



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 314 259**

51 Int. Cl.:
H04N 7/50 (2006.01)
H04N 7/30 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **03775577 .4**
96 Fecha de presentación : **18.11.2003**
97 Número de publicación de la solicitud: **1563689**
97 Fecha de publicación de la solicitud: **17.08.2005**

54 Título: **Transmisión de vídeo.**

30 Prioridad: **18.11.2002 GB 0226872**
29.09.2003 GB 0322767

45 Fecha de publicación de la mención BOPI:
16.03.2009

45 Fecha de la publicación del folleto de la patente:
16.03.2009

73 Titular/es: **BRITISH TELECOMMUNICATIONS**
public limited company
BT Group Legal, Intellectual Property Department
PP C5A, BT Centre, 81 Newgate Street
London EC1A 7AJ, GB

72 Inventor/es: **Ghanbari, Mohammed y**
Sun, Kai

74 Agente: **Curell Suñol, Marcelino**

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

Transmisión de vídeo.

5 La presente invención se refiere a la transmisión de señales de vídeo codificadas digitalmente (por ejemplo, a través de una red de telecomunicaciones) y, más particularmente, de señales de vídeo que se han codificado mediante un algoritmo de compresión.

10 La base lógica de los algoritmos de compresión consiste en explotar la redundancia inherente de la señal de vídeo original para reducir el número de bits que se necesita transmitir. Muchos de dichos algoritmos se definen en las normas internacionales, tales como las normas ITU H. 263 e ISO MPEG. En el documento de Ghanbari, M., Video Coding, an introduction to Standard codecs, IEE, Londres, 1999, se proporciona una reseña útil de dichos algoritmos.

15 Como es natural, el grado de redundancia (y por lo tanto la eficacia de compresión) varía con el contenido de la imagen, hecho que da por resultado un número variable de bits codificados por trama. Una opción consiste en transmitir los bits a medida que se generan, como en los denominados sistemas de tasa de bits variable (VBR) en los cuales la tasa de bits transmitidos varía considerablemente con el tiempo. Otra opción - la de los sistemas de tasa de bits constante (CBR) - consiste en emplear una memoria tampón tanto en el transmisor como en el receptor para eliminar estas fluctuaciones y transmitir desde la memoria tampón de transmisión hasta la memoria tampón de recepción a una velocidad constante. Los sistemas CBR utilizan un mecanismo de retroalimentación para variar la tasa de generación de los datos (por ejemplo, ajustando el nivel de detalle de la cuantificación utilizada o descartando tramas) y prevenir el desbordamiento de la memoria tampón. La utilización del almacenamiento en memoria tampón conlleva necesariamente la introducción de un retardo, el incremento de la latencia de inicio (LOS) - es decir, el usuario debe esperar a que la memoria tampón de recepción se llene hasta el nivel necesario para poder comenzar la decodificación y la presentación de las imágenes. El mecanismo de retroalimentación conlleva la reducción de la calidad de la imagen.

20 También se ha propuesto emplear un grado de almacenamiento en memoria tampón para reducir, aunque no eliminar por completo, las variaciones de la tasa de bits (véase, por ejemplo, el documento de Furini, M. y Towsley, D.F., "Real-Time Traffic transmissions over the Internet", IEEE Transactions on Multimedia, Vol. 3, nº 1, marzo de 2001).

25 Un factor importante a tener en cuenta cuando se transmite a través de una red de telecomunicaciones, y en particular a través de redes de paquetes tal como Internet, es el efecto de la congestión de la red, en el que la pérdida de paquetes y los retardos impredecibles pueden ocasionar problemas. Esto ha dado lugar a la presentación de propuestas de sistemas de reserva, en los que un transmisor puede solicitar a la red que asigne una tasa de bits garantizada específica para sus transmisiones durante un período de tiempo. Uno de dichos sistemas denominado "RSVP" se describe en el documento RFC 2205 del grupo Internet Engineering Task Force (IETF). No obstante, se pueden utilizar otros sistemas, tales como Expedited Forwarding of Differentiated Service o CR-LDP.

30 En el caso de la emisión de un vídeo que se está grabando actualmente, las características futuras de la corriente de bits que se está codificando se desconocen; en cambio, en un vídeo grabado dichas características sí se conocen. El hecho de que los sistemas de reserva permitan cambiar la cantidad de la tasa de bits reservada ofrece la posibilidad de decidir una política a seguir para reservar la cantidad de capacidad de red adecuada en cada momento, basándose en el conocimiento del material codificado. Una propuesta simple consiste en calcular la tasa de bits de pico (VBR, sin memoria tampón) y solicitarla para toda la duración de la transmisión, aunque debe tenerse en cuenta que esto significa desperdiciar la capacidad de la red, y por supuesto, que cuanto mayor es la capacidad solicitada mayor es la probabilidad que la red sea incapaz de facilitarla y, por consiguiente, que la petición de reserva sea rechazada. Otra propuesta simple, que reduce al mínimo la tasa de bits que se va a solicitar, consiste en calcular la tasa de bits media de toda la transmisión y solicitarla; no obstante, esto determina la necesidad de disponer de una memoria tampón de tamaño muy grande en el receptor y, todavía más importante (puesto que las grandes cantidades de almacenamiento actualmente son relativamente económicas), una gran LOS. En el documento de Furini y Towsley citado anteriormente, se da a conocer una modificación de la propuesta de tasa de pico. Este sistema comprende la determinación del punto de la secuencia de vídeo en el cual la tasa de pico alcanza un valor máximo, y la petición de esta tasa para el período de tiempo hasta ese instante. A continuación, se determina la tasa de pico máxima durante el resto de la secuencia, y se solicita esta tasa (inferior). El procedimiento continúa de la misma manera durante toda la secuencia. El documento también indica que puede aplicarse un grado de almacenamiento en memoria tampón, para reducir de ese modo las tasas de pico efectivas antes de aplicar el algoritmo de reserva. Aunque este sistema aumenta la eficacia de utilización de la red en comparación con el sistema de tasa de pico única, todavía se sigue desperdiciando mucha capacidad de red (es decir, que está reservada pero que no se utiliza), y evidentemente el beneficio es pequeño si la tasa de pico máxima se produce cerca del final de la secuencia. No obstante, un beneficio de este sistema es que la cantidad de capacidad de red solicitada desciende y, en particular, las peticiones de reserva nunca especifican una tasa de bits superior a la de las peticiones previas, reduciéndose de ese modo el riesgo de rechazo de la petición de reserva.

35 Según un aspecto de la presente invención, se proporciona un procedimiento para transmitir una secuencia digital de señales de vídeo que se han codificado mediante un algoritmo de compresión, de tal forma que el número de bits codificados por trama no es constante, que comprende las etapas siguientes:

- (a) dividir la secuencia en segmentos;
- (b) determinar una tasa de bits para cada segmento, no siendo dicha tasa de bits superior a la tasa de bits del segmento inmediatamente precedente y
- (c) transmitir las señales a las tasas de bits determinadas;

caracterizado porque la secuencia se divide en segmentos, de tal forma que el primer segmento es una parte situada al principio de la secuencia que contiene un número medio de bits codificados por trama (A_j) que es superior o igual al número medio de bits codificados por trama de cualquier parte posible empezando por el principio de la secuencia, y porque cada elemento consecutivo es una parte situada inmediatamente a continuación del segmento precedente que contiene un número medio de bits codificados por trama que es superior o igual al número medio de bits codificados por trama de cualquier parte posible situada inmediatamente a continuación del segmento precedente.

En otro aspecto, la presente invención proporciona un procedimiento para transmitir una secuencia digital de señales de vídeo que se han codificado mediante un algoritmo de compresión, de tal forma que el número de bits codificados por trama no es constante, en el que el vídeo de origen se ha codificado como una primera secuencia y una segunda secuencia que presentan respectivos índices de compresión diferentes, que comprende las etapas siguientes:

- (a) dividir por lo menos la primera o la segunda secuencia en segmentos, siendo el primer segmento la parte del principio de la secuencia que contiene un número medio de bits codificados por trama que es superior o igual al número medio de bits codificados por trama de cualquier parte posible empezando por el principio de la secuencia, y siendo cada segmento consecutivo una parte situada inmediatamente a continuación del segmento precedente que contiene un número medio de bits codificados por trama que es superior o igual al número medio de bits codificados por trama de cualquier parte posible situada inmediatamente a continuación del segmento precedente;
- (b) seleccionar un punto de cambio en un límite entre dos segmentos determinados en la etapa (a);
- (c) si la primera secuencia no se ha analizado en la etapa (a), analizar la primera secuencia para dividirla en segmentos;
- (d) determinar la tasa de bits para el segmento o cada segmento desde la primera secuencia hasta el punto de cambio;
- (e) transmitir la señal desde la primera secuencia hasta el punto de cambio a la tasa o las tasas de bits determinadas; y
- (f) transmitir, utilizando el procedimiento del primer aspecto, de una secuencia modificada que comprende la segunda secuencia desde el punto de cambio en adelante.

En las reivindicaciones subordinadas adjuntas, se exponen otros aspectos de la presente invención.

A continuación, se describen algunas formas de realización de la presente invención a título de ejemplo, haciendo referencia a los dibujos adjuntos, en los cuales:

las figuras 1A a 3C representan gráficamente los resultados de las pruebas realizadas;

la figura 4 es un diagrama de bloques de una forma de aparato para implementar la presente invención;

la figura 5 es un diagrama de flujo que ilustra el funcionamiento del aparato de la figura 4 y

las figuras 6 a 10 son unos gráficos que representan los resultados de otras pruebas.

Se considerará, en el receptor, cierto segmento de tiempo arbitrario (pero igual al número total de períodos de trama) comprendido desde el tiempo t_g , en el que el receptor empieza a decodificar la trama g , hasta el tiempo t_h , en el que el receptor empieza a decodificar la trama h . La duración de este segmento es $h - g$. Se supondrá además que la tasa de transmisión durante este segmento es A bits/período de trama.

Obviamente, en el tiempo t_g , el receptor deberá haber recibido los bits para todas las tramas hasta la trama g inclusive, es decir:

$$\sum_{j=0}^g d_j \text{ bits}$$

siendo d_j el número de bits codificados generados por el codificador para la trama j .

ES 2 314 259 T3

No obstante, a continuación se supondrá que el receptor también ha recibido, antes del tiempo g , p bits adicionales, es decir, en total:

$$\sum_{j=0}^g d_j + p \text{ bits.}$$

En cualquier momento de tiempo $t_k (t_g \leq t_k \leq t_h)$, en el que el receptor empiece a decodificar la trama k , el receptor habrá recibido también $(k - g)A$ bits y entonces:

$$\text{Número total de bits recibidos en el tiempo } t_k = \sum_{j=0}^g d_j + p + (k - g)A.$$

En este momento, el receptor necesita disponer de todos los bits para todas las tramas hasta la trama k inclusive, es decir:

$$\text{Número total de bits necesarios en el tiempo } t_k = \sum_{j=0}^k d_j.$$

Puesto que el número de bits recibidos debe ser por lo menos igual al número necesario, la condición que debe satisfacerse para evitar el desbordamiento de la memoria también es:

$$\sum_{j=0}^g d_j + p + (k - g)A \geq \sum_{j=0}^k d_j$$

O

$$p + (k - g)A \geq \sum_{j=g+1}^k d_j$$

Si este resultado debe alcanzarse sin la transmisión de bits de precarga p , entonces debe cumplirse:

$$(k - g)A \geq \sum_{j=g+1}^k d_j$$

O

$$A \geq \frac{1}{(k - g)} \sum_{j=g+1}^k d_j.$$

Por lo tanto, la tasa transmitida A debe ser superior o igual al número medio de bits por trama generados durante las tramas $g+1$ a k , para cualquier valor de $k (g + 1 \leq k \leq h)$, condición que se cumple si

$$A \geq \max_{k=g+1}^h \left\{ \frac{1}{(k - g)} \sum_{j=g+1}^k d_j \right\}.$$

La utilización de esta tasa significa que el número de bits $(h-g)A$ transmitidos durante el segmento sobrepasará el número de bits generados para el segmento, a menos que se alcance el valor máximo para $k = h$, es decir, al final del segmento. Partiendo de la base de que la utilización continuada de la tasa de transmisión calculada de este modo, una vez que se ha pasado el valor máximo, parece suponer la utilización de una velocidad superior a la absolutamente necesaria, la primera versión de la presente invención descrita a continuación tiene como objetivo dividir los datos en segmentos, de tal manera que los valores máximos se produzcan siempre al final del segmento.

