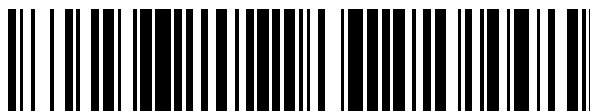


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 623 635**

51 Int. Cl.:

H04N 19/44 (2014.01)

H04N 19/46 (2014.01)

H04N 19/149 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.08.2002 E 02019056 (7)**

97 Fecha y número de publicación de la concesión europea: **15.02.2017 EP 1298938**

54 Título: **Decodificador de referencia generalizado de procesamiento de imagen o video**

30 Prioridad:

19.09.2001 US 955731

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

11.07.2017

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
ONE MICROSOFT WAY
REDMOND, WA 98052, US**

72 Inventor/es:

**RIBAS-CORBERA, JORDI y
CHOU, PHILIP A.**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 623 635 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 de referencia generalizado de procesamiento de imagen o video

La presente invención se refiere a la decodificación señales de imagen y video, así como otras señales que varían en el tiempo tales como voz y audio.

5 En las normas de codificación de video, un flujo de bits es compatible si puede decodificarse, al menos conceptualmente, mediante un modelo matemático de un decodificador que se conecta a la salida de un codificador. Un decodificador modelo de este tipo es conocido como el decodificador ficticio de referencia (HRD) en la norma de codificación H.263 y el verificador de memoria intermedia de video (VBV) en la norma de codificación MPEG. En general, un dispositivo (o terminal) de decodificador real comprende una memoria intermedia de decodificador, un
10 decodificador y una unidad de visualización. Si un dispositivo de decodificador real se construye de acuerdo con el modelo matemático del decodificador y se transmite un flujo de bits compatible al dispositivo en condiciones específicas, entonces la memoria intermedia de decodificador no se desbordará o infrutilizará y la decodificación se realizará correctamente.

15 Decodificadores de referencia (modelo) anteriores asumen que un flujo de bits se transmitirá a través de un canal a una tasa de bits constante dada y se decodificarán (tras un retardo de almacenamiento temporal dado) mediante un dispositivo que tiene algún tamaño de memoria intermedia dado. Por lo tanto, estos modelos son bastante inflexibles y no abordan los requisitos de muchas de las aplicaciones de video importantes actuales tales como difusión de video en directo o flujo continuo de video precodificado sobre demanda en trayectorias de red con diversas tasas de bit máximas, a dispositivos con diversos tamaños de memoria intermedia.

20 En decodificadores de referencia anteriores, el flujo de bits de video se recibe a una tasa de bits constante dada (normalmente la tasa media en bits por segundo del flujo) y se almacena en la memoria intermedia de decodificador hasta que la memoria intermedia alcanza algún nivel deseado de llenado. Por ejemplo, al menos los datos correspondientes a un fotograma inicial de información de video se necesitan antes de que la decodificación pueda reconstruir un fotograma de salida del mismo. Este nivel deseado se representa como el llenado de memoria intermedia de decodificador inicial y a una tasa de bits constante es directamente proporcional al retardo de
25 transmisión o de puesta en marcha (memoria intermedia). Una vez que este llenado se alcanza, el decodificador instantáneamente (en esencia) elimina los bits para el primer fotograma de video de la secuencia y decodifica los bits para mostrar el fotograma. Los bits para los siguientes fotogramas también se eliminan, codifican y muestran instantáneamente en intervalos de tiempo posteriores.

30 Un decodificador de referencia de este tipo opera a una tasa de bits fija, tamaño de memoria intermedia y retardo inicial. Sin embargo, en muchas aplicaciones de video contemporáneas (por ejemplo, flujo continuo de video a través de la Internet o redes ATM), el ancho de banda máximo varía de acuerdo con la trayectoria de red. Por ejemplo, el ancho de banda máximo difiere a base de si la conexión a la red es por módem, ISDN, DSL, cable y así sucesivamente. Además, el ancho de banda máximo también puede fluctuar en el tiempo de acuerdo con las
35 condiciones de red, por ejemplo, a base de congestión de red, el número de usuarios conectados y otros factores conocidos. Además todavía, el flujo de bits de videos se distribuyen a una diversidad de dispositivos con diferentes capacidades de memoria intermedia, incluyendo microteléfonos, Asistentes Digitales Personales (PDA), PC, dispositivos informáticos de bolsillo, caja de conexión de televisión, dispositivos de tipo DVD y similares y se crean para escenarios con diferentes requisitos de retardo, por ejemplo, flujo continuo de bajo retardo, descarga progresiva y similares.
40

Los decodificadores de referencia existentes no se ajustan para tales variables. Al mismo tiempo, los codificadores habitualmente no saben y no pueden saber de antemano cuáles serán las condiciones variables para un receptor dado. Como resultado, a menudo se desperdician recursos y tiempo de retardo innecesariamente o son inadecuados en muchos casos.

45 El documento EP-A-O 852 445 (NCR INT INC), 8 de julio 1998 (08-07-1998) desvela un procedimiento de optimización de ancho de banda para la transmisión de flujos de datos de video comprimidos. Un servidor en el sistema determina primero el tamaño de la memoria intermedia de receptor y a continuación el número mínimo de bloques consecutivos que cabrían en la memoria intermedia. El servidor adicionalmente ordena al receptor que espere antes de descomprimir y mostrar datos durante una cantidad de tiempo suficiente para recibir la cantidad de
50 datos de los primeros bloques de datos comprimidos que son iguales a la cantidad de datos en el bloque de datos comprimidos más largo en el archivo.

55 HSU C-Y ET AL: 'JOINT SELECTION OF SOURCE AND CHANNEL RATE FOR VBR VIDEO TRANSMISSION UNDER ATM POLICING CONSTRAINTS' IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE INC. NUEVA YORK, ESTADOS UNIDOS, vol. 15, n.º 6, 1 de agosto 1997 (01-08-1997), páginas 1016-1028 describe una transmisión de tasa de bits variable de video en redes ATM.

REIBMAN A R ET AL: 'CONSTRAINTS ON VARIABLE BIT-RATE VIDEO FOR ATM NETWORKS' IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE INC. NUEVA YORK, ESTADOS UNIDOS vol. 2, n.º 4, 1 de diciembre 1992 (01-12-1992), páginas 361-372 considera restricciones en la

tasa de bits codificados de una señal de video que se imponen por un canal y memorias intermedias de codificador y decodificador.

ITU: 'Annex B - Hypothetical Reference Decoder (HRD)' ITU-T RECOMMENDATION H.263, febrero 1998 (02-1998), páginas 49-50 desvela una infraestructura de servicios audiovisuales y codificación de video en movimiento.

- 5 ITU: "Annex C - Video Buffering Verifier (VBV)" ITU-T RECOMMENDATION H.262, febrero 2000 (02-2000), páginas 138-142 desvela la infraestructura de servicios audiovisuales y codificación de video en movimiento de acuerdo con H.262.

10 El documento el documento EP 0 515 101 A2 desvela un control de memoria intermedia para canal de tasa de bits variable. Se desvela un procesamiento de imagen de video que evita el desbordamiento e infrautilización de memoria intermedia de codificador y decodificador cuando la imagen se transmite en un canal de tasa de bits variable, que incluye la supervisión de una indicación de llenado de una memoria intermedia de codificador y una memoria intermedia de decodificador.

15 Por lo tanto el objeto de la presente invención es proporcionar un procedimiento mejorado y un correspondiente medio legible por ordenador y sistema informático para proporcionar datos para la distribución a un dispositivo decodificador para la decodificación de un flujo de bits de video.

Este objeto se resuelve mediante el objeto de las reivindicaciones independientes.

Realizaciones preferidas se definen mediante el objeto de las reivindicaciones dependientes.

20 Brevemente, la presente invención proporciona un decodificador de referencia generalizado mejorado que opera de acuerdo con cualquier número de conjuntos de tasas y parámetros de memoria intermedia para un flujo de bits dado. Cada conjunto caracteriza lo que se denomina como un modelo de depósito con fugas, o conjunto de parámetros, y contiene tres valores (R, B, F), en el que R es la tasa de bits de transmisión, B es el tamaño de memoria intermedia y F es el llenado de memoria intermedia de decodificador inicial. Como se aprecia, F/R es la puesta en marcha o retardo de memoria intermedia inicial.

