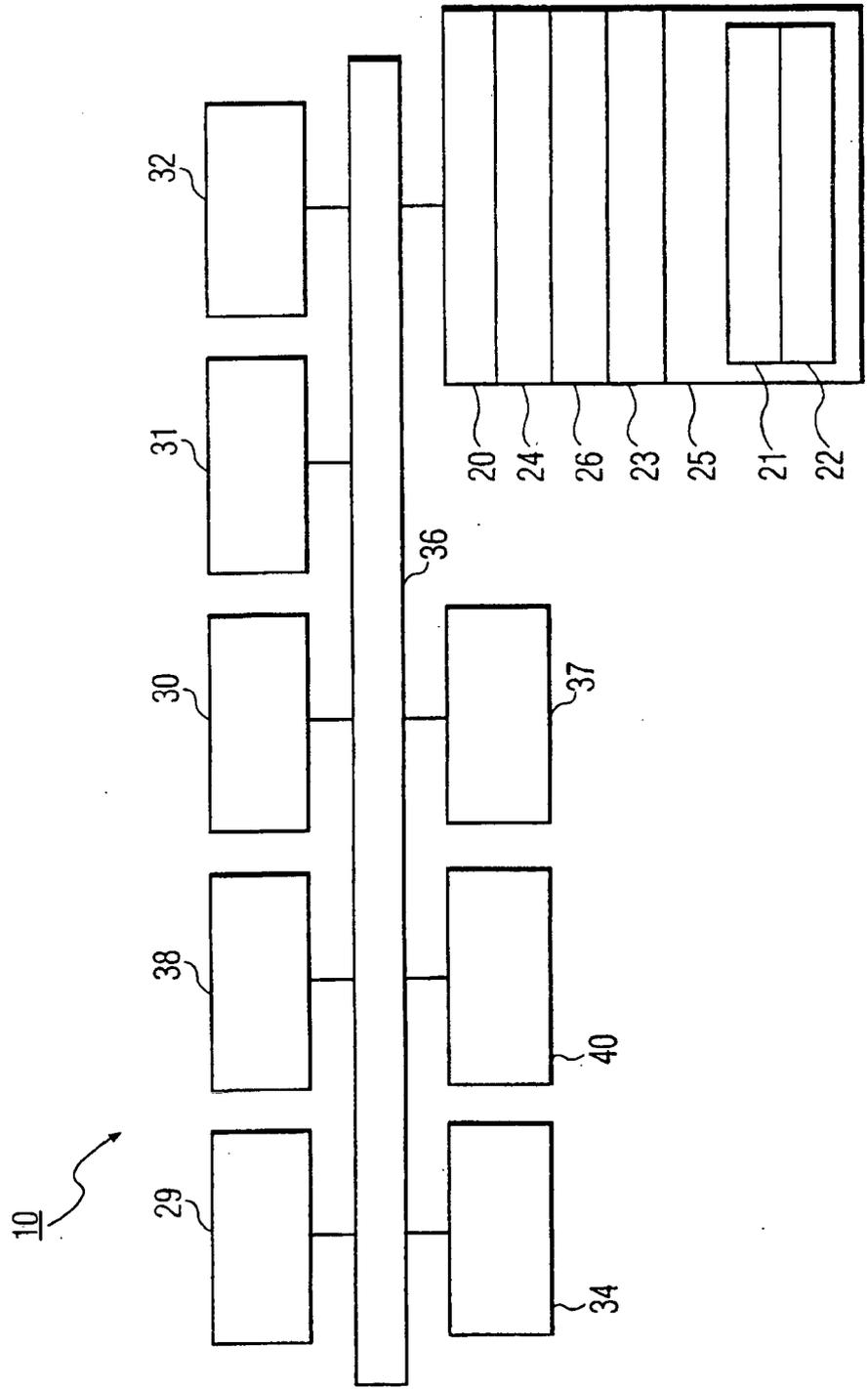


FIG. 5