El primer procedimiento que se va a describir es para la transmisión de material de vídeo almacenado que ya se ha codificado mediante un algoritmo de compresión, tal como el MPEG, a través de una red de paquetes, tal como Internet. Se presupone que la red dispone de medios para reservar capacidad de tasa de bits. Lo que se pretende es determinar la tasa de bits que se va a utilizar como una función del tiempo, para obtener de ese modo:

- una baja latencia de inicio;
- una baja tasa de bits transmitidos y
- una alta eficacia de transmisión (es decir, un menor desperdicio),

aunque, puesto que los requisitos anteriores son opuestos, cualquier solución deberá ser necesariamente una solución de compromiso.

En este ejemplo, se supone que no existe ninguna restricción sobre la tasa de bits que se puede elegir y que la tasa de bits utilizada para la transmisión y la tasa de bits reservada en la red son iguales.

Esta primera versión también está sujeta a una restricción, en la medida en que la tasa de bits solicitada no puede incrementarse nunca, es decir, es una función del tiempo que decrece monotónicamente. Como se ha indicado anteriormente, esto es deseable para reducir el riesgo de rechazo de la reserva.

Puesto que el hardware de almacenamiento de gran tamaño no resulta problemático para los usuarios actuales, en esta solución, la reducción del tamaño de la memoria tampón necesaria en el decodificador no constituye un tema de interés principal. No obstante, de hecho, el tamaño de memoria tampón necesario que resulta de este procedimiento también se reduce sumamente en comparación con el tamaño obtenido utilizando la tasa de bits media para realizar una transmisión de vídeo VBR. Aun en el peor de los casos, que raramente se produce en la práctica, el tamaño de memoria tampón necesario no será superior al que se requiere cuando se transmite una secuencia de vídeo VBR a la tasa de bits media.

El algoritmo siguiente determina la función de transmisión ("FOT") que se debe utilizar.

Se supone que existen N tramas en la secuencia de vídeo y que el número de bits codificados para cada trama es $d_0, d_1, \dots, d_{N-1}$, respectivamente.

Como se ha indicado anteriormente, el algoritmo está limitado en la medida en que la función de transmisión nunca puede incrementarse, sino solo decrecer.

Conceptualmente, se pueden producir cambios en la tasa de transmisión en cualquier intervalo de tramas de la FOT. En la práctica, puede imponerse un límite a la frecuencia a la que se puede cambiar la tasa, dependiendo de las restricciones del sistema de reserva particular que se está utilizando. No obstante, con una FOT que decrece monotónicamente, un retardo en el cambio de la tasa no da por resultado una pérdida de calidad (aunque supone desperdiciar capacidad de la red), ya que su único efecto es la reserva de más capacidad de la que realmente se necesita. La primera etapa del algoritmo tiene por finalidad hallar cuántos "escalones" presentará la FOT y cuando se producirá cada escalón.

En primer lugar, se define:

$$A_i = \frac{\sum_{j=0}^i d_j}{(i+1)}$$

que representa la tasa de bits media de la secuencia de vídeo desde el inicio hasta la trama i inclusive. A continuación, se calculan $A_0, A_1, \dots, A_{N-1}$, y a partir de éstas se calcula el valor de i que presenta la A_i más elevada. Si se supone que este valor es k_0 , el borde del primer "escalón" se establece al final de la trama k_0 , lo cual significa que, hasta el final de la trama k_0 , la FOT necesita su tasa de transmisión más alta.

Una vez determinado el primer "escalón", la trama (k_0+1) se considera la "primera" trama para las tramas siguientes, y se calcula $A_{i-1}^{(1)}$ para $i = k_0+1, k_0+2, \dots, N-1$. La fórmula que se aplica es:

$$A_i^{(1)} = \frac{\sum_{j=k_0+1}^i d_j}{(i-k_0)}$$

o en el caso general,

$$A_i^{(q)} = \frac{\sum_{j=k_{q-1}+1}^i d_j}{(i - k_{q-1})}$$

Asimismo en este caso, se elige el valor más elevado como borde del segundo “escalón”, al final de la trama k_1 , siendo k_1 el correspondiente valor de i . El procedimiento anterior se repite hasta que se llega al borde del último “escalón” en la trama $N-1$. En general, esto da por resultado M valores k_m , $m = 0, \dots, M-1$ (siendo k_{M-1} siempre igual a $N-1$), que se puede considerar que dividen la secuencia de vídeo en $M-1$ segmentos, de los cuales el segmento 0 comprende las tramas 0 a k_0 , y cada uno de los otros m segmentos comprende las tramas $k_{m-1} + 1$ a k_m .

El propósito del segundo paso del algoritmo es el de elegir una tasa de transmisión adecuada para el “nivel” de cada “escalón”. Teóricamente, en este caso, la tasa más baja que permite asegurar que se transmitan todos los bits necesarios antes de que finalice cada “escalón”, aunque se carezca de bits de precarga, es la media de las tasas de bits de las tramas que componen el segmento. Una tasa inferior requerirá necesariamente disponer de bits de precarga y, en consecuencia, una LOS superior, mientras que las tasas más altas, podrán provocar un desperdicio de la capacidad de la red. Asimismo, es de prever que las tasas más altas incrementen el riesgo de que la reserva de recursos resulte infructuosa.

Existen M segmentos, siendo $m = 0, 1, \dots, M-1$. Asimismo se establece lo siguiente:

S_i es la suma de los bits generados en el segmento i , es decir $\sum_{q=k_{i-1}+1}^{k_i} d_q$;

R_i es la tasa de transmisión de la FOT en el segmento i (obsérvese que $K_0=k_0+1$) y

K_i es el número de tramas del segmento i , es decir, $k_i - k_{i-1}$.

En este caso, las tasas necesarias son simplemente las tasas medias $R_i = S_i / K_i$; $i = 1, 2, \dots, M-1$.

Este procedimiento también puede utilizarse para calcular la tasa R_0 para el segmento 0, si se establece $k_{-1} = -1$.

Debe observarse que, en la codificación de vídeo MPEG, la primera trama siempre es una trama I que genera más bits que las tramas P o B . Por consiguiente, a menudo, los resultados ponen de manifiesto que el primer segmento solo comprende una trama y que la tasa de transmisión R_0 es muy inferior a la tasa de transmisión R_1 . Puesto que los usuarios pueden esperar fácilmente a que transcurra un intervalo de algunas tramas para obtener una mejor oportunidad para realizar con éxito una reserva de recursos, es preferible establecer la igualdad $R_0 = R_1$.

En la tercera etapa, una vez que se ha determinado la FOT completa, puede calcularse el tamaño de memoria tampón que se necesita en el decodificador.

A continuación, se describe una segunda versión modificada, en la que la selección de las tasas está sujeta a una limitación. Por ejemplo, la limitación puede consistir en la necesidad de que la tasa sea de un número entero de bits por trama o, más generalmente, que sea una tasa de un grupo de tasas discretas. En el análisis, se utilizan los operadores de cuantificación definidos a continuación:

$Q^+(X)$ significa la tasa más baja permitida que es superior o igual a X (también denominada tasa de “techo”) y

$Q^-(X)$ significa la tasa más alta permitida que es inferior o igual a X (también denominada tasa de “suelo”).

A continuación, se describen dos opciones:

- (a) redondeo hacia arriba hasta la tasa de techo: en este caso, la tasa utilizada puede resultar más alta de lo que es estrictamente necesario para un segmento particular, hecho que puede brindar la oportunidad de utilizar una tasa inferior para el segmento siguiente y
- (b) redondeo hacia abajo hasta la tasa de suelo: en este caso, la tasa utilizada puede resultar más baja de lo que es necesario para un segmento particular, hecho que determina la necesidad de utilizar una tasa más alta para el segmento precedente.

Se considerará en primer lugar la opción de techo y se definirá el valor de techo de la “altura” del primer “escalón” de la FOT original como la “altura” del primer “escalón” delimitado en la nueva FOT. Debe observarse que, de esta forma, después del primer “escalón”, se habrán transmitido más bits al receptor que la suma de bits de las tramas que

pertenecen al “primer” escalón. Por lo tanto, cuando se delimita el segundo “escalón”, debe excluirse el número de bits que pertenecen al “escalón” siguiente pero que se han transmitido en el “escalón” o los “escalones” anteriores, y recalcularse la tasa media del segundo “escalón”. Si el valor de techo de la nueva tasa de bits media no es inferior al valor de techo de la tasa media del antiguo tercer “escalón”, éste se define como la “altura” del segundo “escalón” delimitado. En caso contrario, se define el valor de techo de la tasa de bits media del antiguo tercer “escalón” como la “altura” del segundo “escalón” delimitado. Se sigue este procedimiento hasta que se fija la “altura” del último “escalón” delimitado. Puesto que siempre se adopta el valor de techo de cada “escalón”, es posible que la transmisión de la corriente de vídeo VBR se realice con algunos intervalos de trama menos que la duración de la secuencia de vídeo. Con la simulación de la transmisión basada en la nueva FOT, el tiempo de duración de la FOT puede especificarse con exactitud. Una vez que se ha realizado la transmisión de la corriente de vídeo VBR, los recursos de red reservados pueden liberarse de inmediato. Por lo tanto, la utilización del 100% del ancho de banda sigue estando asegurada. Con la “altura” del primer “escalón” delimitado, la LOS puede recalcularse con precisión. Por último, a través de la simulación del procedimiento de transmisión, también puede fijarse el tamaño de memoria tampón necesario para prevenir el desbordamiento de esta.

El procedimiento adoptado tiene lugar de la forma indicada a continuación. La división en segmentos se realiza de la forma indicada anteriormente.

Aparte de las cantidades S_i , R_i y K_i definidas anteriormente, también se introduce R_i^1 , un valor temporal para la tasa de transmisión en el segmento i .

- I. Cálculo de todas las tasas medias $R_i^1 = S_i / K_i$, $i = 0, 1, \dots, M-1$,
- II. Establecimiento de la tasa para el segmento 0 como $R_0 = Q^+(R_0^1)$. (Debe tenerse en cuenta que si, como se ha indicado anteriormente, es deseable utilizar una tasa inferior para el primer segmento, en lugar de empezar por el segmento 0 se puede empezar por el segmento 1),
- III. Establecimiento de la tasa para el segmento 1 restando, antes de la cuantificación, los bits adicionales enviados durante el segmento anterior:

$$R_1 = Q^+ \{ R_1^1 - (R_0 - R_0^1) \}$$

$$\hat{R}_1 = Q^+ \{ R_1^1 \}$$

seleccionando la mayor de las dos.

- IV. Para los segmentos anteriores $i = 2, \dots, M-1$:

$$R_i = Q^+ \{ R_i^1 - (R_{i-1} - R_{i-1}^1) \}$$

$$\hat{R}_i = Q^+ \{ R_{i+1}^1 \}$$

seleccionando la mayor de las dos. Como es obvio, la segunda alternativa no se aplica para $i = M-1$.

En la tercera versión que se describe a continuación, se utilizan las tasas de “suelo”. En este caso, el procesamiento debe realizarse en orden inverso, empezando por el último “escalón”. Esto es necesario para que los bits que no pueden transmitirse en un segmento particular puedan pre-transmitirse en los segmentos previos. El procedimiento específico primero define el valor de suelo de la tasa de bits media del último “escalón” como la nueva tasa de transmisión del último “escalón” delimitado de la nueva FOT. Entonces, puede determinarse el número de bits que se necesitan para el último “escalón” delimitado pero que no se pueden transmitir. El “escalón” o los “escalones” previos deben asegurar que dicho número de bits adicionales se transmita antes de que se inicie el nuevo último “escalón” de la FOT. Por lo tanto, cuando se delimita el penúltimo “escalón”, debe procurarse que este contenga los bits que necesita, más el número adicional de bits necesarios para el último “escalón”. Por consiguiente, es necesario recalcular una nueva tasa de bits media para el penúltimo “escalón”. Si el valor de suelo de la nueva tasa de bits media del penúltimo “escalón” no es superior al valor de suelo de la tasa media del antepenúltimo “escalón” de la FOT original, éste se define simplemente como la “altura” del nuevo penúltimo “escalón”. En caso contrario, se define el valor de suelo de la tasa de bits media del antiguo antepenúltimo “escalón” como la “altura” del nuevo penúltimo “escalón”. Siguiendo este procedimiento hasta el primer “escalón”, la delimitación se completa y se obtiene la FOT delimitada. Como en el caso del valor de techo, con el número de bits precargados y la “altura” del primer “escalón” delimitado, puede recalcularse con precisión la LOS. Por último, mediante la simulación del procedimiento de transmisión, también puede fijarse el tamaño de memoria tampón necesario para prevenir el desbordamiento de ésta.