25 Un codificador crea un flujo de bits de video que se contiene mediante algún número N deseado de depósitos con fugas o el codificador puede simplemente calcular los N conjuntos después de que el flujo de bits se ha generado. El codificador pasa el número al decodificador (al menos) una vez, con un correspondiente número de conjuntos (R, B, F) de alguna manera, tales como en un encabezamiento de flujo inicial o fuera de banda.

30 Cuando se recibe en el decodificador, si al menos dos conjuntos están presentes, el decodificador de referencia generalizado selecciona uno o interpola entre los parámetros de depósito con fugas y de este modo puede operar en cualquier tasa de bits máxima, tamaño de memoria intermedia o retardo deseados. Más particularmente, dada una tasa de transmisión máxima R' deseada, que se conoce en el extremo de decodificador, el decodificador de referencia generalizado selecciona el tamaño de memoria intermedia y retardo más pequeños (de acuerdo con los conjuntos (R, B, F) disponibles ya sea mediante la selección de uno, interpolación entre dos o más o mediante extrapolación) que será capaz de decodificar el flujo de bits sin sufrir infrautilización o desbordamiento de memoria intermedia. Como alternativa, para un tamaño de memoria intermedia B' dado de decodificador, el decodificador ficticio seleccionará y operará a la mínima tasa de transmisión máxima requerida.

35 Beneficios del decodificador de referencia generalizado incluyen que un proveedor de contenidos puede crear un flujo de bits una vez y un servidor puede distribuir el mismo a múltiples dispositivos de diferentes capacidades, usando una diversidad de canales de diferentes tasas de transmisión máximas. Como alternativa, un servidor y un terminal pueden negociar los mejores parámetros de depósito con fugas para las condiciones de red dadas, por ejemplo, el que produzca el menor retardo de puesta en marcha (memoria intermedia) o el que requiera la menor tasa de transmisión máxima para el tamaño de memoria intermedia dado del dispositivo. En la práctica, el tamaño de memoria intermedia y el retardo para algunos terminales pueden reducirse en un orden de magnitud o la tasa de transmisión máxima puede reducirse en un factor significativo (por ejemplo, cuatro veces) y/o la relación señal-ruido (SNR) puede aumentar quizás por varios dB sin aumentar la tasa de bits media, excepto para una cantidad insignificante de bits adicionales para comunicar la información de depósito con fugas.

Otros beneficios y ventajas serán evidentes a partir de la descripción detallada cuando se tome en conjunción con los dibujos, en los que:

Breve descripción de los dibujos

- 50 La Figura 1 en un diagrama de bloques que representa un sistema informático ilustrativo en el que puede incorporarse la presente invención;
 la Figura 2 en un diagrama de bloques que representa el codificador mejorado y decodificador de referencia generalizado y sus respectivas memorias intermedias de codificado y decodificado de datos de video o imagen de acuerdo con un aspecto de la presente invención;
 55 la Figura 3 es un gráfico de llenado de memoria intermedia con el paso del tiempo cuando se contiene en un

depósito con fugas de parámetros (R, B, F);

la Figura 4 es una representación de una tasa frente a curva de tamaño de memoria intermedia para un clip de video representativo; y

5 la Figura 5 es una representación de la tasa frente a curva de tamaño de memoria intermedia para un clip de video representativo con dos modelos de depósito con fugas (conjuntos de parámetros) proporcionados al decodificador de referencia generalizado para interpolación y extrapolación de acuerdo con un aspecto de la presente invención.

Descripción detallada

ENTORNO OPERATIVO ILUSTRATIVO

10 La Figura 1 ilustra un ejemplo de un entorno 120 operativo adecuado en el que puede implementarse la invención, particularmente para la decodificación de datos de imagen de video. El entorno 120 operativo es únicamente un ejemplo de un entorno operativo adecuado y no pretende sugerir ninguna limitación al ámbito de uso o funcionalidad de la invención. Otros sistemas informáticos bien conocidos y/o configuraciones que pueden ser adecuadas para su uso con la invención incluyen, pero sin limitación, ordenadores personales, ordenadores de servidor, dispositivos
15 portátiles o de mano, sistemas de multiprocesador, sistemas basado en microprocesador, electrónica de consumo programable, PC de red, miniordenadores, ordenadores centrales, entornos informáticos distribuidos que incluyen cualquiera de los sistemas o dispositivos anteriores y similares. Por ejemplo, es probable que datos de imagen decodificada y/o de imagen de video se realicen en un ordenador con más potencia de procesamiento que ordenadores personales de mano contemporáneos, pero no hay razón por la que la codificación no pueda realizarse
20 en el dispositivo ilustrativo o la decodificación en una máquina más potente.

La invención puede describirse en el contexto general de instrucciones ejecutables por ordenador, tales como módulos de programa, ejecutadas por uno o más ordenadores u otros dispositivos. En general, módulos de programa incluyen rutinas, programas, objetos, componentes, estructuras de datos y así sucesivamente que realizan tareas particulares o implementan tipos de datos abstractos particulares. Habitualmente la funcionalidad de los
25 módulos de programa pueden combinarse o distribuirse como se desee en diversas realizaciones. El dispositivo 120 informático habitualmente incluye al menos alguna forma de medio legible por ordenador. Medio legible por ordenador puede ser cualquier medio disponible que puede ser accedido mediante el dispositivo 120 informático. A modo de ejemplo, y no como limitación, medio legible por ordenador puede comprender medio de almacenamiento informático y medio de comunicación. Medio de almacenamiento informático incluye memoria volátil y no volátil y extraíble y no extraíble implementado en cualquier procedimiento o tecnología de almacenamiento de información tales como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Medio de almacenamiento informático incluye, pero sin limitación, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento óptico, casetes magnéticos, cinta magnética, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro
35 medio que puede usarse para almacenar la información deseada y que puede accederse mediante el dispositivo 120 informático. Medio de comunicación habitualmente incorpora instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada tales como una onda portadora u otro mecanismo de transporte e incluye cualquier medio de distribución de información. La expresión "señal de datos modulada", significa una señal que tiene una o más de sus características establecidas o cambiadas de tal manera que codifica información en la señal. A modo de ejemplo, y no como limitación, medio de comunicación incluye medios por cable tales como una red por cable o conexión directa por cable y medios inalámbricos tales como acústica, RF, infrarroja y otros medios inalámbricos. Combinaciones de cualquiera de los anteriores también debería incluirse dentro del ámbito de medio legible por ordenador.

La Figura 1 muestra componentes funcionales de un dispositivo 120 informático de mano de este tipo, que incluye un procesador 122, una memoria 124, una pantalla 126 y un teclado 128 (que puede ser un teclado físico o virtual). La memoria 124 generalmente incluye tanto memoria volátil (por ejemplo, RAM) como memoria no volátil (por ejemplo, ROM, tarjetas PCMCIA y así sucesivamente). Un sistema 130 operativo está residente en la memoria 124 y se ejecuta en el procesador 122, tal como el sistema operativo Windows® CE de Microsoft® Corporation u otro sistema operativo.

50 Uno o más programas 132 de aplicación se cargan en memoria 124 y funcionan en el sistema 130 operativo. Ejemplos de aplicaciones incluyen programas de correo electrónico, programas de planificación, programas PIM (gestión de información personal), programas de procesamiento de textos, programas de hojas de cálculo, programas de navegadores de Internet y así sucesivamente. El ordenador 120 personal de mano puede también incluir un gestor 134 de notificaciones cargado en la memoria 124, que se ejecuta en el procesador 122. El gestor
55 134 de notificaciones trata solicitudes de notificaciones, por ejemplo, de los programas 132 de aplicación.

El ordenador 120 personal de mano tiene una fuente 136 de alimentación, que se implementa como una o más baterías. La fuente 136 de alimentación puede incluir adicionalmente una fuente de potencia externa que anula o recarga las baterías incorporadas, tales como un adaptador de CA o una plataforma de acoplamiento alimentada.