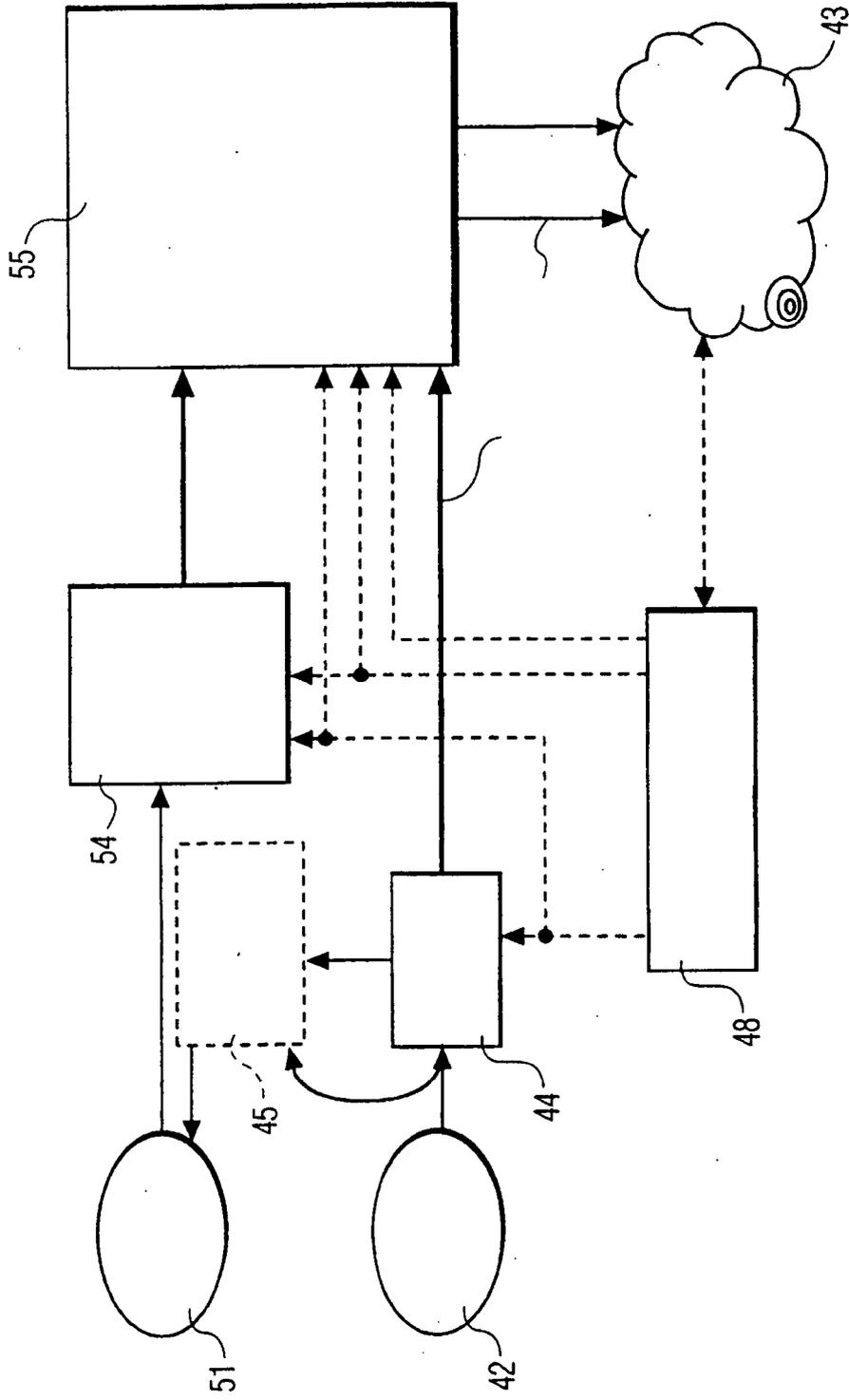


FIG. 6

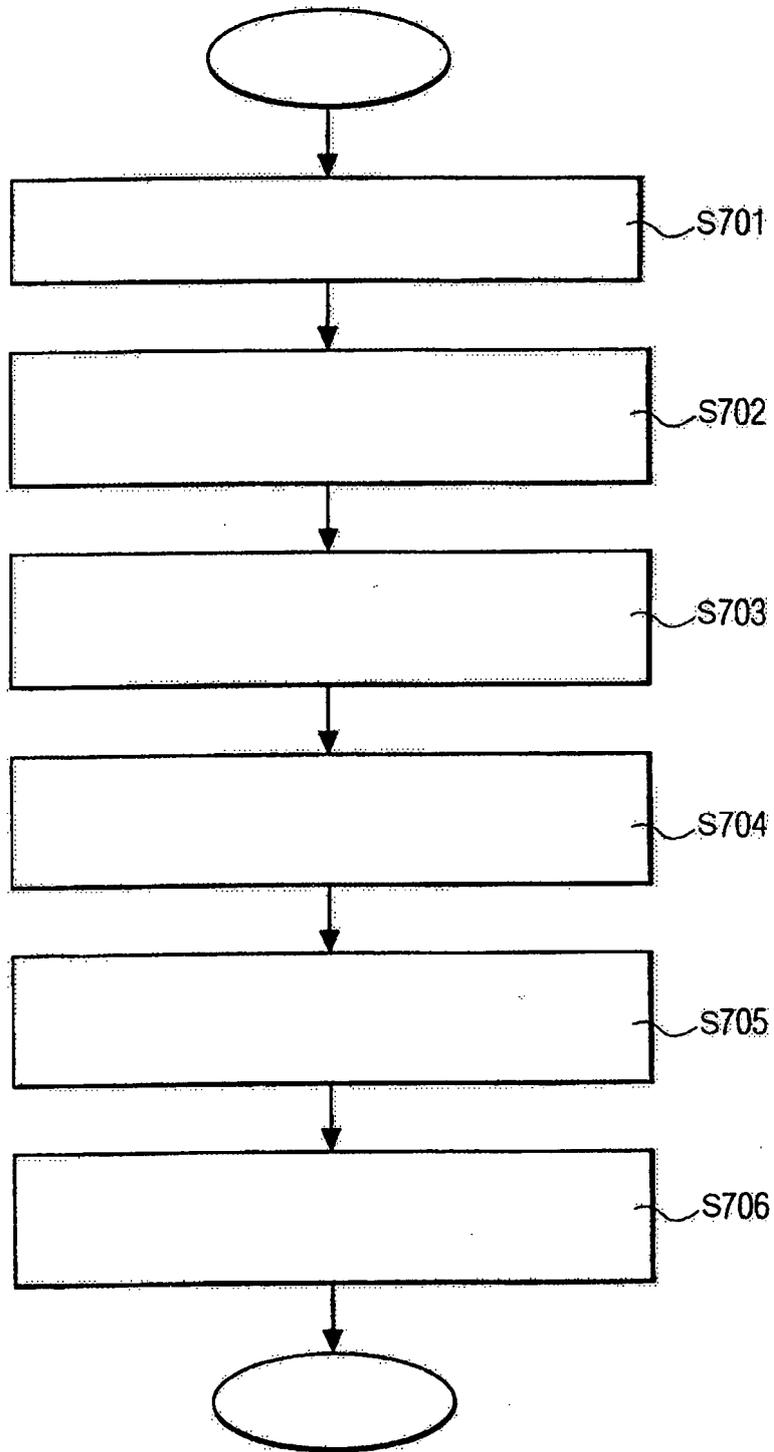