ES 2 314 259 T3

Como en el caso anterior, existen M segmentos $m = 0, 1, \dots, M-1$. Asimismo, se establece lo siguiente:

S_i es la suma de los bits generados en el segmento i , es decir $\sum_{q=k_{i-1}+1}^{k_i} d_q$;

R_i es la tasa de transmisión de la FOT en el segmento i ;

K_i es el número de tramas del segmento i , es decir, $k_i - k_{i-1}$ y

R_i^1 es una tasa de transmisión temporal que se supone para el segmento i .

I. Cálculo de todas las tasas medias $R_i^1 = S_i / K_i$, $i = 0, 1, \dots, M-1$,

II. Establecimiento de la tasa de transmisión R_{M-1} para el segmento $M-1$ igual al valor de suelo de la tasa media para este segmento. Es decir,

$$R_{M-1} = Q^- \{ R_{M-1}^1 \}$$

II. Cálculo del número de bits precargados, P_{M-1} que es necesario que estén presentes en la memoria tampón del receptor al principio del segmento $M-1$ para prevenir el subdesbordamiento en el segmento $M-1$.

$$P_{M-1} = (R_{M-1}^1 - R_{M-1}) * K_{M-1}$$

III. A continuación, la tasa para el siguiente segmento puede calcularse como

$$R_{M-2} = Q^- \{ R_{M-1}^1 + P_{M-1} \}$$

$$\acute{o} = Q^- \{ R_{M-3}^1 \}$$

seleccionando la menor de las dos, con

$$P_{M-2} = (R_{M-2}^1 - R_{M-2}) * K_{M-2}$$

IV. Este procedimiento se repite utilizando la fórmula general, $m = M-3, \dots, 0$:

$$R_m = Q^- \{ R_{m+1}^1 + P_{m+1} \}$$

$$\acute{o} = Q^- \{ R_{m-1}^1 \}$$

seleccionando la menor de las dos.

Y

$$P_m = R_{m+1}^1 - R_m$$

ES 2 314 259 T3

Nuevamente, si se desea, esta iteración puede detenerse cuando $m = 1$ y cuando se utiliza R_1 para el segmento 0.

Este procedimiento da por resultado un valor de P_0 que es una precarga para el primer segmento y que es necesario transmitir en primer lugar. En realidad, es conveniente definir una precarga b_0 que comprenda todos los bits que se transmiten antes de que el receptor empiece a decodificar la primera trama en $t=0$.

Suponiendo que R_0 se calcule como se ha indicado anteriormente, entonces

$$b_0 = P_0 + R_0$$

No obstante, si se utiliza la tasa R_1 para el segmento 0, entonces sólo pueden transmitirse $(K_0 - 1)R_1$ bits entre $t=0$ y el final del segmento, siendo entonces la precarga total:

$$P_0 + K_0 R_0 - (K_0 - 1) R_1$$

Si se supone que se utiliza R_1 , la latencia de inicio (LOS) es b_0 / R_1 .

A continuación, se hace referencia a la cuestión del tamaño de la memoria tampón. Indudablemente, con la presente FOT, se puede obtener una tasa de transmisión y una LOS razonable. La eficacia de transmisión de la red puede ser casi del 100% y requiere un tamaño de memoria tampón menor que el que se necesitaría utilizando directamente un ancho de banda de tasa media fijo. No obstante, en algunas situaciones, dicho tamaño todavía es mucho más grande de lo que se necesita debido a que se reserva el ancho de banda de la tasa de pico. En el sistema de reserva del ancho de banda de la tasa de pico, basta con que el tamaño de la memoria tampón del decodificador sea tan grande como el número de bits utilizados en la trama más complicada. No obstante, en el presente sistema, se necesita una memoria tampón de un tamaño superior a ese. Aunque en comparación con una tasa de bits media constante el presente sistema puede conseguir un tamaño de memoria tampón mucho menor en la mayoría de situaciones, debe admitirse que, en la peor situación, el tamaño de memoria tampón necesario para el presente sistema se aproxima al tamaño de memoria tampón necesario para una tasa de bits media constante. Dicha situación se produce cuando la A; más elevada aparece en las últimas tramas de la secuencia de vídeo. En tal situación, la presente curva de “descenso escalonado” prácticamente presenta un solo “escalón”. Por lo tanto, no resulta suficientemente efectivo reducir al mínimo el tamaño de la memoria tampón a través de cambios “escalonados”. A pesar de todo, dicha situación apenas se produce, debido a que cuanto más tarde aparecen los “bits de pico”, menos es el efecto sobre A; y, a menos que al final de la secuencia aparezcan de manera anormal algunas tramas excepcionalmente complicadas, no se producirá nunca. Sea cual sea la situación, la LOS nunca supondrá ningún problema en el presente sistema. Actualmente, se considera que disponer de hardware con una memoria un poco grande no debería ser problemático para el usuario. Lo que debería preocupar más a los usuarios es obtener una baja LOS y una buena eficacia de transmisión de la red.

Además, aun cuando los usuarios no puedan costearse la memoria tampón de gran tamaño que el presente sistema requiere, se puede hallar una solución de compromiso entre la eficacia de transmisión y el tamaño de memoria tampón necesaria en el decodificador. Con dicha solución de compromiso, el tamaño de memoria tampón necesario puede reducirse todavía más según las necesidades de los usuarios.

Aunque la descripción del algoritmo actual solo se basa en los bits por trama como unidad básica, resulta evidente que la unidad puede definirse como un GOP o un determinado número de imágenes o paquetes juntos. Independientemente de la unidad que se desee definir en este algoritmo, el principio es general y debe ser común.

A continuación, se describen algunos ejemplos de secuencias de codificación de vídeo de prueba, mediante el procedimiento de “suelo”. En cada caso, se indican los valores de la función de transmisión $f(t)$ o (R_i) , el valor de b_0 y una tasa de transmisión propuesta para b_0 , (a) cuando se utiliza el algoritmo anterior, (b) cuando se utiliza el procedimiento de Furini y Towsley y (c) cuando se utiliza una única tasa de bits media.

Ejemplo 1

JacknBox

- (a) Se codifica una secuencia de prueba (denominada JacknBox) de formato intermedio común (CIF) y 140 tramas de duración con un cuantificador fijo de tamaño de etapa 16 de conformidad con la norma H.263+, y se calcula la función FOT con el presente algoritmo.

ES 2 314 259 T3

$f(t) =$

5	5100	$0 < t \leq T_{48};$
	3645	$T_{48} < t \leq T_{51};$
	3058	$T_{51} < t \leq T_{52};$
	2830	$T_{52} < t \leq T_{61};$
10	2682	$T_{61} < t \leq T_{70};$
	2651	$T_{70} < t \leq T_{71};$
	2464	$T_{71} < t \leq T_{90};$
15	2447	$T_{90} < t \leq T_{108};$
	2321	$T_{108} < t .$

20 En este documento, se define T_i como el tiempo en el que el decodificador presenta la trama i . Se define la unidad de medida de todas las tasas de medición de este documento como los bits por intervalo de trama $b_0 = 39824$ bits, La tasa de transmisión propuesta para b_0 es de 5100 bits por intervalo de trama.

(b) Utilizando el procedimiento de Furini y Towsley, se obtiene

25

$f(t) =$

	9896	$T_0 < t \leq T_{29};$
30	9432	$T_{29} < t \leq T_{40};$
	7272	$T_{40} < t \leq T_{41};$
	6552	$T_{41} < t \leq T_{46};$
35	6184	$T_{46} < t \leq T_{47};$
	5328	$T_{47} < t \leq T_{48};$
	3696	$T_{48} < t \leq T_{51};$
40	3632	$T_{51} < t \leq T_{106};$
	3552	$T_{106} < t \leq T_{138};$
45	2896	$T_{138} < t .$

$$b_0 = 39824 \text{ bits}$$

En este sistema de transmisión, el valor b_0 se obtiene mediante 39824 por intervalo de trama.

50 (c) Con una tasa de bits media constante, la función será:

$$f(t) = 3669$$

$$b_0 = 108488 \text{ bits};$$

55

b_0 se alcanza mediante 3669 bits por intervalo de trama.

La figura 1 representa estos resultados en un gráfico.

60 Los resultados del análisis se indican en la tabla 1:

65

ES 2 314 259 T3

TABLA 1

JacknBox de 140 tramas según H.263+

Sistemas	(c) Canal de ancho de banda fijo con tasa de bits media	(a) Presente sistema	(b) Sistema de Furini y Towsley
Utilización de ancho de banda (%)	100	100	63,46
Tasa de reserva inicial (bits por intervalo de trama)	3,669	5,100	39,824
LOS (tramas)	108488/3669=29,57	39824/5100=7,8	39824/39824=1
Tamaño de memoria tampón (bits)	108,488	60,336	39,824

Con el presente sistema, también se ha codificado la misma secuencia de vídeo con un control de tasa para CBR. En este caso, la LOS sería de $29656/3735 = 7,94$ tramas. No obstante, con el control de tasa para CBR normal se saltan 10 tramas, y el cálculo de bits que se proporciona es el mismo que el número medio de bits de la codificación VBR.

Ejemplo 2

Programa de TV de 8400 tramas según H.263+

En esta prueba, se utiliza una secuencia QCIF (quarter-CIF) de programa de TV normal con 8400 tramas, codificada con un tamaño de paso de cuantificador fijo de 16 de conformidad con la norma H.263+. El tipo de imagen es IPPPP... con actualización forzada cada 132 tramas según la recomendación H.263+.

(a)

$f(t) =$

$$\begin{aligned}
 &4977 & T_0 < t \leq T_{3173} ; \\
 &4218 & T_{3173} < t \leq T_{3679} ; \\
 &3968 & T_{3679} < t \leq T_{3680} ; \\
 &3848 & T_{3680} < t \leq T_{3681} ; \\
 &3844 & T_{3681} < t \leq T_{4752} ; \\
 &3090 & T_{4752} < t \leq T_{8392} ; \\
 &992 & T_{8392} < t \leq T_{8393} ; \\
 &816 & T_{8393} < t \leq T_{8394} ; \\
 &644 & T_{8394} < t \leq T_{8396} ; \\
 &544 & T_{8396} < t \leq T_{8397} ; \\
 &384 & t > T_{8397} ;
 \end{aligned}$$

Como antes, b_0 puede obtenerse mediante la primera tasa de 4977 bits por intervalo de trama.

ES 2 314 259 T3

(b)

$f(x) =$

27672	$T_0 < t \leq T_{8339} ;$	21952	$T_{8358} < t \leq T_{8359} ;$
26704	$T_{8339} < t \leq T_{8340} ;$	21744	$T_{8359} < t \leq T_{8369} ;$
26560	$T_{8340} < t \leq T_{8341} ;$	20448	$T_{8369} < t \leq T_{8373} ;$
26488	$T_{8341} < t \leq T_{8342} ;$	20344	$T_{8373} < t \leq T_{8384} ;$
26240	$T_{8342} < t \leq T_{8344} ;$	19960	$T_{8384} < t \leq T_{8385} ;$
25832	$T_{8344} < t \leq T_{8345} ;$	19016	$T_{8385} < t \leq T_{8391} ;$
25136	$T_{8345} < t \leq T_{8346} ;$	11656	$T_{8391} < t \leq T_{8392} ;$
24168	$T_{8346} < t \leq T_{8347} ;$	992	$T_{8392} < t \leq T_{8393} ;$
23816	$T_{8347} < t \leq T_{8352} ;$	816	$T_{8393} < t \leq T_{8394} ;$
23760	$T_{8352} < t \leq T_{8353} ;$	648	$T_{8394} < t \leq T_{8396} ;$
23616	$T_{8353} < t \leq T_{8356} ;$	544	$T_{8396} < t \leq T_{8397} ;$
22824	$T_{8356} < t \leq T_{8357} ;$	384	$T_{8397} < t \leq T_{8399} ;$
22528	$T_{8357} < t \leq T_{8358} ;$		

$b_0 = 13944$ bits

b_0 puede transmitirse a 29762 bits por intervalo de trama.

(c) Con una tasa de bits media constante, la FOT será:

$f(t) = 3966$

$b_0 = 3348584$ bits;

b_0 puede establecerse a 3669 bits por intervalo de trama.

La figura 2 representa las curvas de la FOT para un programa de TV de 8400 tramas con codificación H.263+.