El ordenador 120 personal de mano ilustrativo representado en la Figura 1 se muestra con tres tipos de mecanismos

externos de notificaciones: uno o más diodos 140 emisores de luz (LED) y un generador 144 de audio. Estos dispositivos pueden acoplarse directamente a la fuente 136 de alimentación de modo que cuando se activan, permanecen encendidos durante una duración dictada por un mecanismo de notificaciones incluso si el procesador 122 informático personal de mano y otros componentes pudieran apagarse para conservar energía de batería. El LED 140 preferentemente permanece encendido indefinidamente hasta que el usuario actúa. Obsérvese que versiones contemporáneas del generador 144 de audio usan demasiada energía para baterías actuales de ordenador personal de mano y por lo tanto se configura para apagarse cuando en resto del sistema se apaga o en alguna duración finita tras la activación.

DECODIFICADOR DE REFERENCIA GENERALIZADO

10 Un depósito con fugas es un modelo conceptual del estado (o llenado) de un codificador o memoria intermedia de decodificador como una función de tiempo. La Figura 2 muestra este concepto, en el que los datos 200 de entrada se suministran a un codificador 202 mejorado (descrito a continuación) que codifica los datos en una memoria 204 intermedia de codificador. Los datos codificados se transmiten a través de algún medio 206 de transmisión (conducto) a una memoria 208 intermedia de decodificador, que se decodifica a continuación mediante un decodificador 210 en datos 212 de salida, tales como una imagen o fotograma de video. Para fines de simplicidad, la memoria 208 intermedia de decodificador se describirá esencialmente en el presente documento, porque el llenado de las memorias intermedias de codificador y decodificador son conceptualmente complementarias entre sí, es decir, cuantos más datos en la memoria intermedia de decodificador, menos en la memoria intermedia de codificador y viceversa.

20 Un modelo de depósito con fugas se caracteriza por un conjunto de tres parámetros, R, B, y F, en el que R es la tasa de bits máxima (en bits por segundo) a la que los bits entran en la memoria 208 intermedia de decodificador. En escenarios de tasa de bits constante, R es a menudo la tasa de bits de canal y la tasa de bits media del clip de video o audio y conceptualmente puede considerarse como que corresponden al ancho del conducto 206. B es el tamaño del contador o memoria 208 intermedia de decodificador (en bits) que suaviza las fluctuaciones de tasa de bits de video. Este tamaño de memoria intermedia no puede ser mayor que la memoria intermedia física del dispositivo de decodificación. F es el llenado de memoria intermedia de decodificador inicial (también en bits) que necesita estar presente antes de que el decodificador comience a eliminar de la memoria intermedia. F es al menos tan grande como la cantidad de datos codificados que representa el fotograma inicial. Dejando de lado el tiempo de procesamiento, que para fines del presente ejemplo pueden considerarse instantáneos, F y R determinan el retardo D inicial o de puesta en marcha, en el que $D = F/R$ segundos.

Por lo tanto, en un modelo de depósito con fugas, los bits entran en la memoria 208 intermedia de decodificador a la tasa R hasta que el nivel de llenado es F (es decir, durante D segundos) y a continuación los bits necesarios para el primer fotograma, b_0 se eliminan (instantáneamente en el presente ejemplo). Los bits siguen entrando en la memoria intermedia a la tasa R y el decodificador elimina b_1, b_2, \dots, b_{n-1} bits para los siguientes fotogramas en algún instante de tiempo dado, habitualmente (pero no necesariamente) cada $1/M$ segundos, en el que M es la tasa de fotogramas del video.

La Figura 3 es un gráfico que ilustra el llenado de memoria intermedia de decodificador con el paso del tiempo para un flujo de bits que se contiene en un depósito con fugas de parámetros (R, B, F), como se ha descrito anteriormente, en el que el número de bits para el $i^{\text{ésimo}}$ fotograma es b_i . En la Figura 3, los fotogramas de video codificados se eliminan de la memoria intermedia (habitualmente de acuerdo con la tasa de fotogramas de video), como se muestra mediante los descensos en llenado de memoria intermedia. Más particularmente, siendo B_i el llenado de memoria intermedia de decodificador inmediatamente antes de la eliminación de b_i bits en el momento t_i . Un modelo de depósito con fugas genérico opera de acuerdo con las siguientes ecuaciones:

$$B_0 = F$$

$$B_{i+1} = \min (B, B_i - b_i + R (t_{i+1} - t_i)), \quad i = 0, 1, 2, \dots$$

45 Típicamente, $t_{i+1} - t_i = 1/M$ segundos, en el que M es la tasa de fotogramas (en fotogramas/seg) para el flujo de bits.

Un modelo de depósito con fugas con parámetros (R, B, F) contiene un flujo de bits si no existe infrutilización de la memoria 208 intermedia de decodificador (Figura 2). Debido a que el codificador y llenado de memoria intermedia de decodificador son complementarias entre sí, esto es equivalente a no desbordamiento de la memoria 204 intermedia de codificador. Sin embargo, se permite que la memoria 204 intermedia de codificador (el depósito con fugas) se vacíe o de forma equivalente la memoria 208 intermedia de decodificador puede llenarse, en cuyo punto no se transmiten bits adicionales desde la memoria 204 intermedia de codificador a la memoria 208 intermedia de decodificador. Por lo tanto, la memoria 208 intermedia de decodificador para de recibir bits cuando está llena, que es el motivo por el que se usa el operador min en la segunda ecuación anterior. Debido a que son complementarias, una memoria 208 intermedia de decodificador llena significa que la memoria 204 intermedia de codificador está vacía, como se describe a continuación con respecto a flujos de tasa de bits variable (VBR).

Obsérvese que un flujo de video dado puede contenerse en diversas configuraciones de depósito con fugas. Por ejemplo, si un flujo de video se contiene en un depósito con fugas con parámetros (R, B, F) , también se contendrá en un depósito con fugas con una mayor memoria intermedia (R, B', F) , en la que B' es mayor que B , o en un depósito con fugas con una mayor tasa de transmisión máxima (R', B, F) , en la que R' es mayor que R . Además, para cualquier tasa de bits R' , existe un tamaño de memoria intermedia que contendrá flujo de bits de video (de tiempo limitado). En el pero caso, a saber R' se aproxima a cero, el tamaño de memoria intermedia necesitará ser tan grande como el propio flujo de bits. En otras palabras, un flujo de bits de video puede transmitirse a cualquier tasa (independientemente de la tasa de bits media del clip) siempre que el tamaño de memoria intermedia sea lo suficientemente grande.

La Figura 4 es un gráfico de tamaño de memoria intermedia mínimo B_{\min} contra tasa de bits máxima R_{\min} para un flujo de bits dado, utilizando la segunda ecuación anterior, en la que el llenado de memoria intermedia inicial deseado se establece a una fracción constante del tamaño de memoria intermedia total. La curva en la Figura 4 indica que para transmitir el flujo a una tasa de bits máxima r , el decodificador necesita almacenar temporalmente al menos $B_{\min}(r)$ bits. Además, como se entiende a partir del gráfico, mayores tasas máximas requieren menores tamaños de memoria intermedia y por lo tanto retardos de puesta en marcha de memoria intermedia más cortos. Como alternativa, el gráfico indica que si el tamaño de la memoria intermedia de decodificador es b , la tasa máxima mínima requerida para la transmisión del flujo de bits es la asociada $R_{\min}(b)$. Además, la curva de los pares (R_{\min}, B_{\min}) para cualquier flujo de bits (tal como el de la Figura 4) es lineal por tramos y convexa.

De acuerdo con un aspecto de la presente invención, si al menos dos puntos de la curva se proporcionan mediante el codificador 202 mejorado, el decodificador 210 de referencia generalizado puede seleccionar un punto o linealmente interpolar entre puntos o extrapolar los puntos para llegar a algunos puntos $(R_{\text{interp}}, B_{\text{interp}})$ que son un poco más grandes aunque de forma segura que (R_{\min}, B_{\min}) . Como una consecuencia significativa, el tamaño de memoria intermedia puede reducirse de forma segura, en muchos casos aproximadamente en un orden de magnitud relativo a un único depósito con fugas que contiene el flujo de bits a su tasa media, con lo que el retardo asimismo se reduce. Como alternativa, para el mismo retardo, la tasa de transmisión máxima puede reducirse en un factor de (posiblemente) cuatro o mejorarse la relación señal-ruido (SNR) por (posiblemente) varios dB.

