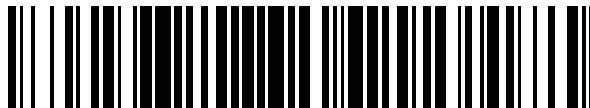


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 390 596**

51 Int. Cl.:

H04N 7/50 (2006.01)

H04N 7/26 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **10162205 .8**

96 Fecha de presentación: **26.03.2004**

97 Número de publicación de la solicitud: **2209319**

97 Fecha de publicación de la solicitud: **21.07.2010**

54 Título: **Codificador de vídeo**

30 Prioridad:
31.03.2003 US 404947

45 Fecha de publicación de la mención BOPI:
14.11.2012

45 Fecha de la publicación del folleto de la patente:
14.11.2012

73 Titular/es:
**SHARP KABUSHIKI KAISHA (100.0%)
22-22, NAGAIKE-CHO ABENO-KU
OSAKA-SHI, OSAKA 545-8522, JP**

72 Inventor/es:
DESHPANDE, SACHIN GOVIND

74 Agente/Representante:
CARVAJAL Y URQUIJO, Isabel

ES 2 390 596 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

Codificador de vídeo

5 Antecedentes de la invención

La presente invención se refiere a un descodificador hipotético de referencia.

10 Un sistema de vídeo digital incluye un transmisor y un receptor que ensambla vídeo que consta de audio, imágenes y componentes auxiliares para una presentación coordinada al usuario. El sistema transmisor incluye subsistemas para recibir y comprimir los datos fuente digitales (los flujos de datos elementales o de aplicación, que representan el audio, vídeo y componentes de datos auxiliares de un programa); multiplexar los datos procedentes de diversos flujos de datos elementales en un solo flujo de bits de transporte; y transmitir los datos al receptor. En el receptor, el flujo de bits de transporte es desmultiplexado en sus flujos de datos elementales constitutivos. Los flujos de datos elementales son descodificados, y los flujos de datos de audio y vídeo son distribuidos como elementos de programa sincronizados, al subsistema de presentación del receptor, para su visualización como partes de un programa coordinado.

20 En muchos estándares de codificación de vídeo, un flujo de bits compatible con el descodificador es descodificado mediante un descodificador hipotético que está conectado conceptualmente a la salida de un codificador, y consiste en una memoria tampón de descodificador, un descodificador y una unidad de visualización. Este descodificador virtual se conoce como descodificador hipotético de referencia (HRD, hypothetical reference decoder) en H.263, y como verificador de memoria tampón de vídeo (VBV, video buffering verifier) en MPEG-2. El codificador crea un flujo de bits de tal forma que la memoria tampón del descodificador hipotético no produce un desbordamiento ni un subdesbordamiento.

30 Como resultado, la cantidad de datos que puede solicitarse al receptor que almacene en memoria tampón, puede exceder su capacidad (una condición desbordamiento de memoria), o sus capacidades de caudal de tráfico. Alternativamente el receptor puede no recibir, en una unidad de acceso de datos, todos los datos a tiempo para la descodificación y la presentación sincronizada, lo que en un instante concreto tiene como resultado que se pierde datos en los flujos de datos de audio y vídeo, así como un funcionamiento inconsistente (una condición de subdesbordamiento de memoria).

35 En las descodificadores hipotéticos de referencia existentes, el flujo de bits de vídeo se recibe a una velocidad binaria constante, dada (usualmente la velocidad promedio en bits/segundo del flujo), y se almacena en la memoria tampón del descodificador hasta que el llenado alcanza un nivel deseado. Tal nivel deseado se denota como el llenado inicial de memoria tampón del descodificador, y es directamente proporcional al retardo de transmisión o de puesta funcionamiento (memoria tampón). En este punto, el descodificador retira instantáneamente los bits para la primera imagen de vídeo de la secuencia, descodifica los bits y muestra la imagen. Los bits de las siguientes imágenes son también retirados, descodificados, y visualizados instantáneamente en los subsiguientes intervalos temporales.

45 Las descodificadores hipotéticos tradicionales funcionan a velocidad binaria, memoria tampón y retardo inicial fijos. Sin embargo, en muchas de las aplicaciones de vídeo actuales (por ejemplo, emisión de vídeo en tiempo real a través de Internet o de redes ATM) el ancho de banda disponible varía de acuerdo con la ruta de acceso de la red (por ejemplo, dependiendo de como el usuario conecte con la red: por módem, ISDN, DLS, cable, etc.) y también fluctúa con el tiempo de acuerdo con las condiciones de la red (por ejemplo congestión, número de usuarios conectados, etc.). Además, los flujos de bits de vídeo son distribuidos a una variedad de dispositivos con diferentes capacidades de memoria tampón (por ejemplo microteléfonos, PDAs, PCs, módulos de conexión, reproductores de tipo DVD, etc.), y son creados para escenarios con diferentes requisitos de retardo (por ejemplo emisión directa de bajo retardo, descarga progresiva, etc.). Como resultado, estas aplicaciones necesitan un descodificador hipotético de referencia más flexible, que pueda descodificar un flujo de bits a diferentes velocidades binarias máximas, y con diferentes tamaños de memoria tampón y retardos de puesta en funcionamiento.

55 Jordi Ribas-Corbera y Philip A. Chou, en un documento titulado "A Generalized Hypothetical Reference Decoder For H.26L", el 4 de septiembre de 2001 propusieron el descodificador hipotético de referencia modificado. El descodificador funciona de acuerdo con N conjuntos de velocidad y parámetros de memoria tampón para un flujo de bits dado. Cada conjunto caracteriza lo que se conoce como un modelo de cubeta con goteo (leaky bucket) y contiene tres valores (R, B, F), donde R es la velocidad binaria de transmisión, B es el tamaño memoria tampón y F es el llenado inicial de memoria tampón del descodificador (F/R es el retardo de memoria tampón inicial o de puesta en funcionamiento). Un codificador puede crear un flujo de bits de vídeo que esté contenido en ciertas cubetas con goteo deseadas N, o puede simplemente computar los N conjuntos de parámetros después de que ha sido generado el flujo de bits. El descodificador hipotético de referencia puede interpolar entre los parámetros de cubeta con goteo, y puede funcionar a cualquier velocidad binaria, memoria tampón o retardo, máximos, deseados. Por ejemplo, dada una velocidad de transmisión máxima R' el descodificador de referencia puede seleccionar la memoria tampón y el retardo mínimos (de acuerdo con los datos disponibles de la cubeta con goteo) con los que será posible descodificar

el flujo de bits sin sufrir desbordamiento o subdesbordamiento de la memoria tampón. A la inversa, para un tamaño dado de memoria tampón B', el descodificador hipotético puede seleccionar y operar en la mínima velocidad de transmisión máxima requerida.