FIG. 7

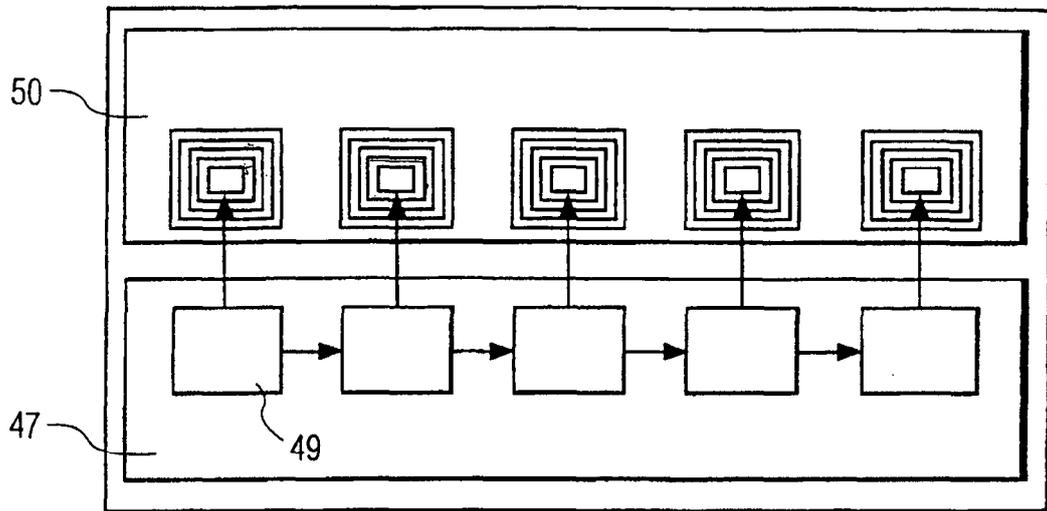


FIG. 8