Para este fin, el codificador 202 se mejora disponiéndose para generar al menos dos conjuntos de parámetros 214 de depósito con fugas, por ejemplo, (R_1, B_1, F_1) , (R_2, B_2, F_2) ..., (R_N, B_N, F_N) , correspondientes a al menos dos puntos en la curva Tasa-Memoria intermedia que son útiles (por ejemplo, separados razonablemente con respecto al intervalo de R y/o B) para el clip de imagen o video dado. El codificador 202 mejorado a continuación proporciona estos conjuntos de parámetros de depósito con fugas, junto con el número N de los mismos, al decodificador 210 de referencia generalizado, tal como insertando los mismos en un encabezamiento de flujo inicial o como alternativa de alguna manera de fuera de banda. Obsérvese que incluso para un N relativamente grande, (por ejemplo, docenas de depósitos, mientras que de dos a cuatro normalmente serían suficientes para representar razonablemente una curva R - B), la cantidad de bytes extra necesarios para proporcionar esta información (por ejemplo, un byte para N , más ocho bytes por modelo de depósito con fugas o conjunto de parámetros) es insignificante cuando se compara con datos de video o imagen típicos.

Además, obsérvese que en tasas de bits mayores, el creador de contenido puede decidir especificar diferentes modelos de depósito con fugas en momentos diferentes en el flujo de bits, que sería útil siempre que una conexión falle durante transmisión y se reinicie en medio de un flujo de bits. Por ejemplo, pueden proporcionarse modelos de depósito con fugas durante intervalos de quince minutos, de tal forma que el decodificador puede cambiar sus condiciones operativas (por ejemplo, su tamaño de memoria intermedia o la tasa) como se desee mediante nueva selección, nueva interpolación o nueva extrapolación en momentos adecuados.

El valor deseado de N puede seleccionarse mediante el codificador, (obsérvese que si $N=1$, el decodificador 210 generalizado extrapolará puntos como un verificador de almacenamiento temporal de video MPEG). El codificador puede elegir preseleccionar los valores de depósito con fugas y codificar el flujo de bits con un control de tasa que asegure que las restricciones de depósito con fugas se cumplan, codificar el flujo de bits y a continuación usar la ecuación descrita anteriormente para calcular conjuntos de parámetros de depósito con fugas que contienen el flujo de bits a N diferentes valores de R o hacer ambas. El primer enfoque puede aplicarse a transmisión en directo y bajo demanda, mientras que las otras se aplica a bajo demanda.

De acuerdo con la presente invención, una vez recibido en el decodificador 210 de referencia generalizado, el decodificador 210 puede determinar qué depósito con fugas desea usar, conociendo la tasa de bits máxima disponibles para el mismo y/o su tamaño de memoria intermedia física. Como alternativa, el decodificador 210 de referencia generalizado puede interpolar linealmente entre o extrapolar linealmente desde estos puntos para encontrar un conjunto de parámetros adecuado para una configuración dada. La Figura 5 muestra dos parámetros de depósito con fugas conjuntos y sus valores linealmente interpolados (R, B) . Para referencia, la curva R - B calculada se representa como la línea discontinua fina, mientras que los valores R y B proporcionados en los modelos de depósito con fugas (R_x, B_x) y (R_y, B_y) se representan mediante asteriscos. La línea continua desde (R_x, B_x) a (R_y, B_y) representa los valores interpolados. Cualquier emparejamiento R o B elegido en esta línea continua mantendrá de forma adecuada (por ejemplo, sin desbordamiento o infrutilización) la memoria 208 intermedia de decodificador. Parámetros de depósito con fugas también pueden extrapolarse desde estos puntos, representados

por las líneas discontinuas gruesas en la Figura 5, y de nuevo, cualquier emparejamiento R o B elegido en esta línea continua mantendrá de forma adecuada la memoria 208 intermedia de decodificador.

El tamaño de memoria intermedia B interpolada entre puntos k y k+1 sigue la línea recta:

$$B = \frac{R - R_k}{R_{k+1} - R_k} B_k + \frac{R_{k+1} - R}{R_{k+1} - R_k} B_{k+1}$$

5 en la que $R_k < R < R_{k+1}$.

Análogamente, el llenado de memoria intermedia de decodificador inicial F puede interpolarse linealmente:

$$F = \frac{R - R_k}{R_{k+1} - R_k} F_k + \frac{R_{k+1} - R}{R_{k+1} - R_k} F_{k+1}$$

en la que $R_k < R < R_{k+1}$.

10 Se garantiza que el depósito con fugas resultante con parámetros (R, B, F) contiene el flujo de bits, porque, (como puede demostrarse matemáticamente) el tamaño de memoria intermedia mínimo B_{min} es convexo tanto en R como en F, es decir, el tamaño de memoria intermedia mínimo B_{min} correspondiente a cualquier combinación convexa (R, F) = a(R_k, F_k) + (1-a) (R_{k+1}, F_{k+1}), 0 < a < 1, es menor que o igual a B = aB_k + (1-a)B_{k+1}.

15 Como se ha descrito anteriormente, cuando R es mayor que R_N, el depósito con fugas (R, B_N, F_N) también contendrá el flujo de bits, con lo que B_N y F_N son el tamaño de memoria intermedia y llenado de memoria intermedia de decodificador inicial recomendados cuando R ≥ R_N. Si R es más pequeño que R₁, puede usarse el límite superior B = B₁ + (R₁-R) T, en el que T es la longitud de tiempo del flujo en segundos. Estos valores (R, B) fuera del intervalo de los N puntos pueden extrapolarse.

20 Debería observarse que el decodificador no necesita seleccionar, interpolar o extrapolar los parámetros de depósito con fugas, sino que otra entidad puede seleccionar los parámetros para el envío de un único juego al decodificador, que usará entonces ese único conjunto. Por ejemplo, dada alguna información tal como unos requisitos del decodificador, un servidor puede determinar (a través de selección, interpolación o extrapolación) un conjunto apropiado de parámetros de depósito con fugas para enviar a un decodificador y a continuación el decodificador puede decodificar usando únicamente un único conjunto de parámetros. Un intermediario para el servidor o decodificador también podría hacer la selección, interpolación o extrapolación de la información de depósito con fugas, sin que el decodificador nunca vea más de un depósito con fugas. En otras palabras, en lugar de que el decodificador decida, el servidor puede decidir, posiblemente con el servidor y decodificador de cliente negociando los parámetros. En general sin embargo, y de acuerdo con la presente invención, tiene lugar alguna determinación de un modelo de depósito con fugas apropiado, ya sea por adelantado o dinámicamente, a base de al menos dos modelos de depósito con fugas.

30 Los valores de la curva R-B para un flujo de bits dado puede calcularse a partir de los tiempos en los que suceden los valores de llenado máximos y mínimos en la trama de memoria intermedia de decodificador, tales como las ilustradas en la Figura 3. Más particularmente, considera dos momentos (t_M, t_m), de los valores máximo y mínimo de llenado de memoria intermedia de decodificador, respectivamente, para un flujo de bits contenido en un depósito con fugas de parámetros (R, B, F). Los valores máximo y mínimo de llenado pueden alcanzarse en varias ocasiones, pero considera el par (t_M, t_m) de valores más altos de tal forma que t_M < t_m. Asumiendo que el depósito con fugas se calcula adecuadamente, B es el tamaño de memoria intermedia mínimo que contiene el flujo de bits para los valores R, F, por lo tanto

$$B = \sum_{t=t_M}^{t=t_m} \left(b(t) - \frac{R}{M} \right) = \sum_{t=t_M}^{t=t_m} b(t) - n \frac{R}{M} = -n \frac{R}{M} + c$$

40 en la que b(t) es el número de bits para el fotograma en el momento t y M es la tasa de fotogramas en fotogramas/seg. En esta ecuación, n es el número de fotogramas entre los momentos t_M y t_m, y c es la suma de los bits para esos fotogramas.