5 Hay beneficios en el uso de tal descodificador hipotético de referencia generalizado. Por ejemplo, un proveedor de contenidos puede crear un flujo de bits una vez, y un servidor puede distribuirlo a múltiples dispositivos de diferentes capacidades, utilizando una variedad de canales de velocidades máximas de transmisión, diferentes. O bien un servidor y un terminal pueden negociar la mejor cubeta con goteo para las condiciones de red dadas - por ejemplo, la que produzca el mínimo retardo de puesta en funcionamiento (memoria tampón), o la que requiera la menor
10 velocidad máxima de transmisión para el tamaño de datos de la memoria tampón del dispositivo.

Como se describe en el documento VCEG-58, secciones 2.1 – 2.4, una cubeta con goteo es un modelo para el estado (o llenado) de una memoria tampón de codificador o de descodificador, en función del tiempo. Los llenados de la memoria tampón del codificador y del descodificador son mutuamente complementarios. Un modelo de cubeta
15 con goteo está caracterizado por tres parámetros (R, B, F), donde:

R es la velocidad binaria máxima (en bits por segundo) a la que entran los bits en la memoria tampón del descodificador. En escenarios de velocidad binaria constante, R es a menudo la velocidad binaria del canal y la velocidad binaria promedio del videoclip.
20

B es el tamaño de la cubeta o memoria tampón del descodificador (en bits) que suaviza las fluctuaciones de la velocidad binaria de vídeo. Este tamaño de memoria tampón no puede ser mayor que la memoria tampón física del dispositivo descodificador.

F es el llenado inicial de memoria tampón del descodificador (también en bits) antes de que el descodificador comience a retirar bits desde la memoria tampón. F y R determinan el retardo inicial o de puesta funcionamiento D, donde $D = F / R$ segundos.
25

En un modelo de cubeta con goteo, los bits entran en la memoria tampón a la velocidad R, hasta que el nivel de llenado es F (es decir, durante D segundos), y a continuación se retira instantáneamente b_0 bits para la primera imagen. Los bits siguen entrando en la memoria tampón a la velocidad R, y el descodificador retira b_1, b_2, \dots, b_{n-1} bits para las siguientes imágenes, en ciertos momentos datos, típicamente (pero no necesariamente) cada $1/M$ segundos, donde M es la frecuencia de imagen del vídeo. La figura 1 ilustra el llenado de la memoria tampón del descodificador a lo largo del tiempo, de un flujo de bits que es modulado en una cubeta con goteo, de parámetros
30 (R, B, F).
35

Sea B_i el llenado de la memoria tampón del descodificador inmediatamente antes de retirar b_i bits en el instante t_i . Un modelo genérico de cubeta con goteo funciona de acuerdo con las siguientes ecuaciones:

$$B_0 = F$$

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

40 Típicamente, $t_{i+1} - t_i = 1/M$ segundos, donde M es la frecuencia de imagen (normalmente en imágenes/segundo) para el flujo de bits.

45 Un modelo de cubeta con goteo con parámetros (R, R, F) contiene un flujo de bits, si no hay subdesbordamiento de la memoria tampón del descodificador. Debido a que los llenados de las memorias tampón del codificador y del descodificador son complementarios entre sí, esto es equivalente a la ausencia de desbordamiento de la memoria tampón del codificador. Sin embargo, se permite que la memoria tampón (la cubeta con goteo) del codificador se vacíe, o equivalentemente que la memoria tampón del descodificador se llene, momento en el cual no se transmite
50 más bits desde la memoria tampón del codificador a la memoria tampón del descodificador. Así, la memoria tampón del descodificador detiene la recepción de bits cuando está llena, razón por la cual está incluido el operador min en la ecuación (1). Una memoria tampón de descodificador llena, significa simplemente que la memoria tampón del codificador está vacía.

55 Pueden realizarse las siguientes observaciones.

Un flujo de vídeo dado puede contenerse en muchas cubetas con goteo. Por ejemplo, si un flujo de vídeo está contenido en una cubeta con goteo con parámetros (R, B, F), estará también contenido en una cubeta con goteo con una memoria tampón mayor (R, B', F), $B' > B$, o en una cubeta con goteo con una velocidad de transmisión máxima superior (R', B, F), $R' > R$.
60

Para cualquier velocidad binaria R' el sistema puede siempre encontrar un tamaño de memoria tampón que contenga el flujo de bits de vídeo (limitado en el tiempo). En el peor de los casos (R' próximo a 0), será necesario que el tamaño de la memoria tampón sea tan grande como el propio flujo de bits. Dicho de otra forma, un flujo de bits de vídeo puede ser transmitido a cualquier frecuencia (independientemente de la velocidad binaria promedio del clip), siempre que el tamaño de la memoria tampón sea lo suficientemente grande.

Se asume que el sistema fija la relación $F = a B$ para todas las cubetas con goteo, donde a es alguna fracción deseada del llenado inicial de la memoria tampón. Para cada valor de la velocidad binaria máxima R, el sistema puede hallar el tamaño mínimo de memoria tampón B_{min} que contenga el flujo de bits, utilizando la ecuación (1). En la figura 2 se muestra la representación de la curva de valores R - B.

Por inspección, la curva de (R_{min}, B_{min}) para cualquier flujo de bits (como el de la figura 2) es lineal a tramos, y convexa. Por lo tanto si se proporciona N puntos de la curva, el descodificador puede interpolar linealmente los valores para llegar a algunos puntos (R_{interp}, B_{interp}) que son ligeramente, pero con seguridad, mayores que (R_{min}, B_{min}) . De este modo se puede reducir el tamaño de la memoria tampón, y por consiguiente también el retardo, en un orden de magnitud en relación con una sola cubeta con goteo que contenga el flujo de bits en su frecuencia promedio. Alternativamente, para el mismo retardo se puede reducir la frecuencia de transmisión máxima mediante un factor cuatro, o posiblemente incluso mejorar la relación señal frente a ruido en varios dB.

Verificador de Memoria Tampón de Vídeo MPEG (VBV)

El verificador de memoria tampón de vídeo MPEG (VBV) puede funcionar en dos modos: velocidad binaria constante (CBR, constant bit rate) y velocidad binaria variable (VBR, variable bit rate). MPEG-1 solo soporta el modo CBR, mientras que MPEG-2 soporta ambos modos.

El VBV funciona en modo CBR cuando el flujo de bits está contenido en un modelo de cubeta con goteo de parámetros (R, B, F) y:

$$R = R_{max} = \text{velocidad binaria promedio del flujo.}$$

El valor de B es almacenado en el parámetro sintáctico tamaño_de_memoria_tampón_vbv, utilizando una unidad especial de tamaño (a saber, unidades de 16 x 1024 bits).

El valor de F/R es almacenado en el elemento sintáctico retardo_vbv, asociado con la primera imagen de vídeo en la secuencia, utilizando una unidad especial de tiempo (a saber, número de períodos de un reloj de 90 kHz).

El llenado de la memoria tampón del descodificador sigue las siguientes ecuaciones:

$$B_0 = F$$

$$B_{i+1} = B_i - b_i + R_{max}/M, \quad i = 0, 1, 2, \dots \quad (2)$$

El codificador tiene que asegurar que $B_i - b_i$ es siempre mayor o igual que cero, siendo a la vez B_i es siempre menor o igual que B. En otras palabras, el codificador asegura que la memoria tampón del descodificador no presente un desbordamiento ni un subdesbordamiento.