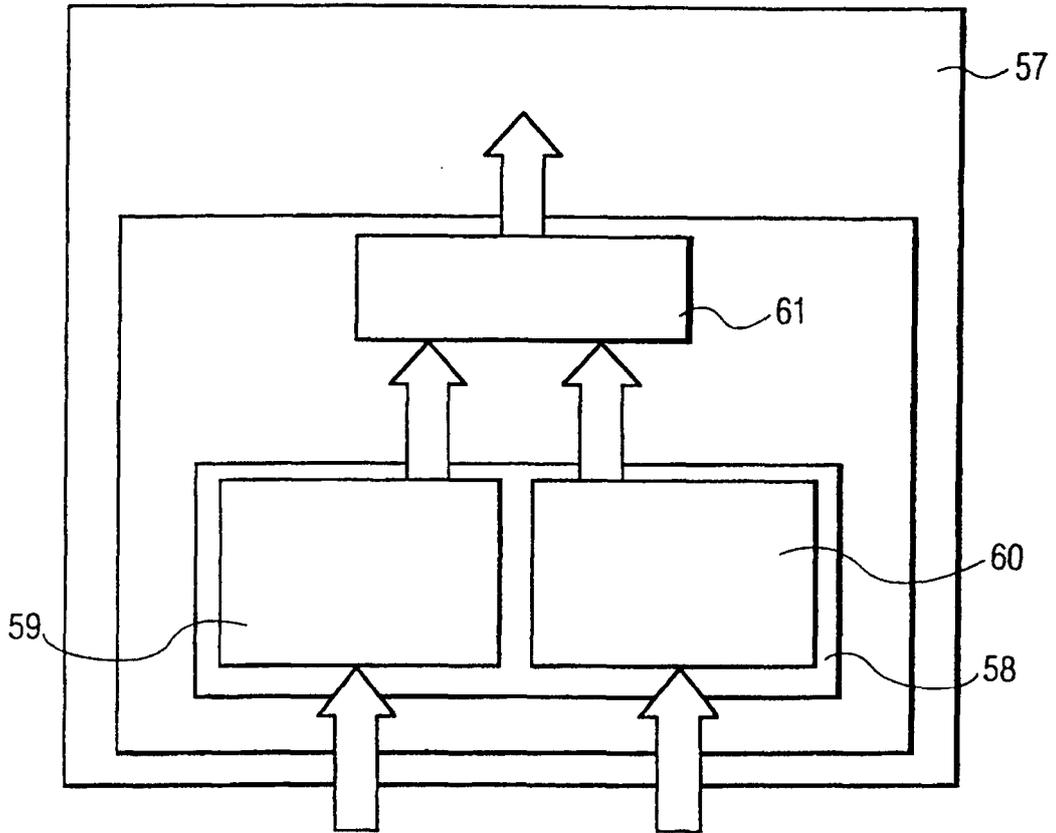


FIG. 9

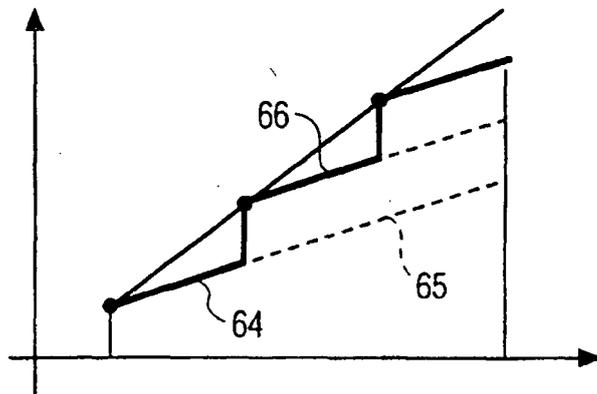