- Esta ecuación puede interpretarse como un punto en una línea recta $B(r)$, en la que $r = R$ y $-n/M$ es la pendiente de la línea. Existe un intervalo de tasa de bits $r \in [R-r1, R+r2]$ de forma que el par más grande de valores t_M y t_m seguirá siendo el mismo, con lo que la anterior ecuación corresponde a una línea recta que define el tamaño de memoria intermedia mínimo B asociada a la tasa de bits r . Si la tasa de bits r está fuera del intervalo anterior, al menos uno de los valores t_M y/o t_m cambiará, con lo que si $r > R+r2$, la distancia de tiempo entre t_M y t_m será más pequeña y el valor de n en la nueva línea recta que define $B(r)$ también será más pequeño y la pendiente de la respectiva línea será mayor (menos negativa). Si $r < R-r1$, la distancia de tiempo entre t_M y t_m será más grande y el valor de n en la línea recta que define $B(r)$ también será más grande. La pendiente de la línea será entonces más pequeña (más negativa).
- Los valores de los pares (t_M, t_m) para un intervalo de tasas de bits (o los valores asociados de n) y algunos valores de c (al menos uno para un par dado) pueden almacenarse en el encabezamiento de un flujo de bits, y por lo tanto la curva $B(r)$ lineal por tramos podría obtenerse usando la anterior ecuación. Además, esta ecuación puede usarse para simplificar el cálculo de parámetros de modelo de depósito con fugas después de que un codificador ha generado un flujo de bits.
- En pruebas, se produjo el flujo de bits en la Figura 5, resultando una tasa de bits media de 797 Kbps. Como se muestra de forma general en la Figura 5, a una tasa de transmisión constante de 797 Kbps, el decodificador necesitaría un tamaño de memoria intermedia de aproximadamente 18.000 Kbits (R_x, B_x). Con un llenado de memoria intermedia de decodificador inicial igual a 18.000 Kbits, el retardo de puesta en marcha sería aproximadamente de 22,5 segundos. Por lo tanto, esta codificación (producido sin control de tasa) desplaza bits por hasta 22,5 segundos para lograr esencialmente la mejor calidad posible para su longitud codificada total.
- La Figura 5 también muestra que a una tasa de transmisión máxima de 2.500 Kbps (por ejemplo, la porción de tasa de bits de video de un CD 2x), el decodificador necesitaría un tamaño de memoria intermedia de únicamente 2.272 Kbits, (R_y, B_y), que es razonable para un dispositivo de hardware de consumo. Con un llenado de memoria intermedia inicial igual a 2.272 Kbits, el retardo de puesta en marcha será de únicamente aproximadamente 0,9 segundos.
- Por lo tanto para esta codificación, habitualmente podrían ser útiles dos modelos de depósito con fugas, por ejemplo, ($R=797$ Kbps, $B=18.000$ Kbits, $F=18.000$ Kbits) y ($R=2.500$ Kbps, $B=2.272$ Kbits, $F=2.272$ Kbits). Este primer conjunto de parámetros de depósito con fugas permitiría la transmisión del video en un canal de tasa de bits constante, con un retardo de aproximadamente 22,5 segundos. Mientras este retardo puede ser demasiado largo para muchos escenarios, es probablemente aceptable para flujo continuo por Internet de películas, por ejemplo. El segundo conjunto de parámetros de depósito con fugas permitiría la transmisión del video en una red compartida con tasa máxima de 2.500 Kbps o permitiría reproducción local desde un CD 2x, con un retardo de aproximadamente 0,9 segundos. Este sub-segundo retardo es aceptable para reproducción de acceso aleatorio con funcionalidad del tipo VCR (grabador de video).
- Los beneficios son evidentes cuando se considera lo que sucede cuando únicamente se especificó el primer depósito con fugas en el flujo de bits, pero no el segundo. En un evento de este tipo, incluso cuando se reproduce en un canal con tasa de bits máxima de 2.500 Kbps, el decodificador usaría una memoria intermedia de tamaño de 18.000 Kbits y por lo tanto el retardo sería $F/R = 18.000$ Kbits / 2.500 Kbps = 7,2 segundos. Como puede apreciarse, un retardo de este tipo es inaceptable para reproducción de acceso aleatorio, tal como con funcionalidad del tipo VCR. Sin embargo, si el segundo depósito con fugas se especifica adicionalmente, entonces a una tasa de 2.500 Kbps el tamaño de memoria intermedia desciende a 2.272 Kbits y el retardo desciende a 0,9 segundos, como se ha descrito anteriormente.
- Por otra parte, si únicamente el segundo depósito con fugas se especificó, pero no el primero, entonces a una tasa de transmisión constante de 797 Kbps, incluso un decodificador inteligente se vería forzado a usar una memoria intermedia que es mucho más grande que la necesaria, para garantizar que la memoria intermedia no se desbordará, a saber $B' = B + (R - R')T = 2.272$ Kbits + (2.500 Kbps - 797 Kbps) x 130 segundos = 223.662 Kbits. Incluso si tanta memoria está disponible en un dispositivo dado, esto corresponde a un retardo inicial de 281 segundos, o casi cinco minutos, que está lejos de ser aceptable. Sin embargo, si el primer depósito con fugas también se especifica, entonces a una tasa de 797 Kbps, el tamaño de memoria intermedia desciende a 18.000 Kbits y el retardo desciende a 22,5 segundos, como se ha descrito anteriormente.
- Además, cuando ambos depósitos con fugas se especifican, entonces el decodificador puede interpolar linealmente entre ellos (usando las formulas de interpolación anteriores), para cualquier tasa de bits R entre 797 Kbps y 2.500 Kbps, de este modo logrando casi tamaño de memoria intermedia y retardo mínimos a cualquier tasa dada. La extrapolación (representada en la Figura 5 por la línea discontinua gruesa) también es más eficiente tanto por debajo de 797 Kbps como por encima de 2.500 Kbps, comparado con la extrapolación con únicamente un único depósito con fugas en cualquier sitio entre 797 Kbps y 2.500 Kbps, inclusive.
- Como se ha demostrado en el ejemplo anterior, incluso solo dos conjuntos de parámetros de depósito con fugas pueden proporcionar un orden de magnitud de reducción en tamaño de memoria intermedia (por ejemplo, de 223.662 a 18.000 Kbits en un caso y de 18.000 a 2.272 Kbits en otro) y un orden de magnitud de reducción en

retardo (por ejemplo, de 281 a 22,5 segundos en un caso y e 7,2 a 0,9 segundos en otro) a una tasa de transmisión máxima dada.

5 Como alternativa, también posible reducir la tasa de transmisión máxima para un dado tamaño de memoria intermedia de decodificador. De hecho, como se desprende de la Figura 5, si la curva R-B puede obtenerse mediante la interpolación y/o extrapolación de múltiples depósitos con fugas, entonces es posible para un decodificador con un tamaño de memoria intermedia física fijo elegir la tasa de transmisión máxima mínima necesaria para decodificar de forma segura el flujo de bits sin infrutilización de memoria intermedia de decodificador. Por ejemplo, si el decodificador tiene una memoria intermedia fija de tamaño de 18.000 Kbits, entonces la tasa de transmisión máxima para la codificación puede ser tan baja como 797 Kbps. Sin embargo, si únicamente se especifica el segundo depósito con fugas, pero no el primero, entonces el decodificador puede reducir la tasa de bits a no menos de $R' = R - (B' - B)/T = 2.500 \text{ Kbps} - (18.000 \text{ Kbits} - 2.272 \text{ Kbits}) / 130 \text{ segundos} = 2.379 \text{ Kbps}$. En este caso, comparado con usar un único depósito con fugas, usar solo dos depósitos con fugas reduce la tasa de transmisión máxima en un factor de cuatro, para el mismo tamaño de memoria intermedia de decodificador.

15 Tener múltiples parámetros de depósito con fugas también puede mejorar la calidad del video reconstruido, a la misma tasa de codificación media. Considera la situación en la que ambos depósitos con fugas están disponibles para la codificación. Como se ha descrito anteriormente, con esta información en el decodificador, es posible reproducir la codificación con un retardo de 22,5 segundos si la tasa de transmisión máxima es 797 Kbps y con un retardo de 0,9 segundos si la tasa de transmisión máxima es 2.500 Kbps.

20 Sin embargo, si el segundo depósito con fugas no está disponible, entonces el retardo aumenta desde 0,9 a 7,2 segundos a 2.500 Kbps. Una manera de reducir el retardo de vuelta a 0,9 segundos sin el beneficio del segundo depósito con fugas es recodificar el clip con un control de tasa, reduciendo el tamaño de memoria intermedia (del primer depósito con fugas) desde 18.000 Kbits a $(0,9 \text{ segundos}) \times (2.500 \text{ Kbps}) = 2.250 \text{ Kbits}$. Esto garantizaría que el retardo fuera de únicamente 0,9 segundos si la tasa de transmisión máxima fuera de 2.500 Kbps, aunque el retardo a 797 Kbps también disminuiría, desde 22,5 a 2,8 segundos. Sin embargo, como consecuencia, la calidad (SNR) también disminuirá en una cantidad estimada de varios dB, especialmente para un clip con un intervalo dinámico grande.