El VBV funciona en modo VBR cuando el flujo de bits está limitado en un modelo de cubeta con goteo de parámetros (R, B, F) y:

$R = R_{max}$ = pico de la frecuencia máxima. R_{max} es mayor que la frecuencia promedio del flujo de bits.

$F = B$, es decir la memoria tampón se llena inicialmente.

El valor de R está representado en el parámetro sintáctico tamaño_de_memoria_tampón_vbv, como en el caso de CBR.

El llenado de la memoria tampón del descodificador sigue las siguientes ecuaciones:

$$B_0 = B$$

$$B_{i+1} = \min (B, B_i - b_i + R_{\max}/M), i = 0, 1, 2, \dots \quad (3)$$

5 El codificador asegura que $B_i - b_i$ es siempre mayor o igual que cero. Es decir, el codificador debe asegurar que en la memoria tampón del descodificador no presenta subdesbordamiento. Sin embargo, en este caso de VBR el codificador no necesita asegurar que la memoria tampón del descodificador no presenta desbordamiento. Si la memoria tampón del descodificador se llena, entonces se asume que la memoria tampón del codificador está vacía y por tanto se deja de transmitir bits desde la memoria tampón del codificador a la memoria tampón del descodificador.

10 El modo VBR es útil para dispositivos que pueden leer datos hasta la frecuencia máxima R_{\max} . Por ejemplo, un DVD incluye clips de VBR en los que R_{\max} es aproximadamente 10 Mb/s, lo que corresponde a la máxima velocidad de lectura de la unidad de disco, incluso aunque la frecuencia promedio del flujo de vídeo DVD sea de solo unos 4 Mb/s.

15 En referencia la figura 3A y a la figura 3B, se muestra gráficos de llenado de la memoria tampón del descodificador, para algunos flujos de bits operativos en modo CBR y VBR.

En líneas generales, el modo CBR puede considerarse un caso especial de VBR donde se tiene que R_{\max} es la frecuencia promedio del clip.

20 Descodificador Hipotético de Referencia (HRD) H.263

El modelo hipotético de referencia para H.263 es similar al modo CBR del VBV de MPEG discutido previamente, excepto por lo siguiente.

25 El descodificador inspecciona el llenado de la memoria tampón en ciertos intervalos temporales, y descodifica una imagen tan pronto estén disponibles todos los bits para la imagen. Este enfoque tiene como resultado un par de beneficios: (a) el retardo se minimiza debido a que F es usualmente solo un poco mayor que el número de bits para la primera imagen, y (b) si es común el salto de imágenes, el decodificador simplemente espera hasta la siguiente imagen disponible. Lo segundo se habilita también en el modo de bajo retardo del VBV de MPEG.

30 La comprobación del desbordamiento de la memoria tampón se realiza después de que los bits para una imagen son retirados de la memoria tampón. Esto relaja la limitación de enviar grandes imágenes I una vez cada cierto tiempo, pero hay un valor máximo para la imagen mayor.

35 El HRD de H.263 puede esencialmente ser mapeado a un tipo de modelo de cubeta con goteo, de bajo retardo.

Limitaciones de los Descodificadores Hipotéticos de Referencia Previos

40 Los descodificadores hipotéticos de referencia existentes previamente, funcionan en un solo punto (R, B) de la curva de la figura 2. Como resultado, estos descodificadores tienen los siguientes inconvenientes:

45 si la velocidad binaria disponible en el canal R' es menor que la disponible en R (por ejemplo, esto es común en la emisión en tiempo real por Internet y en la descarga progresiva, o cuando se requiere transmitir un clip VBR en MPEG a una frecuencia inferior al máximo), estrictamente hablando el descodificador hipotético no sería capaz de descodificar el flujo de bits.

50 Si el ancho de banda disponible R' es mayor que R (por ejemplo, esto es común también para la emisión en tiempo real por Internet, así como para la reproducción local), los descodificadores hipotéticos previos podrían funcionar en el modo VBR y descodificarían el flujo de bits. Sin embargo, si hubiera disponible más información sobre la curva Frecuencia - Memoria Tampón, se reduciría significativamente el tamaño de memoria tampón y el retardo asociado de puesta en funcionamiento, requerido para descodificar el flujo de bits.

55 Si el tamaño físico de la memoria tampón en un dispositivo descodificador es menor que R, el dispositivo no será capaz de descodificar tal flujo de bits.

60 Si el tamaño de memoria tampón es mayor que B, el dispositivo será capaz de descodificar el flujo de bits pero el retardo de puesta en funcionamiento será el mismo.

De forma más general, un flujo de bits que se ha generado de acuerdo con una cubeta con goteo (R, B, F) usualmente no será susceptible de ser distribuido a través de diferentes redes de velocidad binaria menor que R, y a una variedad de dispositivos con tamaños de memoria tampón menores que B. Además, no será minimizado el retardo de puesta en funcionamiento.

5 Descodificador Hipotético de Referencia Generalizado (GHRD)

Un decodificador hipotético de referencia generalizado (GHRD) puede funcionar dada la información de N modelos de cubeta con goteo,

$$10 \quad (R_1, B_1, F_1), (R_2, B_2, F_2), \dots, (R_N, B_N, F_N), \quad (4)$$

cada uno de los cuales contiene el flujo de bits. Sin perder generalidad, asumimos que estas cubetas con goteo están ordenadas desde la menor a la mayor velocidad binaria, es decir que $R_i < R_{i+1}$. Asumimos también que el codificador computa estos modelos de cubeta de goteo correctamente, y por lo tanto $B_i < B_{i+1}$.

15 El valor deseado de N puede ser seleccionado mediante el codificador. Si $N = 1$, el GHRD es esencialmente equivalente al VBV del MPEG. El codificador puede elegir: (a) preseleccionar valores de cubeta con goteo y codificar el flujo de bits con un control de velocidad, asegurando que se satisface la totalidad de las limitaciones de cubeta de goteo, (b) codificar el flujo de bits y después utilizar la ecuación (1) para computar un conjunto de cubetas con goteo que contiene el flujo de bits a N diferentes valores de R, o (c) hacer ambas cosas. El primer enfoque (a) puede aplicarse a transmisión en directo o bajo demanda, mientras que (b) o (c) aplican sólo a transmisión bajo demanda.

20 El número N de cubetas con goteo y los parámetros de cubeta con goteo (4), son insertados en el flujo de bits. De este modo, el decodificador puede determinar qué cubeta con goteo desea utilizar, conociendo la velocidad binaria máxima disponible para esta, y/o su tamaño físico de memoria tampón. Los modelos de cubeta con goteo en (4), así como todos los modelos linealmente interpolados o extrapolados, son de uso viable. La figura 4 ilustra un conjunto de N modelos de cubeta con goteo, y sus valores interpolados o extrapolados (R, B).