Los resultados del análisis se indican en la tabla 2:

TABLA 2

8400 tramas, codificación H.263+

Sistemas	(c) Canal de ancho de banda fijo con tasa de bits media	(a) Presente sistema	(b) Sistema de Furini y Towsley
Utilización de ancho de banda (%)	100	100	14,36
Tasa de reserva de inicio (bits por intervalo de trama)	3.966	4.977	27.672
LOS (intervalo de trama)	$3348584/3966=844,322$	$13944/4977=2,8$	$13944/27672=0,5$
Tamaño de memoria tampón (bits)	6.116.362	3.908.218	27.672

ES 2 314 259 T3

Ejemplo 3

Secuencia QCIF de programa de TV de 8400 tramas codificada con MPEG4

5 Se codifica la misma secuencia QCIF de programa de TV de 8400 tramas mediante codificación MPEG4, con un tamaño de paso de cuantificador fijo de 10. El tipo de imagen es IBBPBBPBBPBB (N=12, M=3). Debe observarse que, con las imágenes B, la secuencia de codificación de las imágenes es diferente de la secuencia de presentación de las imágenes. Por consiguiente, las imágenes I o P relacionadas deben transmitirse antes de la imagen B. Se necesita cierta cantidad de preprocesamiento antes de utilizar el presente algoritmo.

10

(a) Por último, la presente FOT es:

	$f(t) =$	
15	7426	$T_0 < t \leq T_{4750} ;$
	6938	$T_{4750} < t \leq T_{4786} ;$
	66470	$T_{4786} < t \leq T_{4798} ;$
20	6309	$T_{4798} < t \leq T_{4870} ;$
	6190	$T_{4870} < t \leq T_{4900} ;$
	6083	$T_{4900} < t \leq T_{4918} ;$
25	6026	$T_{4918} < t \leq T_{8398} ;$
	168	$T_{8398} < t.$

30

$b_0 = 16548$ bits;

b_0 puede establecerse mediante 7426 bits por intervalo de trama.

35

(b)

	$f(x) =$	
40	57472	$T_0 < t \leq T_{8338} ;$
	50616	$T_{8338} < t \leq T_{8350} ;$
	49504	$T_{8350} < t \leq T_{8368} ;$
45	48608	$T_{8368} < t \leq T_{8371} ;$
	48536	$T_{8371} < t \leq T_{8383} ;$
	44968	$T_{8383} < t \leq T_{8386} ;$
50	31752	$T_{8386} < t \leq T_{8389} ;$
	28696	$T_{8389} < t \leq T_{8398} ;$
	168	$T_{8398} < t.$

55

$b_0 = 16040$ bits;

b_0 puede establecerse a 57472 bits por intervalo de trama.

60

(c) Con una tasa de bits media constante, la FOT será:

$f(x) = 6825.$

65

$b_0 = 2874758$ bits;

b_0 puede establecerse a 6825 bits por intervalo de trama.

ES 2 314 259 T3

La figura 3 representa las curvas de la FOT para un programa de TV de 8400 tramas con codificación MPEG4 (N=12, M=3).

Los resultados del análisis se indican en la tabla 3:

TABLA 3

8400 tramas, codificación MPEG4

Sistemas	(c) Canal de ancho de banda fijo con tasa de bits media	(a) Presente sistema	(b) Sistema de Furini y Towsley
Utilización de ancho de banda (%)	100	100	11,897
Tasa de reserva de inicio (bits por intervalo de trama)	6.825	7.426	57.472
LOS (tramas)	2874758/6825=421,21	16548/7426=2,228	16040/57472=0,279
Tamaño de memoria tampón (bits)	6.236.252	3.997.072	57.472

A partir de los resultados experimentales anteriores, puede deducirse que la LOS se ha reducido en gran medida, mientras se sigue manteniendo el 100% de eficacia. Debe tenerse en cuenta, asimismo, que no se desperdicia ningún recurso de red. Lo único que todavía se puede mejorar es el tamaño de la memoria tampón necesario para el decodificador, que podría reducirse todavía más.

La figura 4 es un diagrama de bloques de un servidor operativo de conformidad con la presente invención. El servidor contiene los componentes de ordenador habituales, es decir, un procesador 10, una memoria 11, una memoria de disco 12, un teclado 13, una pantalla 14 y una interfaz de red 15 para la conexión con una red de telecomunicaciones 16. Las secuencias de vídeo disponibles para transmitir se almacenan en la memoria de disco 12 de una manera convencional, en forma de archivos codificados 20.

En la memoria de disco 12, también está almacenado un programa informático 21 para implementar el control de la operación del servidor. A continuación, se describe el funcionamiento de este programa, en el que se utiliza el procedimiento de "suelo", con referencia al diagrama de flujo representado en la figura 5.

Etapas 100

Desde un terminal remoto, se recibe una petición para la transmisión de una secuencia de vídeo deseada, por medio de la interfaz 15. Dicha petición comprende el nombre de archivo de uno de los archivos 20 que contiene la secuencia.

Etapas 101

El procesador 10 lee el archivo en cuestión en la memoria de disco 12 y determina el número de bits codificados d_j del archivo para cada una de las N tramas de la secuencia almacenada, y almacena el valor de N y d_j ($j=0, \dots, N-1$) en la memoria 11.

Etapas 102

El procesador calcula $k_0, \dots, k_{M-1}$ de la forma descrita anteriormente y almacena M y $k_0, \dots, k_{M-1}$ en la memoria 11.

Etapas 103: cálculo de $R^1 i =$ para todos los valores de i

Etapas 104: establecimiento de $R_{M-1} = Q - \{R^1_{M-1}\}$ y cálculo de P_{M-1}

Etapas 106: establecimiento de un puntero $m = M-2$

Etapas 107: cálculo de R_m y P_m

Etapas 109: decremento de m. Si $m \geq 0$, regreso a etapas 107

Etapas 111: cálculo de $b_0 = P_0 + R_0$

ES 2 314 259 T3

Etapla 112: cálculo de las duraciones de los segmentos - en esta implementación, la precarga y el segmento 0 se consideran un único segmento a efectos de transmisión. Por lo tanto

$$\tau_0 = (b_0/R_0 + k_0 + 1) * \tau$$

$$\tau_i = (k_i - k_{i-1}) * \tau$$

$$i = 1 \dots M-1$$

siendo τ la longitud de un período de trama.

Etapla 113: establecimiento de i en 0

Etapla 114: transmisión de una petición de reserva que especifica una tasa de R_i y una duración de por lo menos τ_i

Etapla 115: transmisión de segmento i a una tasa R_i (precedido, cuando $i = 0$, por P_0 bits de precarga)

Etapla 116: si se han transmitido todos los segmentos, parada; en caso contrario, incremento de i en la etapa 117 y retroceso a la etapa 114.

En algunos sistemas de reserva, tales como el sistema RSVP mencionado anteriormente, es necesario que el terminal de recepción formule una petición de reserva para adaptarse a la multidifusión. En tal caso, la etapa 113 se modifica para especificar la transmisión de un mensaje al terminal de recepción, en el que se indican R_i y τ_i , y a continuación el terminal transmite la petición de reserva deseada a la red.

En algunas redes, los tiempos permitidos para cambiar la tasa reservada pueden estar sujetos a ciertas restricciones. No obstante, el sistema adoptado anteriormente es resistente ante dichos problemas, debido a que todas las peticiones de reserva, excepto la primera, solicitan una tasa inferior a la anterior. En consecuencia, el retardo en el procesamiento de dichas peticiones determina que la tasa reservada permanezca alta una vez que la tasa de transmisión se ha reducido. En este caso, la eficacia de utilización de la red desciende, pero la calidad de la transmisión no cambia.

Una segunda forma de realización de la presente invención analiza la posibilidad de cambiar la tasa de vídeo. En este caso, se generan dos (o más) corrientes de vídeo con diferentes calidades de imagen y, por lo tanto, diferentes tasas de datos. Habitualmente, estas secuencias pueden generarse utilizando niveles de detalle de cuantificación diferentes; es decir, en la corriente de baja calidad y baja tasa de datos, se utiliza un cuantificador de ajuste grueso y, en una secuencia de calidad más alta y tasa de datos más alta, se utiliza un cuantificador de ajuste menos grueso. La posibilidad de cambiar la tasa de vídeo resulta particularmente interesante en el presente contexto, en el que tal vez se produzca un fallo de reserva de tasa al principio de una transmisión, y en el que dicha situación puede remediarse enviando en primer lugar una corriente de calidad relativamente escasa y, a continuación, cambiando a una corriente de calidad superior cuando la naturaleza de la señal o las condiciones de la red lo permiten. No obstante, el sistema que se va a describir también es útil cuando se realiza el cambio de la tasa de vídeo por algún otro motivo.

Cuando se utiliza la codificación intertrama, el cambio entre dos corrientes diferentes puede ocasionar una alteración importante de la imagen, debido al mal seguimiento de los predictores en el codificador y el decodificador. No obstante, este cambio puede realizarse sin dicha degradación de la calidad de la imagen, generando de vez en cuando tramas codificadas de transición que esencialmente codifican la diferencia entre la trama de la corriente de destino del cambio y la trama de la corriente de origen del cambio. Por lo tanto, la transmisión de las tramas de una primera corriente va seguida de una o más tramas de transición y, a continuación, de las tramas de la segunda corriente. La generación de dichas tramas de transición no es nueva; por este motivo, no se describe en mayor detalle. Puede hallarse una descripción de dicho sistema en la solicitud de patente internacional WO 98/26604 (y la correspondiente patente US nº 6.002.440). Otro sistema de este tipo, en el que se utilizan las denominadas "tramas SP", se describe en el documento de Marta Karczewicz y Ragip Kurceren titulado "A Proposal for SP-frames", VCEG-L-27, ITU-T Video Coding Experts Group Meeting, Eibsee, Alemania, 9-12 de enero de 2001, y el documento de Ragip Kurceren y Marta Karczewicz titulado "SP-frame demonstrations", VCEG-N42, ITU-T Video Coding Experts Group Meeting, Santa Barbara, CA, USA, 24-27 de septiembre de 2001.

En el contexto de la propuesta de la "FOT" descrita anteriormente, la cuestión del cambio entre dos corrientes plantea ciertas dudas que es necesario aclarar. Si en un punto de tiempo determinado se produce un cambio desde una primera corriente hasta una segunda corriente, entonces la memoria tampón del decodificador generalmente contiene tramas de la primera corriente que no son útiles para decodificar la segunda corriente. Por consiguiente, suponiendo que el decodificador vaya a iniciar de inmediato la decodificación de la segunda corriente, estas tramas no se utilizarán y supondrán un desperdicio de la capacidad de transmisión. Peor todavía, las tramas necesarias para decodificar la segunda corriente no estarán presentes en la memoria tampón. Teóricamente, esto se puede resolver recalculando la FOT de la segunda corriente, considerando que el principio de la parte de la segunda corriente que realmente se va a transmitir es el inicio de la corriente, aunque en la práctica esto puede dar por resultado un requisito de tasa de datos transmitidos extremadamente alta si se desea evitar la interrupción de las imágenes presentadas.

El problema del desperdicio de bits puede evitarse dejando que el decodificador continúe decodificando las tramas de la primera corriente que permanecen en la memoria tampón, mientras que la memoria tampón puede acumular durante este período algunas de las tramas que se necesitan para decodificar la segunda corriente (es decir, la trama o las tramas de transición y las tramas de la segunda corriente). A pesar de todo, el peligro de que el requisito de la tasa de bits transmitidos sea excesivo no desaparece.

Idealmente, el cambio de corriente de bits deberá producirse tan pronto como se disponga del ancho de banda. No obstante, debido a los problemas que se acaban de describir, esto no resulta práctico. Asimismo, si es necesario generar tramas de transición (que normalmente no se generan en cada trama, sino solo en puntos seleccionados), entonces deberá planificarse de antemano en qué puntos (los puntos de cambio) deben situarse dichas tramas.

Basándose en dichas cuestiones, se considera primero la posibilidad de cambiar en tiempos que coinciden con el “borde” de un “escalón” de la FOT. Una característica de este sistema es que, en el “borde” de cada “escalón”, la memoria tampón del receptor no almacena ningún bit, puesto que todos los bits transmitidos se han decodificado como imágenes. Por lo tanto, si se desea cambiar en el “borde” de la corriente original, se deberán extraer todos los bits transmitidos de la memoria tampón del receptor y de ese modo no se desperdiciará ningún bit debido al cambio de la corriente de bits.

Aunque con el establecimiento de los puntos de cambio en los “bordes de escalón” de la corriente de bits original se puede impedir que no se desperdicien bits transmitidos, todavía surgirá un problema si el punto de cambio de la nueva corriente no se halla en un “borde de escalón”. Esto se debe a que si el punto de cambio no está situado en un “borde de escalón” en la nueva corriente, tal vez sea necesario transmitir algunos bits preacumulados para la nueva corriente dentro de un espacio de tiempo muy corto para reproducir vídeo de manera continua en el receptor. Esto puede provocar una petición de reserva de tasa muy superior, tal vez superior incluso a la tasa de reserva que implica la nueva corriente. Si el punto de cambio de la nueva corriente de bits se halla en el centro de un “escalón”, la falta de bits acumulados provoca una reserva de tasa elevada. Por lo tanto, idealmente, el punto de cambio de la nueva corriente de vídeo deberá hallarse también en los “bordes de escalón”.