30 Por lo tanto, especificar un segundo depósito con fugas puede aumentar la SNR en posiblemente varios dB, sin cambio en la tasa de bits media, excepto para la insignificante cantidad de bits adicionales por clip para especificar el segundo depósito con fugas. Este aumento en SNR será visible en reproducción para cada tasa de transmisión máxima.

35 Los beneficios de especificar múltiples depósitos con fugas en el decodificador de referencia generalizado se realizan en el que una única codificación se transmite en canales con diferentes tasas máximas o a dispositivos con diferentes tamaños de memoria intermedia física. Sin embargo, en la práctica esto se está volviendo más y más común. Por ejemplo, contenido que se codifica fuera de línea y almacenado en un disco es a menudo reproducido localmente, así como transmitidos en flujo continuo en redes con diferentes tasas máximas. Incluso para reproducción local, diferentes velocidades de discos (por ejemplo, CD x1 a través de DVD 8x) afectan la tasa de transferencia máxima. Además, las tasas de transmisión máximas a través de conexiones de red también varía dramáticamente de acuerdo con la velocidad del enlace limitante, que está habitualmente cerca del usuario final (por ejemplo, 100 o 10 base Ethernet, T1, DSL, ISDN, módems y así sucesivamente). Las capacidades de memoria intermedia de dispositivos reproductores también varían significativamente, desde ordenadores de sobremesa con gigabytes de espacio de memoria intermedia a pequeños dispositivos electrónicos de consumo con espacio de memoria intermedia que es más pequeña en varios órdenes de magnitud. Los múltiples depósitos con fugas y el decodificador de referencia generalizado propuesto de la presente invención hacen posible que el mismo flujo de bits se transmita en una diversidad de canales con el mínimo retardo de puesta en marcha, requisitos mínimos de memoria intermedia de decodificador y máxima calidad posible. Esto se aplica no únicamente a video que se codifica fuera de línea, sino también a video en directo que se difunde simultáneamente a través de diferentes canales a diferentes dispositivos. En breve, el decodificador de referencia generalizado propuesto añade flexibilidad significativa a flujos de bits existentes.

50 Como puede verse en la anterior descripción detallada, se proporciona un decodificador de referencia generalizado mejorado relativo a aquellos en normas anteriores. El decodificador de referencia generalizado requiere únicamente una pequeña cantidad de información del codificador (por ejemplo, en el encabezamiento del flujo de bits) para proporcionar mucha mayor flexibilidad para la distribución de flujo de bits a través de redes contemporáneas en las que el ancho de banda es ancho de banda variable y/o terminales tienen una diversidad de tasas de bits y capacidades de almacenamiento temporal. El decodificador de referencia de la presente invención permite estos nuevos escenarios, mientras que reduce el retardo de transmisión a un mínimo para el ancho de banda disponible y además, prácticamente minimiza el requisito de tasa de bits de canal para la distribución a dispositivos con limitaciones dadas de tamaño de memoria intermedia física.

REIVINDICACIONES

1. Un procedimiento implementado en ordenador que comprende:

con un codificador de vídeo, codificar datos para video dado para producir un flujo de bits de datos codificados para el video dado;

5 generar múltiples conjuntos (214) de parámetros de decodificador de referencia para el flujo de bits, en el que cada uno de los múltiples conjuntos (214) comprende un parámetro de tasa y un parámetro de tamaño de memoria intermedia de decodificador, en el que cada uno de los múltiples conjuntos (214) caracteriza un modelo de depósito con fugas para el flujo de bits, estando el modelo de depósito con fugas para cada uno de los múltiples conjuntos (214), estando adicionalmente **caracterizado por** un llenado de memoria intermedia de decodificador inicial o retardo de memoria intermedia inicial, y

10 en el que, para cada uno de los múltiples conjuntos (214) de parámetros de decodificador de referencia, el conjunto de parámetros de decodificador de referencia contiene el flujo de bits, de tal forma que no existe infrutilización de una memoria (208) intermedia de decodificador de un decodificador (210) de referencia que opera de acuerdo con el conjunto de parámetros de decodificador de referencia para la decodificación de los datos codificados para el video dado;

15 proporcionar el flujo de bits para distribución a uno o más dispositivos de decodificador para la decodificación de los datos codificados para el video dado;

proporcionar un valor numérico N junto con el flujo de bits, en el que el valor numérico N indica cuántos conjuntos de parámetros de decodificador de referencia hay en los múltiples conjuntos (214) de parámetros de decodificador de referencia; y

20 proporcionar los múltiples conjuntos (214) de parámetros de decodificador de referencia junto con el flujo de bits para hacer cada uno de los múltiples conjuntos de parámetros de decodificador de referencia disponibles para su uso en determinación de un tamaño de memoria intermedia mínimo o una tasa máxima mínima para la decodificación de los datos codificados para el video dado, incluyendo la determinación:

- 25
- 1) determinar el tamaño de memoria intermedia a partir de una tasa dada y uno o más de los múltiples conjuntos (214) proporcionados de parámetros de decodificador de referencia, o
 - 2) determinar la tasa máxima a partir de un tamaño de memoria intermedia dado y uno o más de los múltiples conjuntos (214) proporcionados de parámetros de decodificador de referencia.

30 2. El procedimiento de la reivindicación 1 en el que se proporcionan los múltiples conjuntos (214) de parámetros de decodificador de referencia en un encabezamiento de flujo junto con el valor numérico N.

3. El procedimiento de una cualquiera de las reivindicaciones 1 a 2 en el que la determinación incluye seleccionar un parámetro de uno de los múltiples conjuntos (214) de parámetros de decodificador de referencia.

35 4. El procedimiento de una cualquiera de las reivindicaciones 1 a 2 en el que la determinación incluye interpolar entre parámetros de dos de los múltiples conjuntos (214) de parámetros de decodificador de referencia.

5. El procedimiento de una cualquiera de las reivindicaciones 1 a 2 en el que la determinación incluye extrapolar desde un parámetro de uno de los múltiples conjuntos (214) de parámetros de decodificador de referencia.

40 6. Un medio legible por ordenador que almacena instrucciones ejecutables por ordenador para provocar que un sistema informático programado de este modo realice el procedimiento de una cualquiera de las reivindicaciones 1 a 5.