25 El tamaño de memoria tampón interpolado B entre puntos k y k + 1, sigue la línea recta:

$$30 \quad B = \{(R_{k+1}-R)/(R_{k+1}-R_k)\}B_k + \{(R-R_k)/(R_{k+1}-R_k)\}B_{k+1} \quad R_k < R < R_{k+1}$$

Análogamente, el llenado inicial de la memoria tampón del decodificador F puede interpolarse linealmente:

$$F = \{(R_{k+1}-R)/(R_{k+1}-R_k)\}F_k + \{(R-R_k)/(R_{k+1}-R_k)\}F_{k+1} \quad R_k < R < R_{k+1}$$

35 La cubeta con goteo resultante con parámetros (R, B, F) contiene el flujo de bits, debido al que el tamaño mínimo de memoria tampón B_{\min} es convexo tanto en R como en F, es decir el tamaño mínimo de memoria tampón 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 o igual que $B = a B_k + (1 - a) B_{k+1}$.

40 Se observa que si R es mayor que R_N, la cubeta con goteo (R, B_N, F_N) contendrá también el flujo de bits, y por lo tanto B_N y F_N son el tamaño de memoria tampón y el llenado inicial de la memoria tampón del decodificador, recomendados cuando $R \geq R_N$. Si R es menor que R_i puede provocarse el límite superior $B = B + (R_i - R) T$ (y la vez se puede fijar F = B), donde T es la duración temporal del flujo en segundos. Estos valores (R, B) fuera del rango de los N puntos, se muestran también en la figura 4.

45 El documento Working Draft Number 2, revisión 0 (WD-2), del Joint Video Team de ISO/IEC MPEG y ITU-T VCEG, incorpora muchos de los conceptos del decodificador hipotético de referencia propuesto por Jordi Ribas-Cobera et al., de Microsoft Corporation, incorporado aquí como referencia. El documento WD-2 es similar al decodificador propuesto por Jordi Ribas-Cobera et al., de Microsoft Corporation, aunque la sintaxis está algo modificada. Además, WD-2 describe un algoritmo a modo de ejemplo para computar B, y F para una frecuencia dada R.

50 La publicación de aplicación de patente de EE.UU. número US 2003/0 053 416 A, describe el uso de dos modelos de cubeta con goteo. El primer conjunto de parámetros de cubeta con goteo, permitiría la transmisión del vídeo sobre un canal con velocidad binaria constante, con un retardo de unos 22,5 segundos. El segundo conjunto de parámetros de cubeta con goteo permitiría la transmisión del vídeo sobre una red compartida, con una velocidad máxima de 2500 kbps, o permitiría la reproducción local desde un CD 2x, con un retardo de unos 0,9 segundos.

Resumen de la invención

De acuerdo con un aspecto de la presente invención, se proporciona un método de codificación de vídeo que comprende definir un primer conjunto de múltiples valores característicos de una velocidad binaria de transmisión para un primer punto de acceso en un punto de comienzo de una secuencia de vídeo; definir un segundo conjunto de múltiples valores característicos de un tamaño de memoria tampón para dicho primer punto de acceso; definir un tercer conjunto de múltiples valores característicos de un retardo hasta que se presenta una secuencia de vídeo para dicho primer punto de acceso; definir un cuarto conjunto de múltiples valores característicos de una velocidad binaria de transmisión para otro punto de acceso localizado después de dicho primer punto de acceso y a continuación; definir un quinto conjunto de múltiples valores característicos de un tamaño de memoria tampón para dicho otro punto de acceso; definir un sexto conjunto de múltiples valores característicos de un retardo hasta que se presenta una secuencia de vídeo para dicho otro punto de acceso en el que se seleccionan un valor de dicho primer conjunto de múltiples valores, un valor de dicho segundo conjunto de múltiples valores, y un valor de dicho tercer conjunto de múltiples valores tal que dicha secuencia de vídeo está libre de una condición de desbordamiento de memoria tampón en dicho primer punto de acceso; y se seleccionan un valor de dicho cuarto conjunto de múltiples valores, un valor de dicho quinto conjunto de múltiples valores, y dicho sexto conjunto de múltiples valores tal que dicha secuencia de vídeo está libre de una condición de desbordamiento de memoria tampón en dicho otro punto de acceso.

Características preferidas de la invención se exponen en las reivindicaciones dependientes

Breve descripción de los dibujos

La figura 1 ilustra un llenado de memoria tampón del descodificador.

La figura 2 ilustra una curva R - B.

Las figuras 3A y 3B son representaciones del llenado de la memoria tampón del descodificador, para algunos flujos de bits operativos respectivamente en modos CBR y VBR.

La figura 4 ilustra un conjunto de N modelos de cubeta con goteo, y sus valores interpolados o extrapolados (R, B).

La figura 5 ilustra el almacenamiento inicial en memoria tampón B_j del descodificador, para cualquier punto que el usuario busque cuando la velocidad es R_j .

La figura 6 ilustra conjuntos de (R, B, F) definidos en modo avance, para el flujo de vídeo concreto.

La figura 7 ilustra el llenado inicial de memoria tampón (en bits) para un segmento de vídeo.

La figura 8 ilustra el criterio de selección en un conjunto de 10 puntos para la figura 7.

La figura 9 ilustra criterios de selección.

La figura 10 ilustra reducciones en el retardo.

Descripción detallada de la realización preferida

Como se ha descrito previamente, el estándar JVT (WD-2) permite el almacenamiento de ($N \geq 1$) cubetas con goteo, valores $(R_1, B_1, F_1), \dots, (R_N, B_N, F_N)$ que están contenidos en el flujo de bits. Estos valores pueden ser almacenados en el encabezado. Utilizando F_i como el llenado inicial de la memoria tampón y B_i como el tamaño de la memoria tampón, garantiza que la memoria tampón del descodificador no presentará subdesbordamiento cuando el flujo de entrada entre a la velocidad R_i . Este será el caso si el usuario desea presentar el vídeo codificado de principio a fin. En una típica aplicación de vídeo bajo demanda, el usuario puede querer buscar diferentes partes del flujo de vídeo. El punto que el usuario desea buscar puede ser aludido como el punto de acceso. Durante el proceso de recepción de datos de vídeo y construcción de imágenes de vídeo, la cantidad de datos en la memoria tampón fluctúa. Tras su consideración, el presente inventor llega a la comprensión de que si el valor F_i de llenado inicial de la memoria tampón (cuando la velocidad del canal es R_i) se utiliza antes de comenzar a descodificar el vídeo desde el punto de acceso, entonces es posible que el descodificador presente un subdesbordamiento. Por ejemplo en el punto de acceso o algo después, la cantidad de bits necesarios para la reconstrucción de vídeo puede ser mayor que los bits que hay actualmente en la memoria tampón, lo que tiene como resultado un subdesbordamiento y la imposibilidad de presentar imágenes de vídeo de forma oportuna. Igualmente, puede demostrarse que en un flujo de vídeo el valor de llenado inicial de memoria tampón, necesario para asegurarse de que no hay subdesbordamiento en el descodificador, varía en función del punto que busca el usuario. Este valor está limitado por el B_i . Por consiguiente, la combinación de B y F proporcionados para toda la secuencia de vídeo, probablemente no será apropiada si se utiliza para un punto intermedio en el vídeo, teniendo como resultado un subdesbordamiento y por tanto congelando las imágenes.