Según los análisis anteriores, puede parecer que la única oportunidad de disponer de los puntos de cambio óptimos para las dos corrientes se tiene cuando estas presentan los mismos “puntos de borde”. De no ser así, o bien se desperdician bits, o bien se necesita una tasa de bits muy alta tras el cambio de corriente de bits. Afortunadamente, en posteriores investigaciones se ha comprobado que en realidad las curvas de FOT generadas a partir de cuantificadores diferentes presentan “bordes de escalón” situados en posiciones similares, aunque puede que no sean absolutamente iguales. Esto se debe a que, en una secuencia de vídeo, las imágenes complejas deben ocupar más bits que las normales, independientemente del cuantificador que se seleccione.

Este resultado se ha obtenido a través de algunos experimentos, en los que se ha seleccionado una secuencia de vídeo 140 CIF JacknBox.

En el primer experimento, se desea aclarar si unas corrientes de vídeo diferentes basadas en la misma secuencia de vídeo acercan o no sus “bordes de escalón” en las FOT. En la figura 6, se representa la similitud de las curvas de FOT basadas en diferentes cuantificadores. Las curvas corresponden a los tamaños de paso de cuantificador de 2, 3, 4, 10, 16 y 31 y se marcan como Q2, Q3, etc. Puede observarse que cuando se incrementa el tamaño de paso del cuantificador, la FOT va adquiriendo una forma cada vez más plana. No obstante, dichas corrientes siguen presentando los “bordes de escalón” casi al mismo tiempo. Además, debe observarse que, aunque similares, los puntos de “escalón” de FOT diferentes no son exactamente iguales. Las figuras 7 y 8 dan a conocer más detalles de curvas FOT diferentes en los “bordes de escalón”. Aunque dichos puntos no son exactamente iguales, resulta poco perjudicial que el cambio de corriente de bits tenga lugar en un lugar aproximado. El siguiente experimento puede aclarar todavía más esta cuestión.

En el segundo experimento, se supone que se va a cambiar de la corriente de bits (corriente Q16) generada con un cuantificador fijo 16 a una segunda corriente de bits (corriente Q8) generada con un cuantificador fijo 8 en cada intervalo de trama. En la figura 9, se representan las curvas de reserva de cuando se cambia respectivamente la corriente de bits en las tramas 35, 42, 45, 49, 50 y 52. En la figura 10, se representa el número de bits desperdiciados cuando la corriente de bits se cambia en diferentes intervalos de trama. Las figuras 9 y 10 deberán ilustrar por sí solas la diferencia entre cambiar las corrientes en el punto de “borde” y cambiarlas en otros puntos. En la figura 9, puede observarse que, si los puntos de cambio están alejados de los “bordes de escalón”, la tasa de transmisión necesaria es todavía más elevada que la tasa de transmisión necesaria originalmente para la corriente Q8. Esto es exactamente lo que sucedía en los análisis anteriores. En esta situación, es necesario conseguir la acumulación de bits necesaria en un plazo de tiempo corto para obtener una presentación adecuada tras el cambio de corriente de bits. Por lo tanto, la tasa de transmisión necesaria puede llegar a ser muy alta, resultando poco realista finalizar dicho cambio de corriente de bits. Por otro lado, si la corriente de bits se cambia cerca de los puntos de “borde”, no se requiere una tasa de transmisión muy alta para lograr la acumulación de bits necesaria, porque cada “escalón” de la FOT es independiente. En la figura 10, también se puede observar que cambiar la corriente de bits cerca de los puntos de “borde” resulta más aconsejable. En la curva de la FOT, siempre es necesario preacumular algunos bits para las tramas siguientes. Si se aplica el cambio de corriente de bits, los bits preacumulados para la corriente original no tendrán ninguna utilidad, sino que se desperdiciarán.

ES 2 314 259 T3

En la figura 10, es fácil observar que cuando se cambia la corriente de bits solo en los “bordes de escalón” no se puede desperdiciar ningún bit. Cuanto más cerca se está de los “bordes de escalón”, menos bits se desperdician. Tanto a partir de la figura 9 como de la figura 10, se verifica que los mejores puntos de cambio de la FOT son sus “bordes de escalón”.

En cuanto a la cuestión del punto preciso que se debe elegir como punto de cambio para cambiar de una primera a una segunda corriente en la práctica, si los escalones de las dos corrientes coinciden, entonces, por supuesto, no existe ninguna ambigüedad. No obstante, si existe alguna diferencia en los tiempos, pueden emprenderse las siguientes acciones:

- a) elegir un escalón de la primera corriente (fácil de implementar);
- b) elegir un escalón de la segunda corriente (igualmente fácil de implementar);
- c) elegir el primero de los dos escalones (reduciendo al mínimo el desperdicio de bits);
- d) elegir el último de los dos escalones (evitando cualquier incremento del ancho de banda de reserva para la segunda corriente).

No obstante, en la práctica tiene poca relevancia qué opción se selecciona, puesto que las diferencias de rendimiento entre estas son bastante pequeñas. En realidad, si el punto de cambio elegido está desplazado algunas tramas respecto del “escalón”, a menudo se puede obtener un rendimiento satisfactorio.

En vista de lo anterior, el procedimiento propuesto se realiza de la forma indicada a continuación (suponiendo que se elige la opción (a) indicada):

- i) cálculo de la FOT para la primera corriente;
- ii) elección de un punto de cambio que coincide con un escalón de esta FOT;
- iii) generación de una trama de transición;
- iv) cálculo de la FOT para la trama de transición más el resto de la segunda corriente;
- v) transmisión de la primera corriente hasta el punto de cambio y
- vi) transmisión de la trama de transición más el resto de la segunda corriente.

En caso de que se utilice la opción (b), (c) o (d), la etapa i) comprenderá el cálculo de la FOT de la segunda corriente también, y la etapa (ii) comprenderá la selección según la opción elegida. A pesar de todo, todavía es necesario recalcular la FOT para la segunda corriente en la etapa 4. Debe observarse también que el (re)cálculo de la etapa (iv) tendrá en cuenta automáticamente todas las correcciones necesarias debidas a la no coincidencia del punto de cambio con la etapa calculada originalmente para la segunda corriente o debidas a la utilización de las tasas de “techo” o “suelo” descritas anteriormente.

Evidentemente, es posible elegir más de un punto de cambio si así se desea; por ejemplo, para volver a cambiar a la primera corriente o para cambiar a una tercera corriente.

Aunque las cuestiones concernientes al cambio se han descrito dentro del contexto de los sistemas que están sujetos a la limitación de presentar una FOT que decrece monotónicamente, la presente propuesta de cambio también puede utilizarse cuando no se aplica esta limitación. Asimismo, esta propuesta también es útil cuando se realiza el cambio desde una corriente de alta calidad hasta una corriente de baja calidad.

REIVINDICACIONES

5 1. Procedimiento para transmitir una secuencia digital de señales de vídeo que se han codificado mediante un algoritmo de compresión, de tal forma que el número de bits codificados por trama no es constante, que comprende las etapas siguientes:

- (a) dividir (102) la secuencia en segmentos;
- 10 (b) determinar (107) una tasa de bits (R_m) para cada segmento, no siendo dicha tasa de bits superior a la tasa de bits del segmento inmediatamente precedente y
- (c) transmitir (115) las señales a las tasas de bits determinadas;

15 **caracterizado** porque la secuencia se divide en segmentos, de tal forma que el primer segmento es una parte situada al principio de la secuencia que contiene un número medio de bits codificados por trama (A_i) que es superior o igual al número medio de bits codificados por trama de cualquier parte posible empezando por el principio de la secuencia, y en el que cada segmento consecutivo es una parte situada inmediatamente a continuación del segmento precedente que contiene un número medio de bits codificados por trama que es superior o igual al número
20 medio de bits codificados por trama de cualquier parte posible situada inmediatamente a continuación del segmento precedente.

25 2. Procedimiento para transmitir una secuencia digital de señales de vídeo que se han codificado mediante un algoritmo de compresión, de tal forma que el número de bits codificados por trama no es constante, en el que el vídeo de origen se ha codificado como una primera secuencia y una segunda secuencia que presentan respectivos índices de compresión diferentes, que comprende las etapas siguientes:

- (a) dividir por lo menos la primera y la segunda secuencias en segmentos, siendo el primer segmento la parte del principio de la secuencia que contiene un número medio de bits codificados por trama que es superior o
30 igual al número medio de bits codificados por trama de cualquier parte posible empezando por el principio de la secuencia, y siendo cada segmento consecutivo una parte situada inmediatamente a continuación del segmento precedente que contiene un número medio de bits codificados por trama que es superior o igual al número medio de bits codificados por trama de cualquier parte posible situada inmediatamente a continuación del segmento precedente;
- 35 (b) seleccionar un punto de cambio en un límite entre dos segmentos determinados en la etapa (a);
- (c) si la primera secuencia no se ha analizado en la etapa (a), analizar la primera secuencia para dividirla en segmentos;
- 40 (d) determinar la tasa de bits para el segmento o cada segmento desde la primera secuencia hasta el punto de cambio;
- (e) transmitir la señal desde la primera secuencia hasta el punto de cambio a la tasa o las tasas de bits determinadas; y
- 45 (f) transmitir, mediante el procedimiento según la reivindicación 1, de una secuencia modificada que comprende la segunda secuencia desde el punto de cambio en adelante.

50 3. Procedimiento según la reivindicación 2, en el que, en la etapa (b) se selecciona un punto de cambio que se halla en el límite entre dos segmentos de la primera secuencia.

4. Procedimiento según la reivindicación 2, en el que, en la etapa (b), se selecciona un punto de cambio que se halla en el límite entre dos segmentos de la segunda secuencia.

55 5. Procedimiento según la reivindicación 2, en el que, en la etapa (a), tanto la primera como la segunda secuencia se dividen en segmentos, y en la etapa (b), se selecciona un punto de cambio que se halla en el límite entre dos segmentos tanto de la primera como de la segunda secuencias o, en caso de que los límites no coincidan, en el primero de los dos límites.

60 6. Procedimiento según la reivindicación 2, en el que, en la etapa (a), tanto la primera como la segunda secuencia se dividen en segmentos, y en la etapa (b), se selecciona un punto de cambio que se halla en el límite entre dos segmentos tanto de la primera como de la segunda secuencias o, en caso de que los límites no coincidan, en el último de los dos límites.

65 7. Procedimiento según cualquiera de las reivindicaciones 2 a 6, en el que se selecciona un punto de cambio que se encuentra a cuatro tramas del límite pertinente.

ES 2 314 259 T3

8. Procedimiento según la reivindicación 7, en el que se selecciona un punto de cambio que coincide con el límite pertinente.

5 9. Procedimiento según cualquiera de las reivindicaciones 2 a 8, en el que la primera secuencia se codifica a un índice de compresión superior al de la segunda secuencia.

10 10. Procedimiento según la reivindicación 9, en el que la primera secuencia se codifica utilizando una cuantificación más gruesa que la segunda secuencia.

11. Procedimiento según cualquiera de las reivindicaciones 2 a 10, en el que las secuencias se codifican mediante la codificación intertrama, y que comprende la generación en el punto de cambio de una secuencia de transición que comprende o empieza con una trama de la segunda secuencia codificada utilizando una trama decodificada de la primera secuencia como predictor, y en el que la secuencia modificada comprende la secuencia de transición seguida de tramas de la segunda secuencia.

12. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la tasa de bits determinada para cada uno de por lo menos los últimos segmentos de la secuencia o una de las secuencias es un número de bits por período de trama igual al número medio de bits codificados por trama para dicho segmento.

13. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la tasa de bits determinada para cada uno de por lo menos los últimos segmentos de la secuencia o una de las secuencias es un número de bits por período de trama igual al mayor de los siguientes valores:

(i) la más baja de un conjunto de tasas de bits permitidas que es mayor o igual a una tasa nominal para el segmento, siendo dicha tasa nominal el número medio de bits codificados por trama para dicho segmento, menos cualquier reducción permitida como consecuencia de una tasa de bits determinada para la secuencia precedente que sobrepasa la tasa nominal para dicho segmento precedente; y

(ii) la más baja de un conjunto de tasas de bits permitidas que es mayor o igual al número medio de bits codificados por trama para el segmento siguiente.

14. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la tasa de bits determinada para cada uno de por lo menos los últimos segmentos de la secuencia o una de las secuencias es un número de bits por período de trama igual al menor de los siguientes valores:

(i) la más alta de un conjunto de tasas de bits permitidas que es menor o igual a una tasa nominal para el segmento, siendo dicha tasa nominal el número medio de bits codificados por trama para el segmento más cualquier incremento necesario como consecuencia de la tasa de bits determinada para la secuencia siguiente que es inferior a la tasa nominal para dicho segmento precedente y

(ii) la más alta del conjunto de tasas de bits permitidas que es menor o igual al número medio de bits codificados por trama para el segmento precedente.

15. Procedimiento según cualquiera de las reivindicaciones anteriores, que comprende la transmisión, a una red de telecomunicaciones, de mandatos que solicitan la reserva de dichas tasas de bits determinadas.

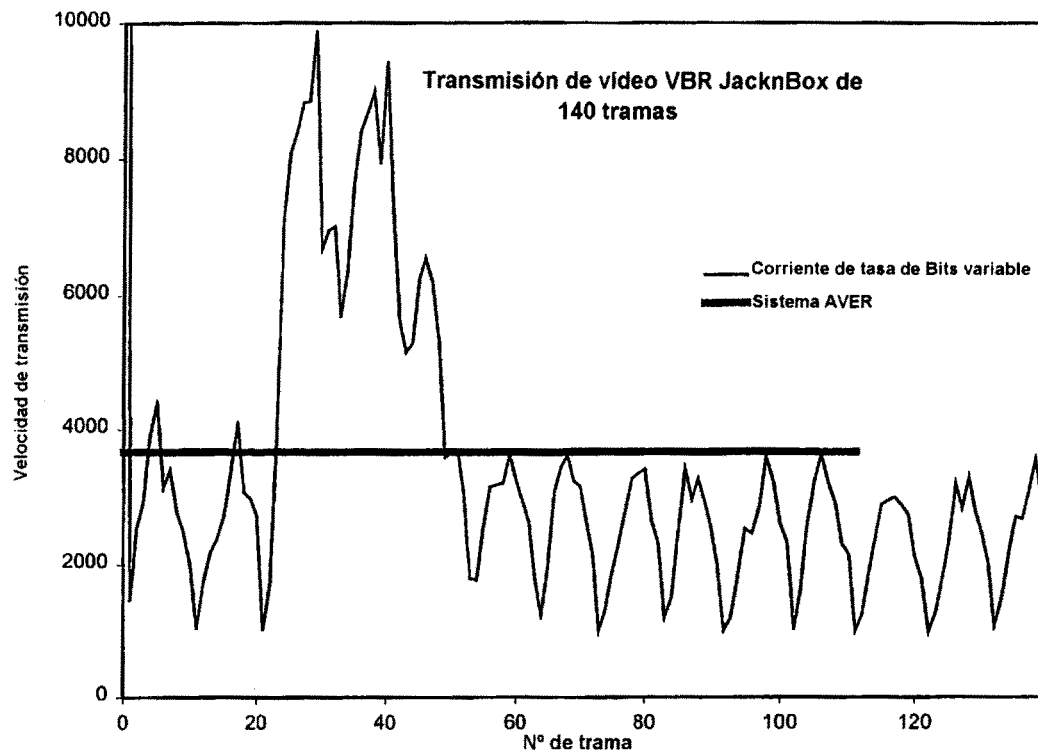


Figura 1A

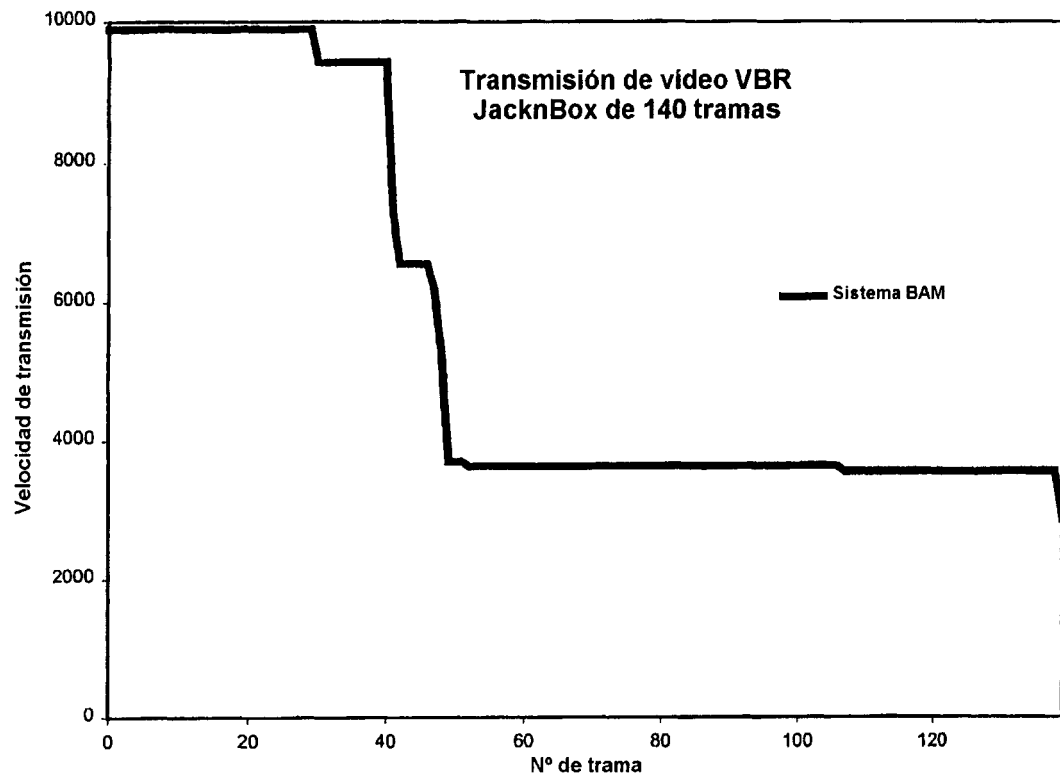


Figura 1B

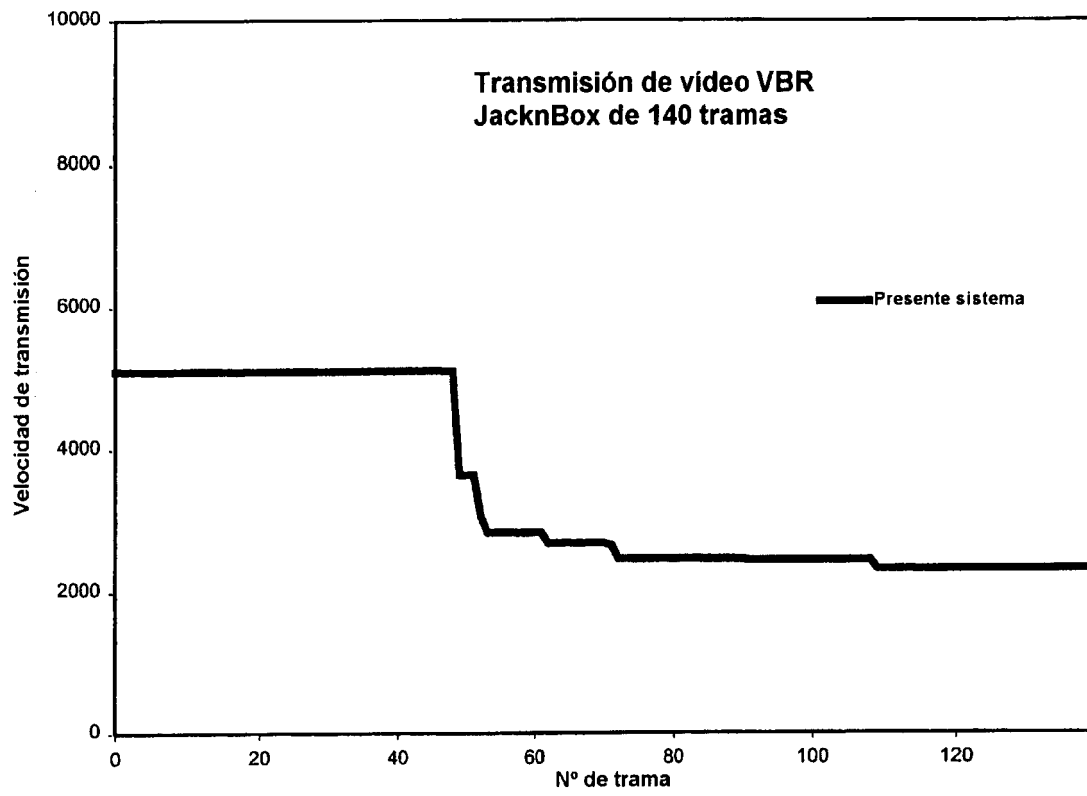


Figura 1C