7. Un sistema informático adaptado para realizar el procedimiento de una cualquiera de las reivindicaciones 1 a 5.

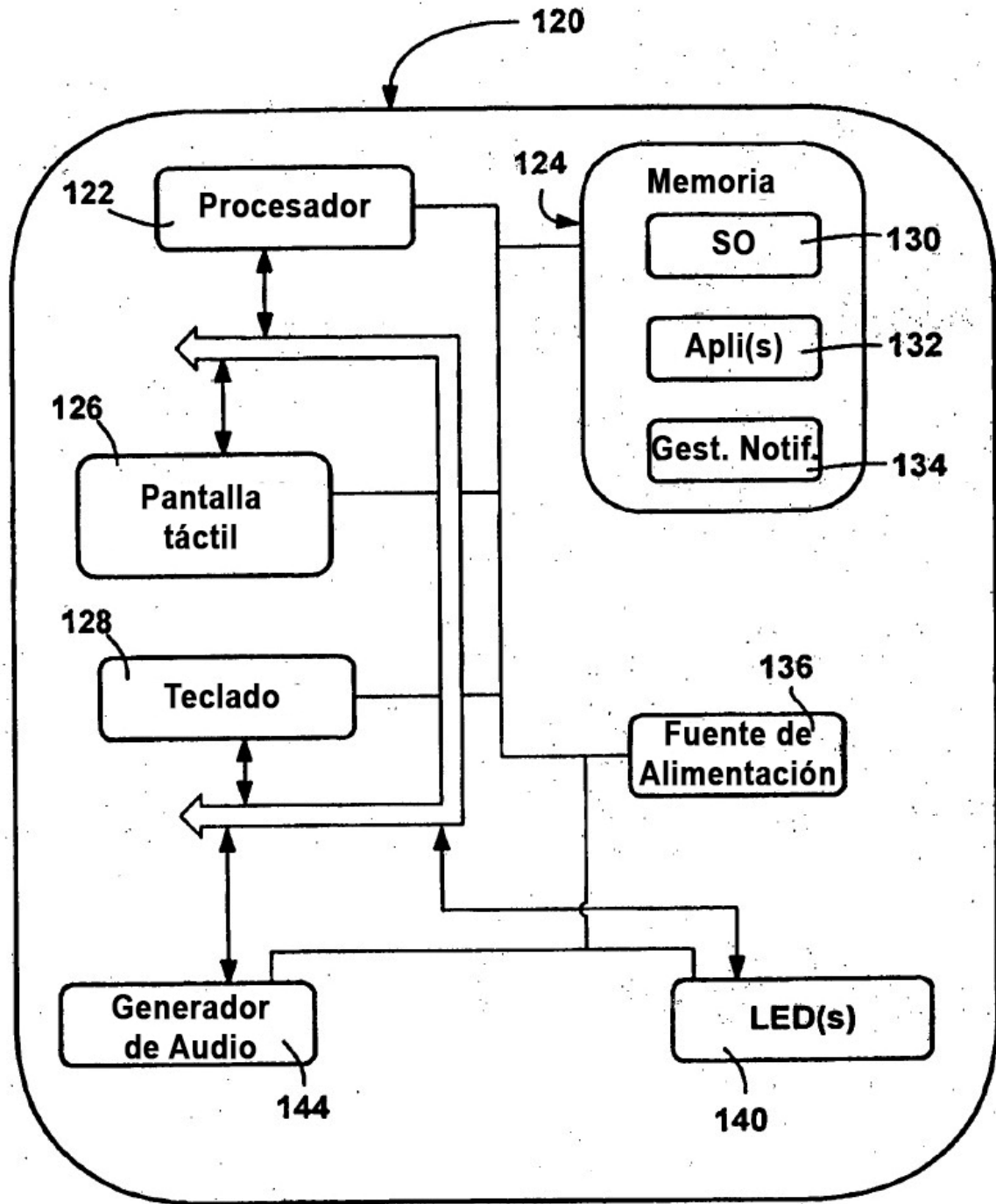


FIG. 1