FIG. 10

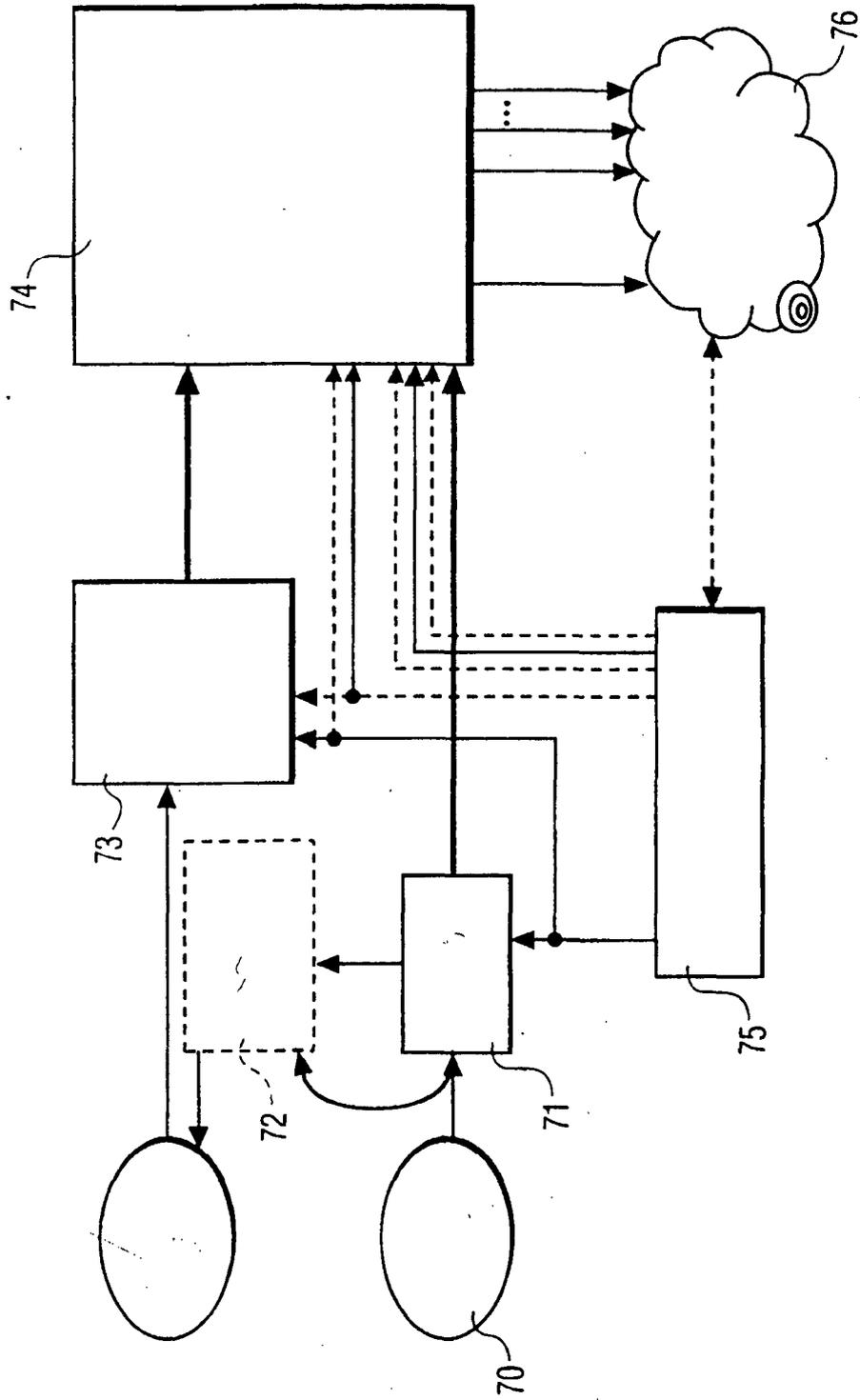


FIG. 11