Basándose en este potencial de subdesbordamiento no comprendido previamente, el presente inventor llega entonces a la constatación de que si se define solo un conjunto de valores de R, B y F para un segmento de vídeo completo, entonces el sistema debería esperar hasta que la memoria tampón B para la correspondiente velocidad R esté llena, o sustancialmente llena (o llena por encima del 90%), para comenzar a decodificar imágenes cuando un usuario salta a un punto de acceso. De este modo el llenado inicial de la memoria tampón estará al máximo, y por lo tanto no existe el potencial de subdesbordamiento durante la subsiguiente decodificación que comience desde el punto de acceso. Esto puede conseguirse sin ningún cambio adicional en el flujo de bits existente, y por lo tanto sin impacto sobre los sistemas existentes. Por consiguiente, el descodificador utilizaría el valor del almacenamiento inicial en memoria tampón B_j , para cualquier punto que busque el usuario, cuando la velocidad es R_j , como se muestra en la figura 5. Sin embargo, desgraciadamente en ocasiones esto tiene como resultado un retardo significativo hasta que se presenta las imágenes, después de seleccionar una localización diferente (por ejemplo, un punto de acceso) desde la que presentar el vídeo.

El llenado inicial de memoria tampón (F) puede análogamente caracterizarse como un retardo hasta que se presenta la secuencia de vídeo. El retardo es de naturaleza temporal, estando relacionado con el tiempo necesario para conseguir el llenado inicial de memoria tampón (F). El retardo y/o F pueden asociarse con todo el vídeo o con los puntos de acceso. Análogamente, debe entenderse que el retardo puede sustituirse por F en todas las realizaciones aquí descritas (por ejemplo (R, B, retardo)). Un valor concreto para el retardo puede calcularse como $\text{retardo} = F/R$, utilizando una unidad especial de tiempo (unidades de reloj de 90 kHz).

Para reducir el retardo potencial, el presente inventor llegó a la comprensión de que puede definirse conjuntos de (R, B, F) para un flujo de vídeo concreto, en cada punto de acceso. En referencia a la figura 6, estos conjuntos de (R, B, F) se definen preferentemente en modo avance, para el flujo de vídeo concreto. Por ejemplo, el conjunto de valores (R, B, F) puede computarse de la forma existente previamente para el flujo de vídeo como un todo, además el conjunto de valores F para los mismos valores (R, B) que los de todo el flujo de vídeo, puede computarse de la forma existente previamente para el flujo de vídeo, con respecto al flujo de vídeo desde la posición "2" en avance, etcétera. El mismo proceso puede utilizarse para los restantes puntos de acceso. Los puntos de acceso pueden ser cualquier imagen dentro de la secuencia de vídeo, imágenes I de la secuencia, imágenes B de la secuencia, o imágenes P de la secuencia (se utiliza típicamente imágenes I, B y P en la decodificación de vídeo basada en MPEG). Por consiguiente, el usuario puede seleccionar uno de los puntos de acceso, y a continuación utilizar el respectivo F_{ij} para el llenado inicial deseado (asumiendo que la memoria tampón B_i y la velocidad R_i permanecen invariables), o en otro caso un conjunto de dos o más de R_i, B_i, F_{ij} .

El índice "i" representa cada cubeta con goteo, y el índice "j" representa cada punto de acceso aleatorio. Asumiendo que la memoria tampón B_i y la velocidad R_i permanecen invariables, en el encabezado se almacena el conjunto múltiple de valores (B_i, R_i, F_{i1}), donde $i = 1, 2, \dots, N$ y F_{i1} representa el llenado inicial de la memoria tampón. Después, en el punto de acceso j se almacena F_{ij} , donde $j = 2, 3, \dots$. Por otra parte, asumiendo que la memoria tampón B_i y la velocidad R_i serán modificadas en cada punto de acceso j, puede almacenarse múltiples conjuntos de valores de (R_{ij}, B_{ij}, F_{ij}) en cada punto de acceso. El beneficio del primer caso es que ahorra en la cantidad de datos, puesto que sólo se ha almacenado un conjunto múltiple de F_{ij} en cada punto de acceso, y el beneficio del último caso es que puede ajustar el conjunto de valores de forma más apropiada para cada punto de acceso. Cuando se utiliza el retardo (D) hasta que se presenta la secuencia de vídeo, en lugar del llenado inicial de la memoria tampón (F), es posible realizar la invención actual mediante sustituir F_{ij} con D_{ij} . En este caso D_{ij} representa el valor del retardo. Así, cuando se asume que la memoria tampón B_i y la velocidad R_i permanecen invariables, se almacena (B_i, R_i, D_{ij}) en el encabezado y se almacena D_{ij} en cada punto de acceso j. Cuando se asume que la memoria tampón B_i y la velocidad R_i cambian en cada punto de acceso j, puede almacenarse un conjunto múltiple de valores (B_{ij}, R_{ij}, D_{ij}) en el punto de acceso.

Los conjuntos de valores R, B, F para cada punto de acceso, pueden estar localizados en cualquier localización apropiada, tal como por ejemplo al comienzo de la secuencia de vídeo, junto con conjuntos de valores (R, B, F) para todo el flujo de vídeo, o antes de cada punto de acceso lo que evita la necesidad de un índice; o estar almacenados de forma externa al propio flujo de vídeo, lo que es especialmente adecuado para un entorno de servidor/cliente.

Esta técnica puede caracterizarse mediante el siguiente modelo:

$$(R_1, B_1, F_1, M_1, f_{11}, t_{11}, \dots, f_{M11}, t_{M11}) \dots, (R_N, B_N, F_N, M_N, f_{1N}, t_{1N}, \dots, f_{MNN}, t_{MNN}),$$

donde f_{kj} denota el valor inicial de llenado de la memoria tampón a la velocidad R_j en el punto de acceso t_{kj} (marca temporal). Los valores de M_j pueden proporcionarse como un parámetro de entrada, o pueden ser seleccionados automáticamente. Por ejemplo, M_j puede incluir las siguientes opciones:

(a) M_j puede ajustarse al valor igual al número de puntos de acceso. De este modo los valores de f_{kj} puede almacenarse para cada punto de acceso, a cada velocidad R_j (bien al comienzo del flujo de vídeo, dentro del flujo de vídeo, distribuido a través del flujo de vídeo, o con cualquier otra localización).

(b) M_j puede fijarse igual a cero, sino se desea soporte de búsqueda.

(c) Puede seleccionarse automáticamente M_j valores para cada velocidad R_j (se describe más abajo).