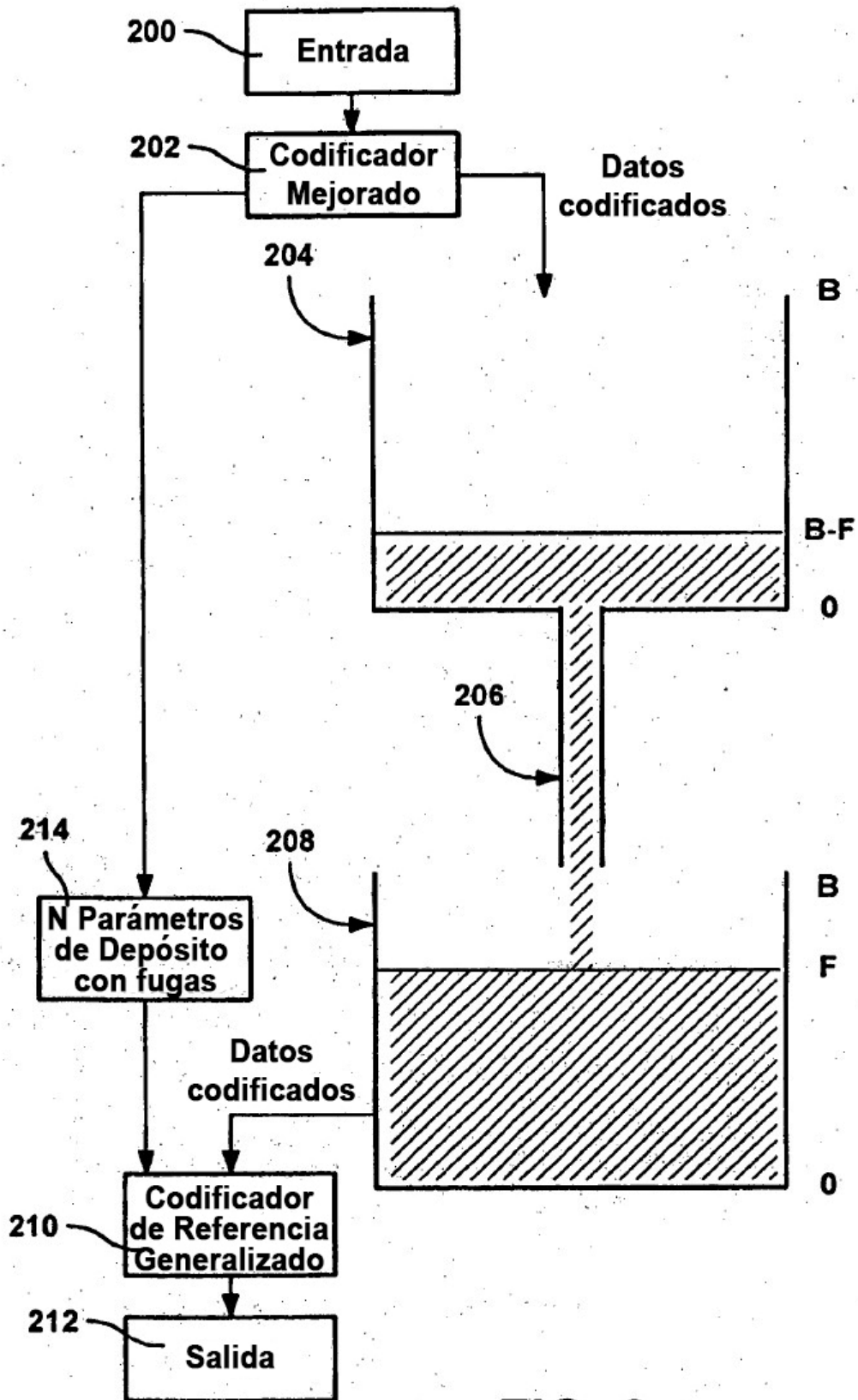


FIG. 2

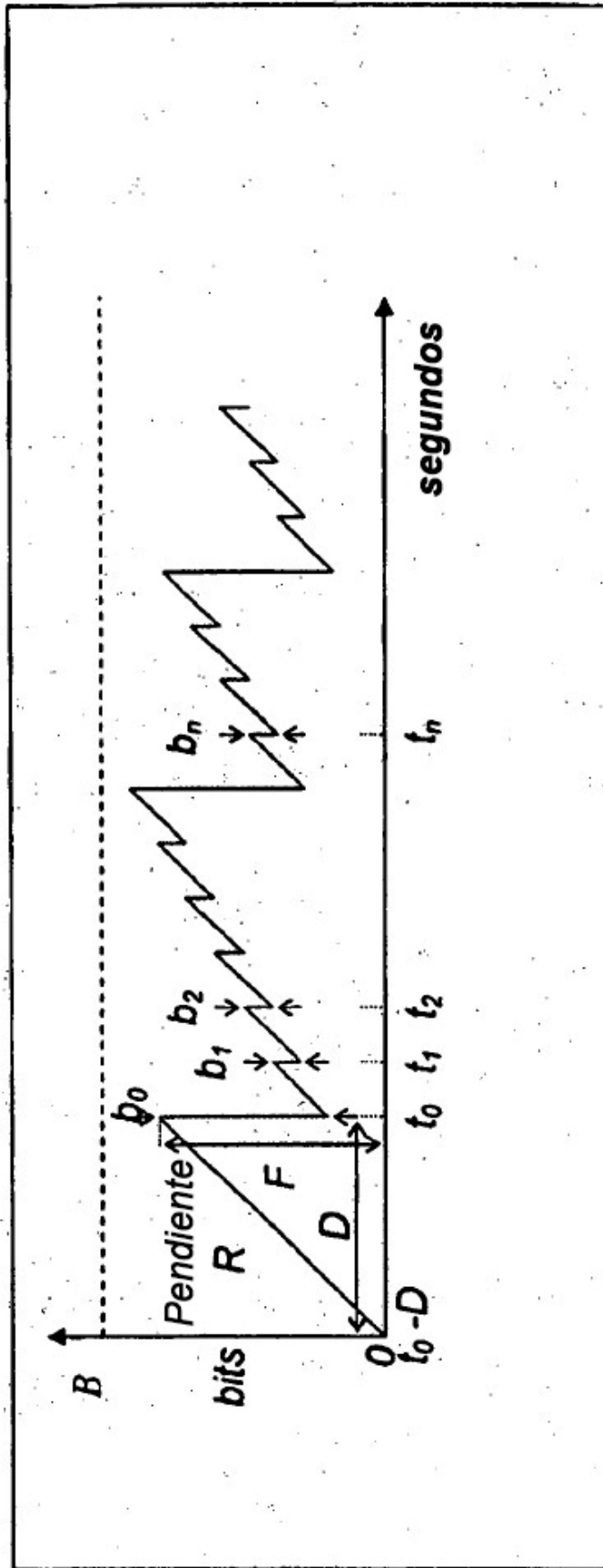


FIG. 3

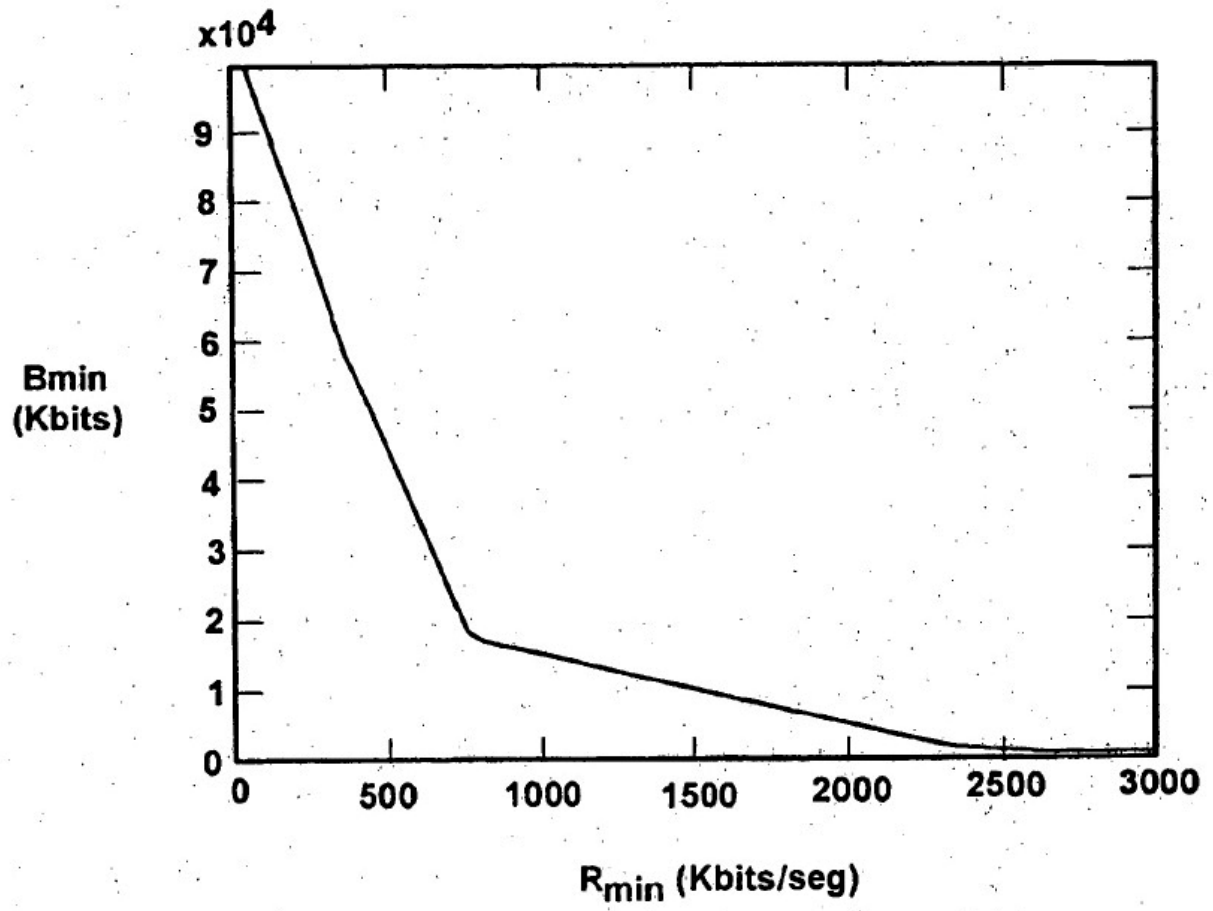


FIG. 4

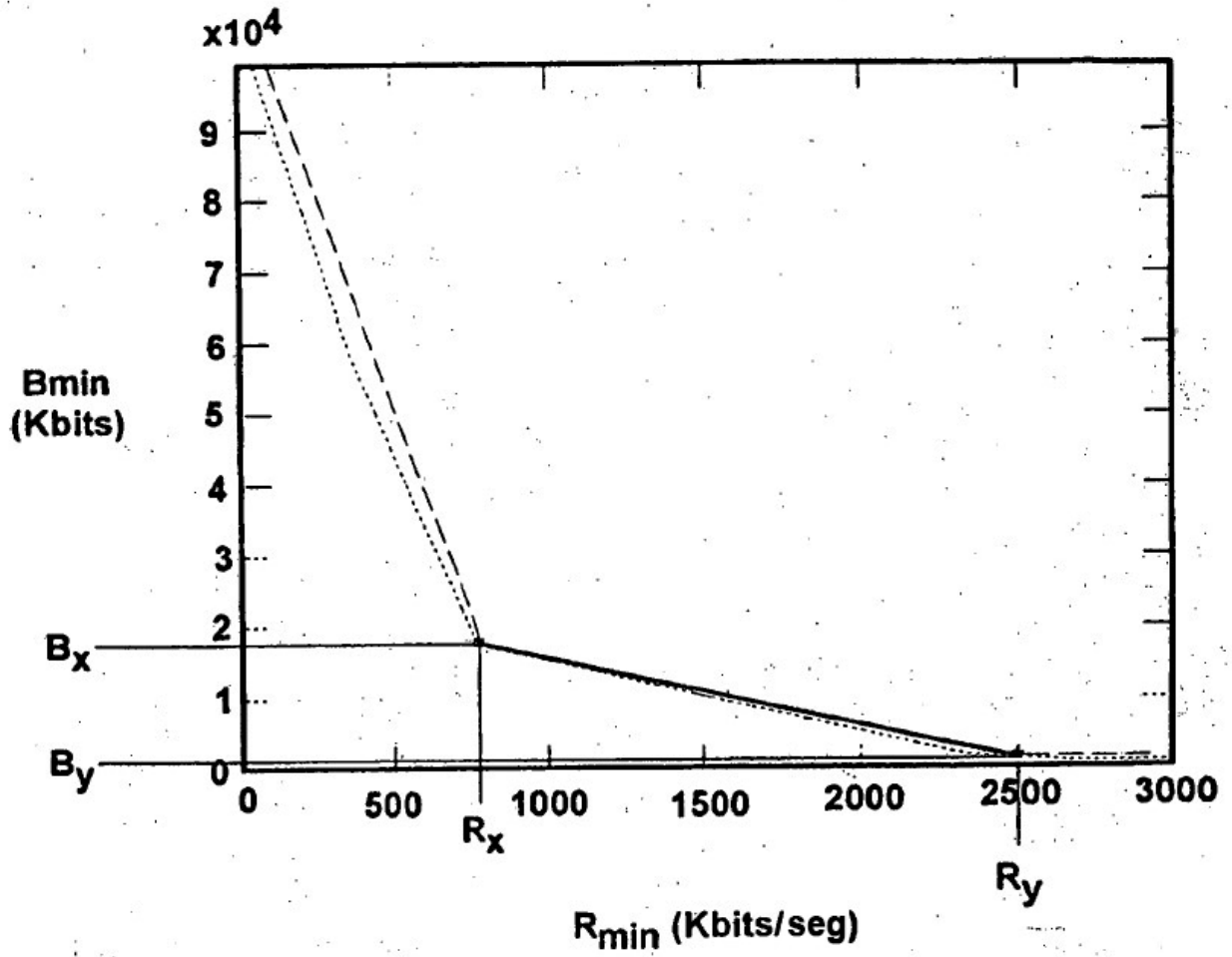


FIG. 5