5 Para un R_j dado, el sistema puede utilizar un llenado inicial de memoria tampón igual a f_{jk} si el usuario busca un punto de acceso t_{kj} . Esto ocurre cuando el usuario selecciona comenzar en un punto de acceso, o bien cuando el sistema ajusta la selección del usuario a uno de los puntos de acceso.

10 Se hace notar que en el caso de que se utiliza una velocidad binaria variable (en flujo de bits), preferentemente el valor de llenado inicial de la memoria tampón (o retardo) es diferente al tamaño de memoria tampón (o el retardo calculado por el tamaño de memoria tampón), aunque puede ser el mismo. En el caso de una velocidad binaria variable en VBV de MPEG-2, la memoria tampón se rellena hasta que está llena, es decir $F = B$ (el valor de B está representado por tamaño_de_memoria_tampón_de_vbv).

15 En la presente invención, el valor de llenado inicial del tamaño de la memoria tampón, puede escogerse apropiadamente en cada punto de acceso aleatorio, para evitar cualquier subdesbordamiento o desbordamiento de la memoria tampón. Cuando se usa el retardo hasta que se presenta la secuencia de vídeo, en lugar del llenado de la memoria tampón, el valor del retardo se escoge apropiadamente en cada punto de acceso aleatorio, para evitar cualquier subdesbordamiento o desbordamiento de la memoria tampón. Generalmente, esto significa que en cada punto de acceso aleatorio se consigue un retardo menor que llenando del todo la memoria tampón VBV. Por lo tanto, la determinación del valor de llenado de la memoria tampón (o el retardo) que es menor que el tamaño de la memoria tampón (o el retardo calculado por el tamaño de la memoria tampón) mediante la presente invención, tiene la ventaja de un retardo reducido, puesto que se requiere almacenar menos datos en la memoria tampón antes del comienzo de la descodificación, que en el arte previo.

25 Si el sistema permite al usuario saltar a cualquier imagen del vídeo, a modo de un punto de acceso, entonces se necesitaría proporcionar el conjunto de datos de descodificación para todas y cada una de las imágenes. En el caso de que fuera permisible, el conjunto de datos resultantes sería excesivamente grande y consumiría una cantidad significativa de la velocidad binaria disponible para los datos. Un enfoque más razonable sería limitar al usuario a puntos de acceso específicos dentro del flujo de vídeo, tal como cada segundo, cada 10 segundos, cada minuto, etcétera. Aun siendo una mejora, el conjunto de datos resultante puede seguir siendo algo extenso, teniendo como resultado excesivos datos para dispositivos con ancho de banda limitado, tales como dispositivos de comunicación móvil.

35 En el caso de que el usuario seleccione una posición que sea uno de los puntos de acceso con un conjunto de datos asociado, entonces el llenado inicial de memoria tampón puede ser igual a $\max(\hat{f}_{kj}, \hat{f}_{(k+1)j})$ para un tiempo entre t_{kj} y $t_{(k+1)j}$, especialmente si los puntos de acceso se seleccionan adecuadamente. De este modo se garantiza que el sistema tiene un conjunto de valores que estará libre de tener como resultado una condición de subdesbordamiento, o en otro caso reduce la probabilidad de una condición de subdesbordamiento como se explicará más abajo.

40 Se hace referencia a la figura 7, para seleccionar un conjunto de valores que asegure que no se produce condición de subdesbordamiento (o en otro caso, que se reduce) cuando se ha utilizado el criterio de selección aludido arriba. La figura 7 ilustra el llenado inicial de la memoria tampón (en bits) para un segmento de vídeo, donde se calcula en llenado inicial de la memoria tampón en avance, para incrementos de 10 segundos. Después, el sistema selecciona preferentemente un punto de acceso en el comienzo de la secuencia de vídeo, y un punto de acceso al término del segmento de vídeo. Entre el comienzo y al final del segmento de vídeo, el sistema selecciona los máximos locales para su inclusión como puntos de acceso. Además, el sistema puede seleccionar los mínimos locales para su inclusión como puntos de acceso. Preferentemente, si se desea un conjunto limitado de puntos de acceso el sistema selecciona primero los máximos locales y después los mínimos locales, lo que ayuda a asegurar que no se produce subdesbordamiento. A continuación, si se desea el sistema puede seleccionar además puntos intermedios.

50 Basándose en el criterio de selección, puede seleccionarse un conjunto de 10 puntos para la figura 7, como se indica en la figura 8. En referencia la figura 9, los 10 puntos seleccionados se muestran mediante la curva a trazos. Los valores resultantes de llenado inicial de la memoria tampón en todos los puntos de acceso, se muestran mediante la curva continua. La curva continua ilustra un conjunto "seguro" de valores para todos los puntos de acceso en el vídeo, de forma que la memoria tampón del descodificador no presente subdesbordamiento. Si se ha producido fluctuaciones extremas en la velocidad binaria del flujo de bits actual, que no se detectaron en el procesamiento, tales como picos abruptos, entonces es posible que se tenga un subdesbordamiento, aunque normalmente es improbable. Los valores óptimos de llenado inicial de la memoria tampón en todos los puntos de acceso, se muestran mediante una curva a trazos y puntos. Se consigue una reducción significativa en el retardo temporal de almacenamiento en memoria tampón, en contraste con requerir una memoria tampón llena cuando se accede a un punto de acceso, como se ilustra en la figura 10.

65 Además, si la velocidad binaria y el tamaño de memoria tampón permanecen iguales mientras se selecciona un punto de acceso diferente, entonces simplemente necesita proporcionarse o determinarse de otro modo el llenado modificado de la memoria tampón, F .

REIVINDICACIONES

1. Método de codificación de vídeo que comprende:
- 5 definir un primer conjunto de múltiples valores (R_{i1}) característicos de una velocidad binaria de transmisión para un primer punto de acceso en un punto de comienzo de una secuencia de vídeo;
- definir un segundo conjunto de múltiples valores (B_{i1}) característicos de un tamaño de memoria tampón para dicho primer punto de acceso;
- 10 definir un tercer conjunto de múltiples valores (D_{i1}) característicos de un retardo hasta que se presenta una secuencia de vídeo para dicho primer punto de acceso;
- definir un cuarto conjunto de múltiples valores (R_{i2}) característicos de una velocidad binaria de transmisión para otro punto de acceso localizado después de dicho primer punto de acceso y a continuación;
- 15 definir un quinto conjunto de múltiples valores (B_{i2}) característicos de un tamaño de memoria tampón para dicho otro punto de acceso;
- 20 definir un sexto conjunto de múltiples valores (D_{i2}) característicos de un retardo hasta que se presenta una secuencia de vídeo para dicho otro punto de acceso; en el que
- se seleccionan un valor de dicho primer conjunto de múltiples valores, un valor de dicho segundo conjunto de múltiples valores, y un valor de dicho tercer conjunto de múltiples valores tal que dicha secuencia de vídeo está libre de una condición de desbordamiento de memoria tampón en dicho primer punto de acceso;
- 25 y
- se seleccionan un valor de dicho cuarto conjunto de múltiples valores, un valor de dicho quinto conjunto de múltiples valores, y un valor de dicho sexto conjunto de múltiples valores tal que dicha secuencia de vídeo está libre de una condición de desbordamiento de memoria tampón en dicho otro punto de acceso.
- 30
2. Método de codificación de vídeo según la reivindicación 1, en el que dicho primer conjunto: de múltiples valores (R_{i1}), dicho segundo conjunto de múltiples valores (B_{i1}), y dicho tercer conjunto de múltiples valores (D_{i1}) definen al menos un modelo de cubeta con goteo para una memoria tampón de un decodificador hipotético de referencia.
- 35
3. Método de codificación de vídeo según la reivindicación 2, en el que dicho modelo de cubeta con goteo usa una velocidad binaria fija.
- 40
4. Método de codificación de vídeo según la reivindicación 2, en el que dicho modelo de cubeta con goteo usa una velocidad binaria variable.

FIG.1

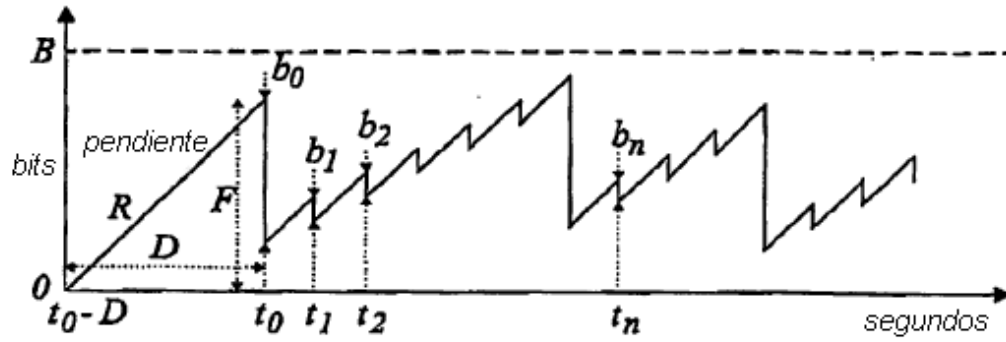


FIG.2

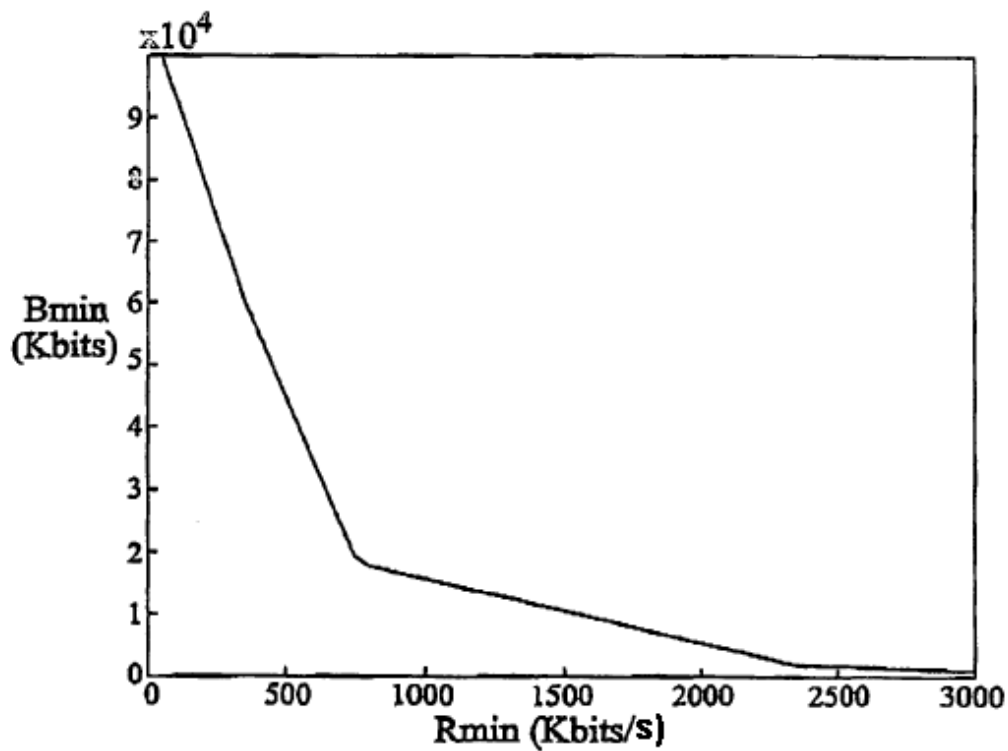


FIG.3A

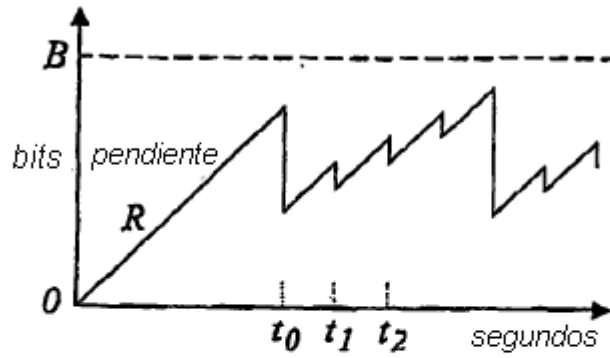


FIG.3B

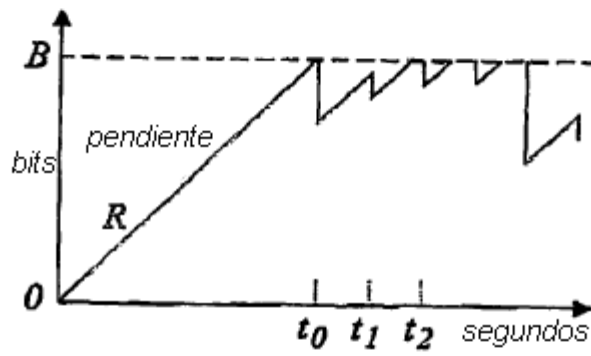


FIG.4

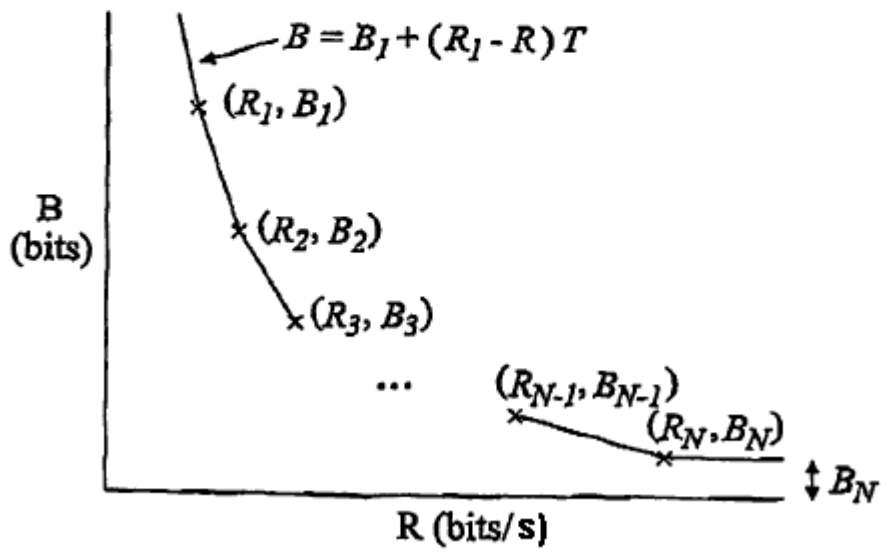


FIG.5

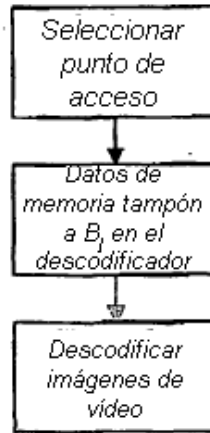


FIG.6

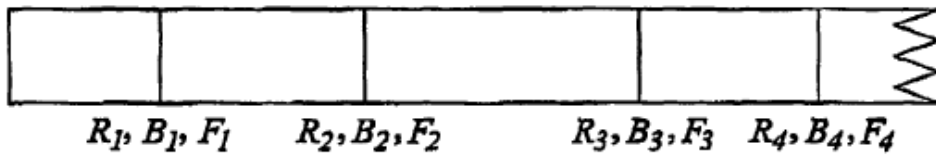


FIG.7

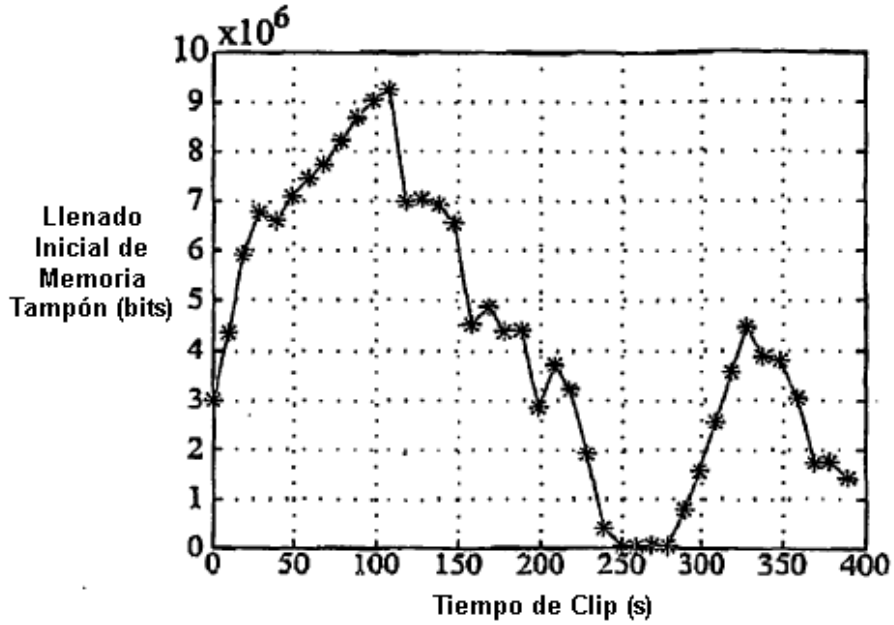


FIG.8

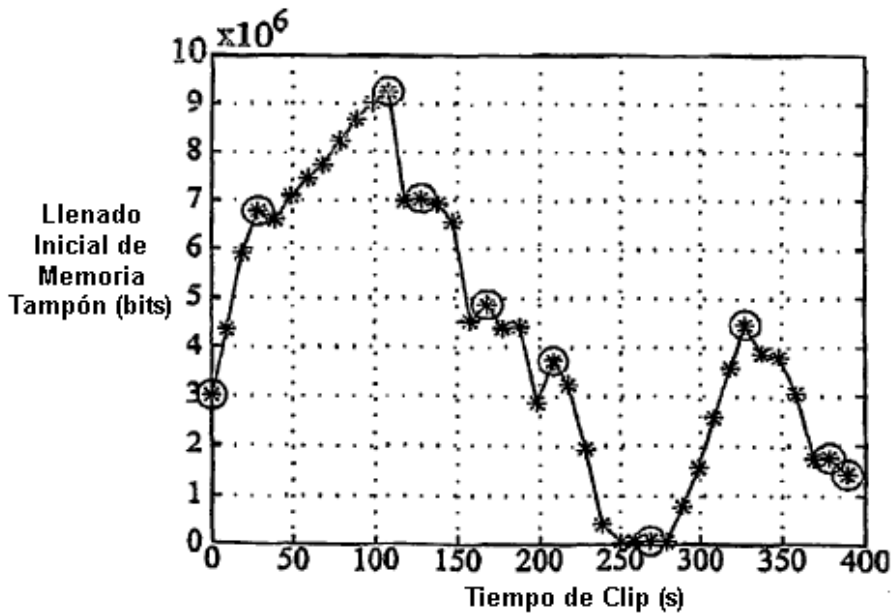


FIG.9

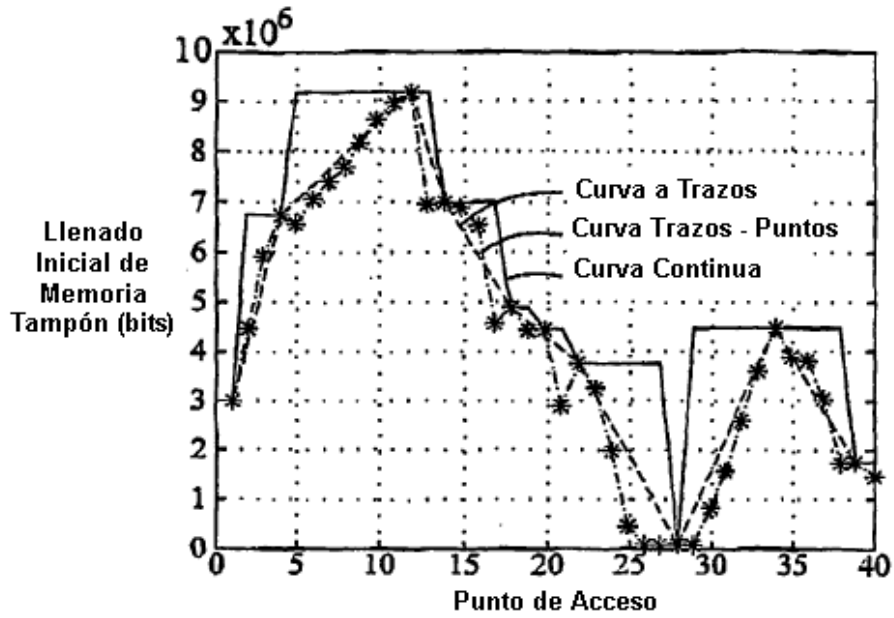


FIG.10