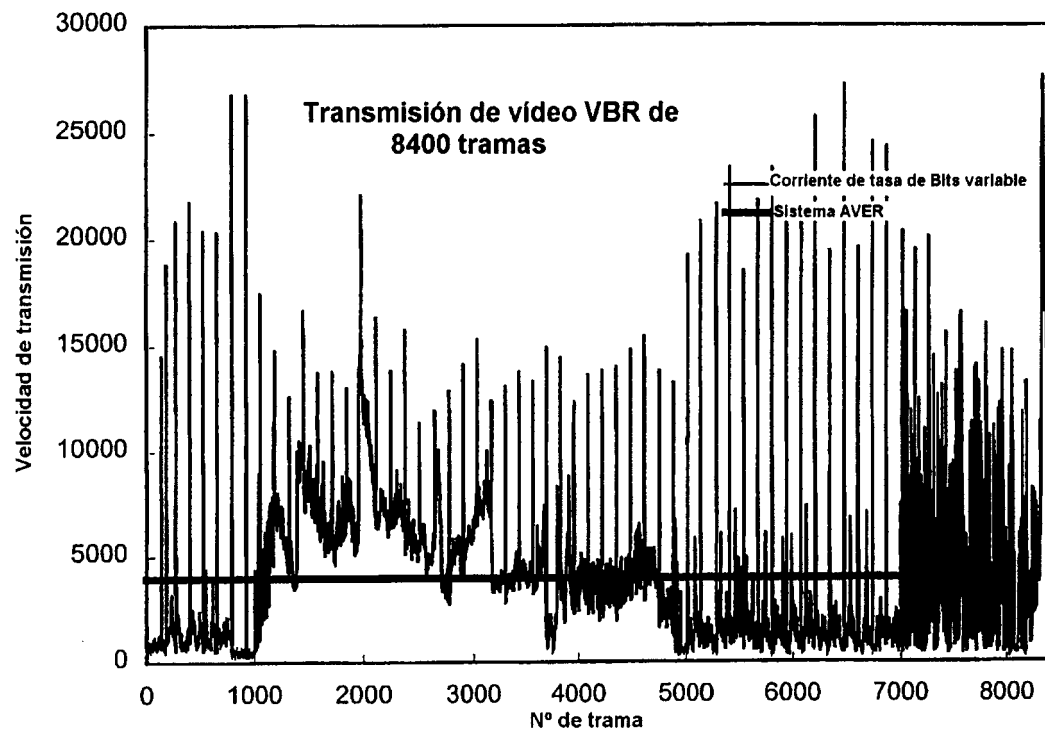


Figura 2A

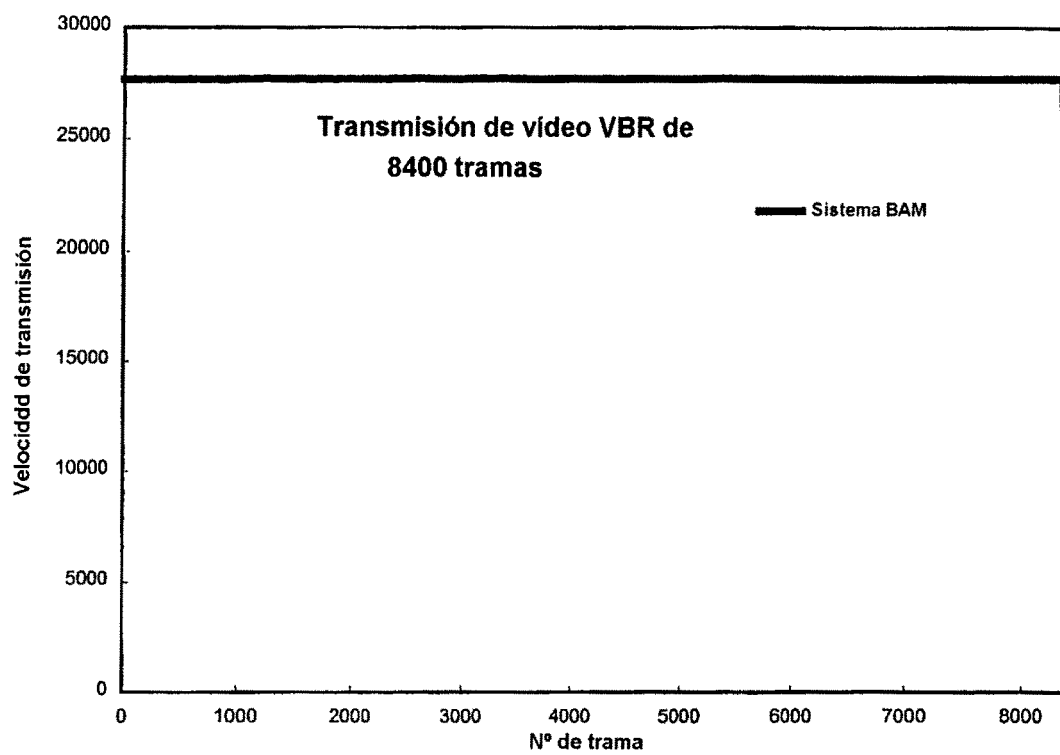


Figura 2B

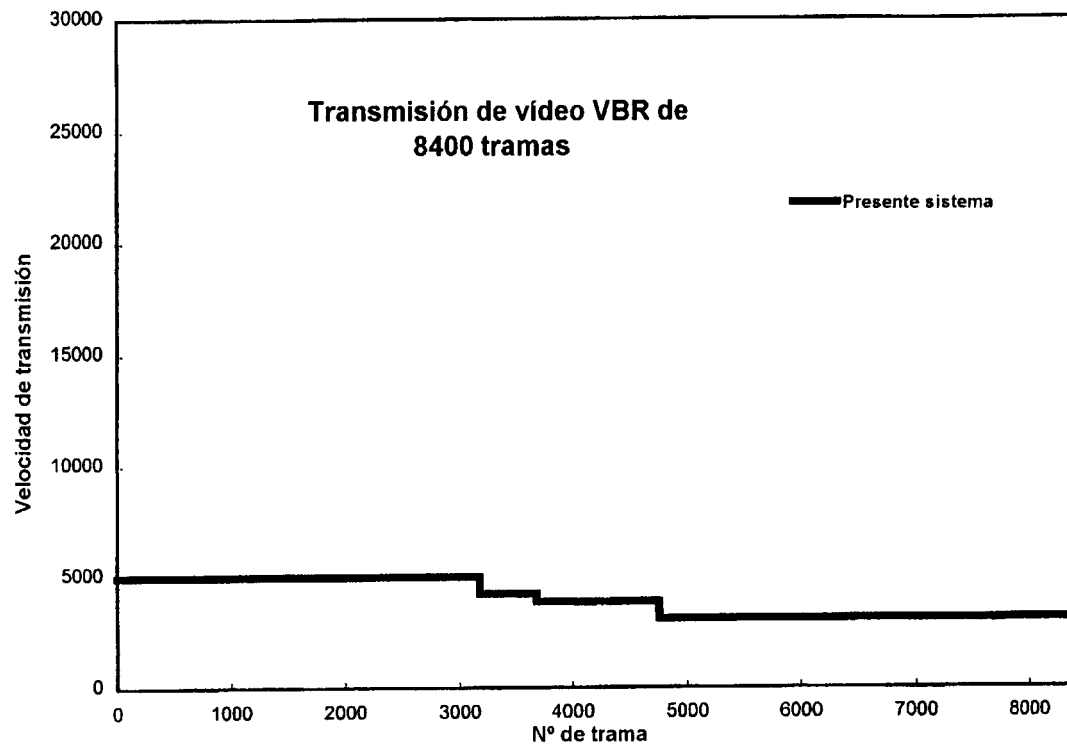


Figura 2C

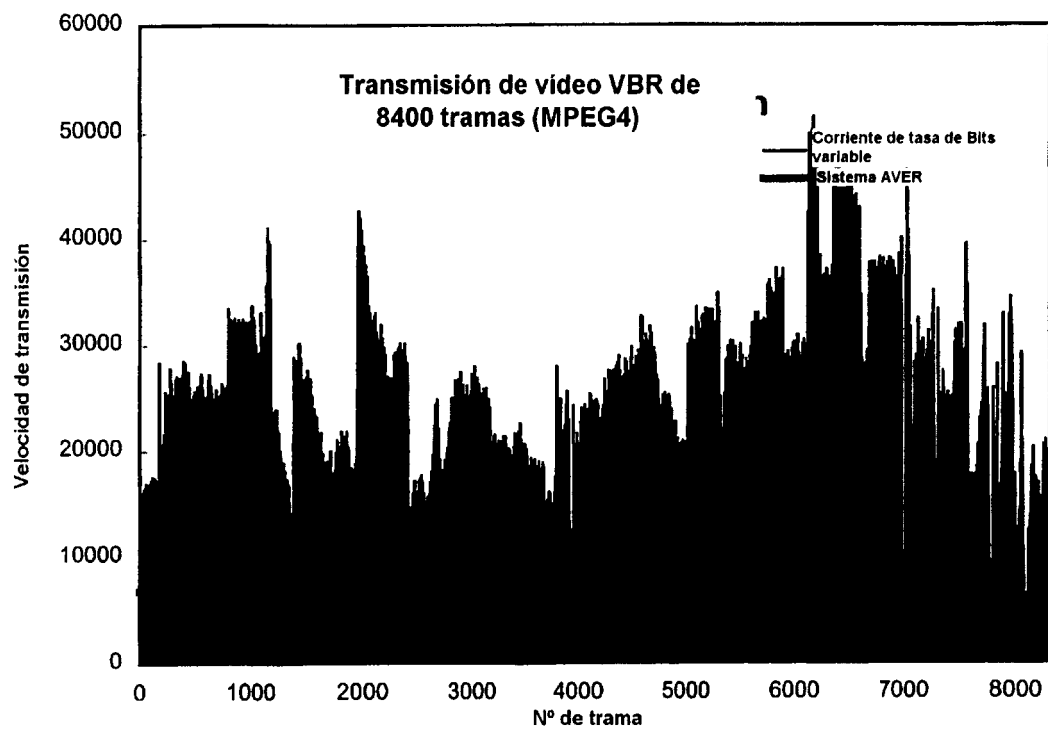


Figura 3A

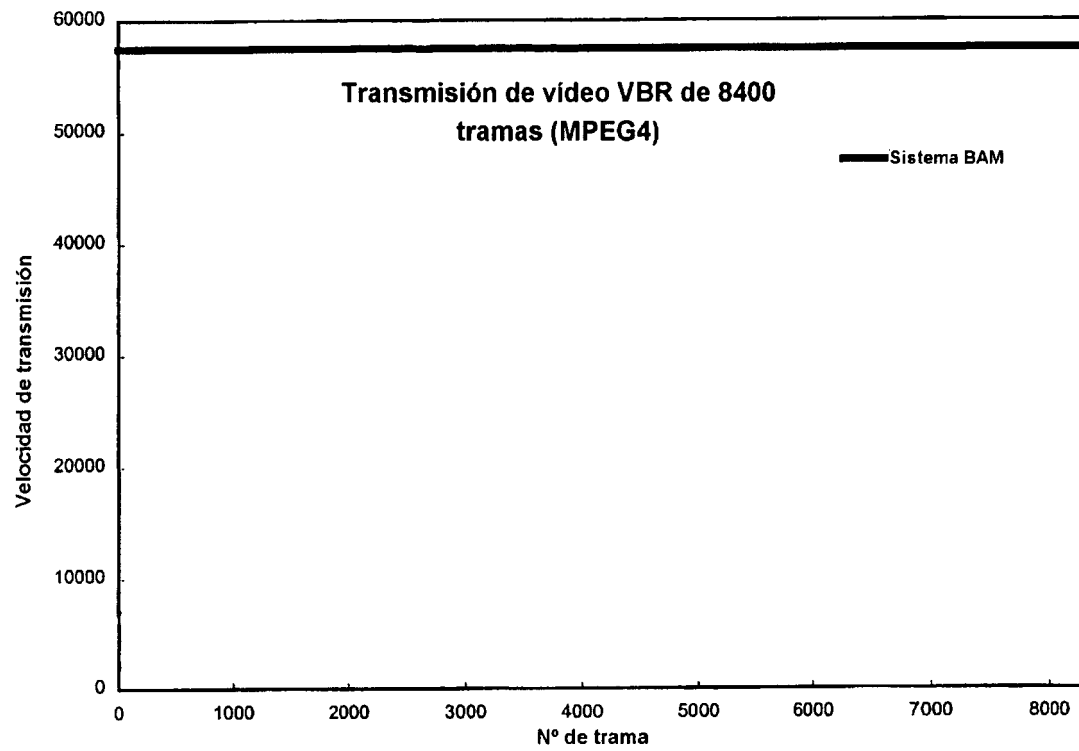


Figura 3B

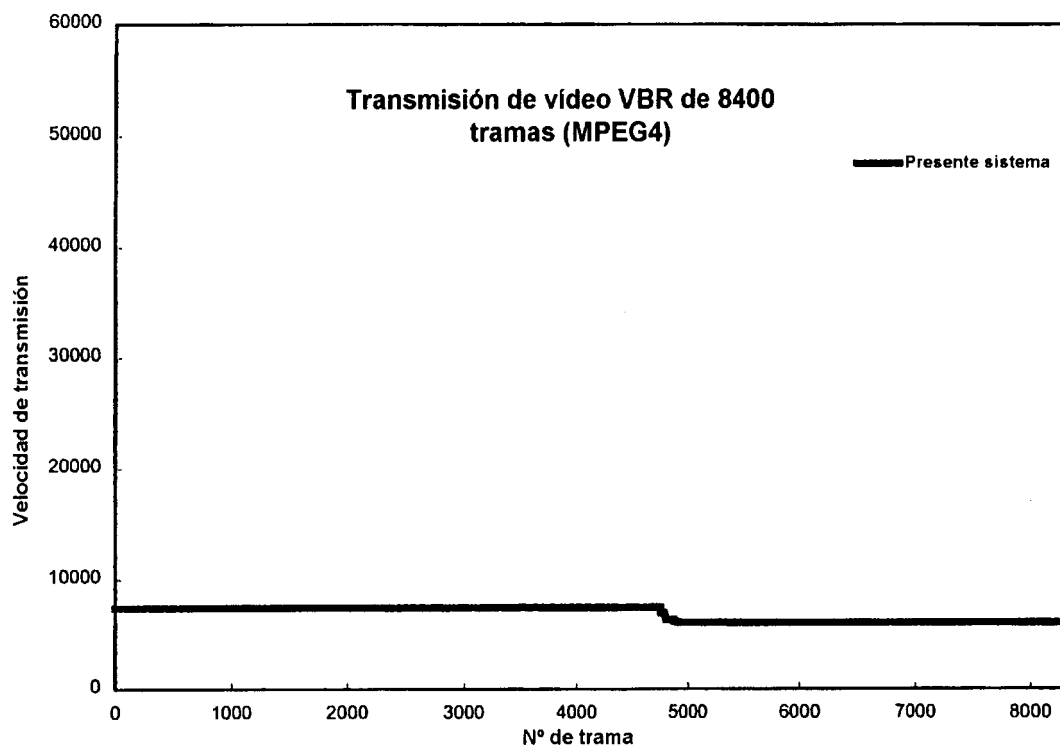


Figura 3C

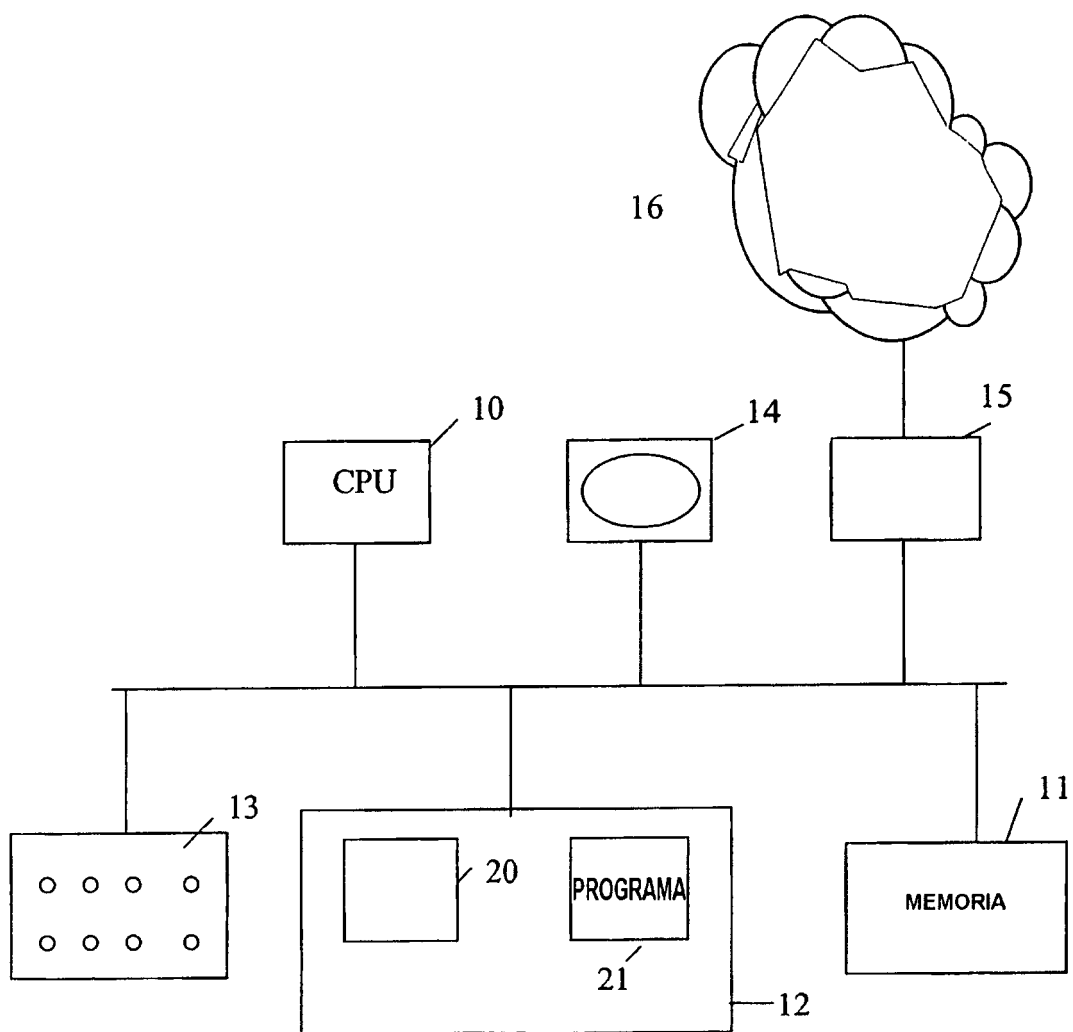


Figura 4

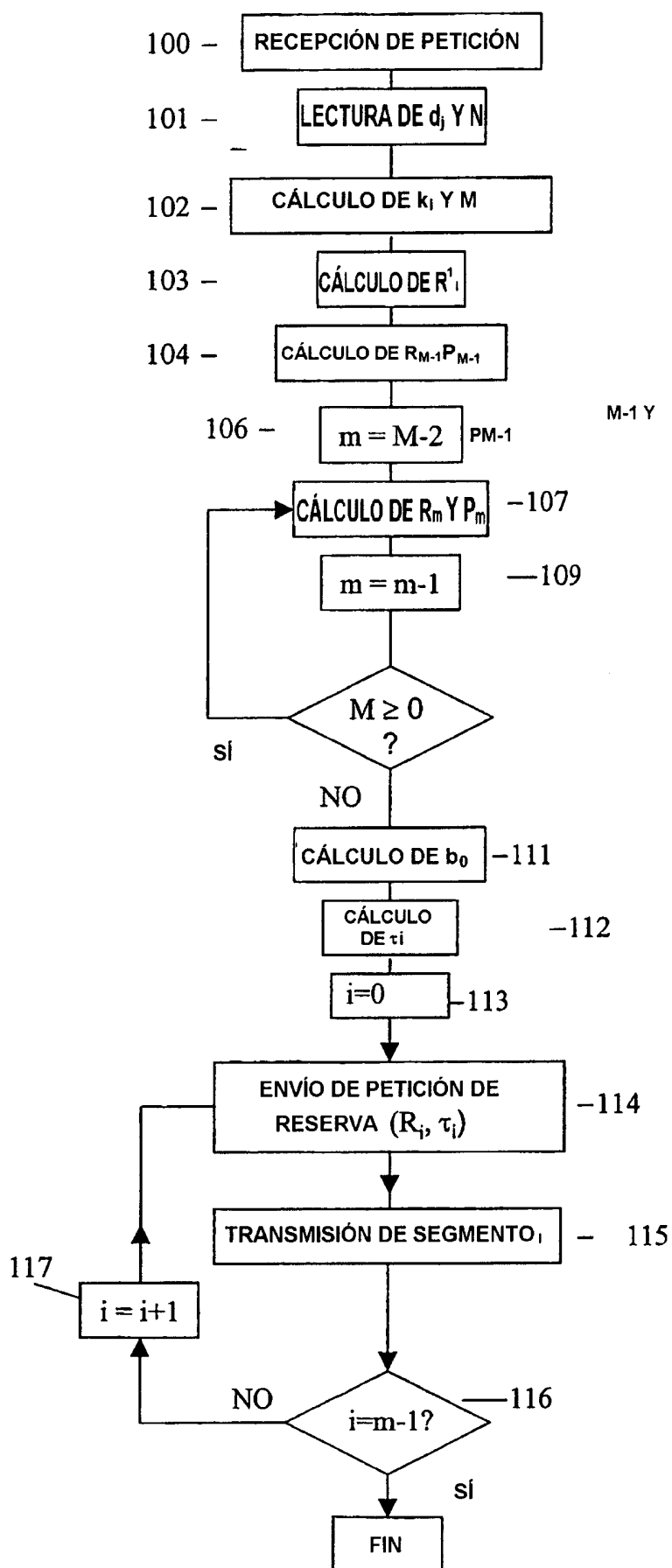


Figura 5

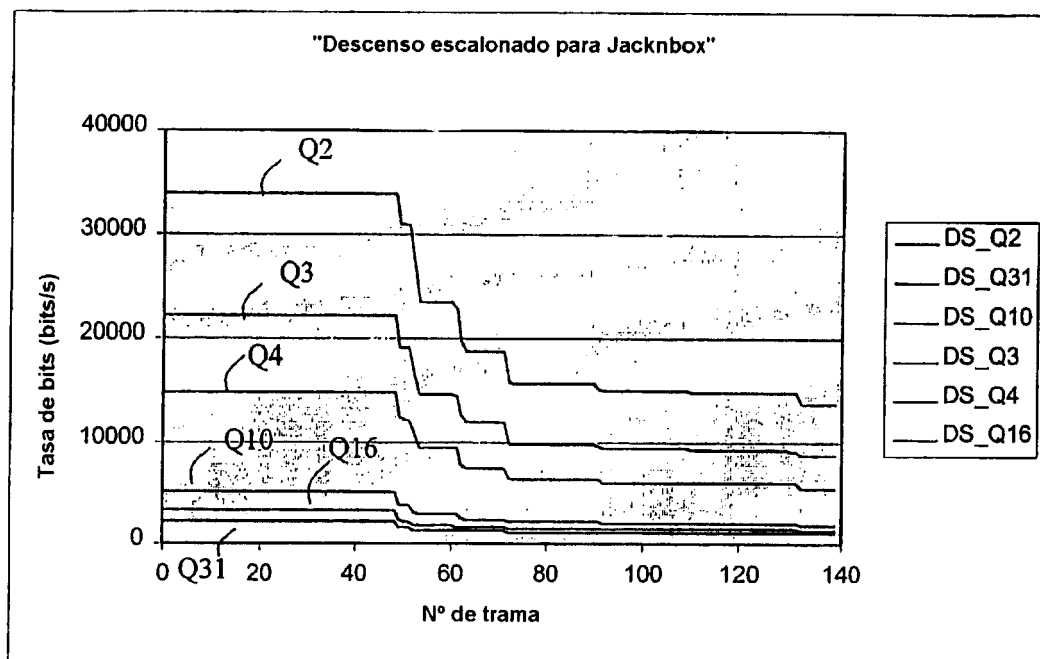


Figura 6

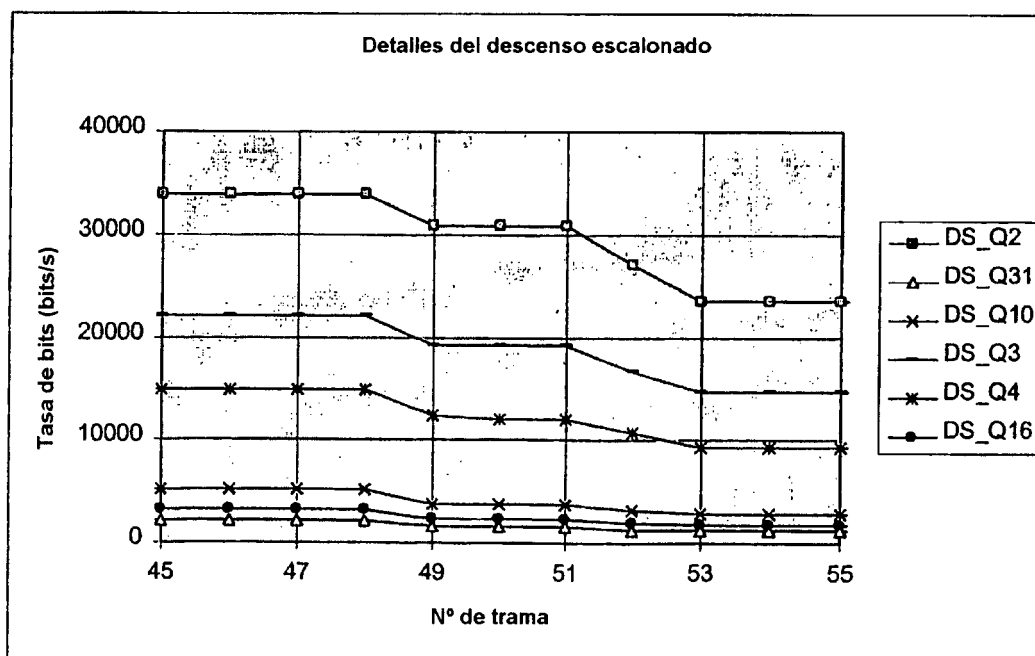


Figura 7

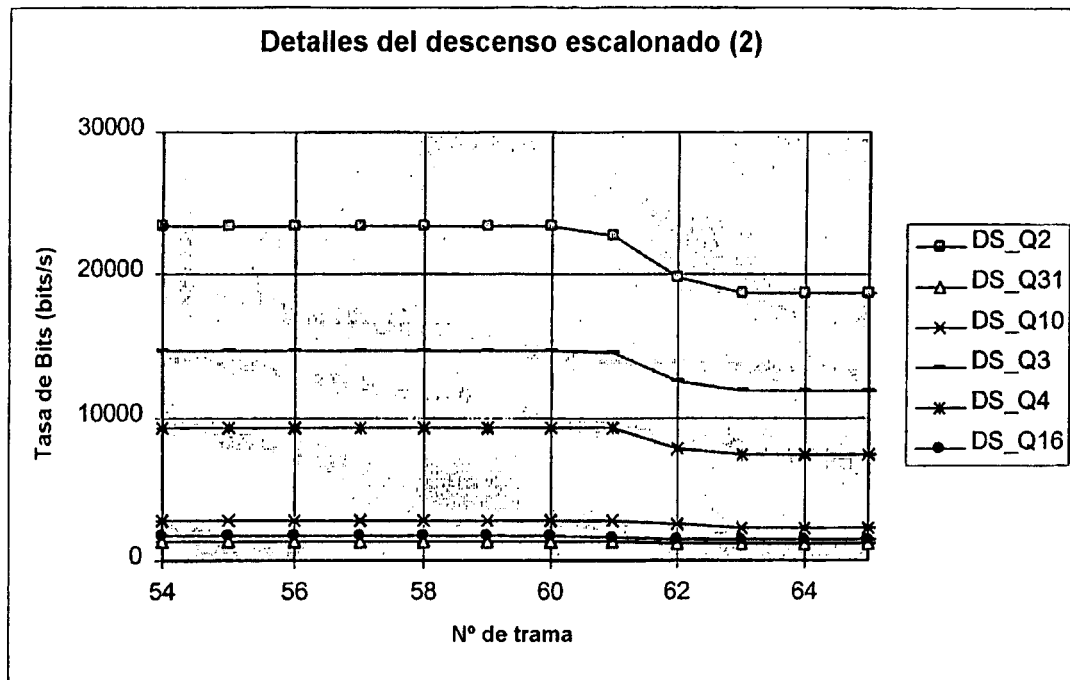


Figura 8

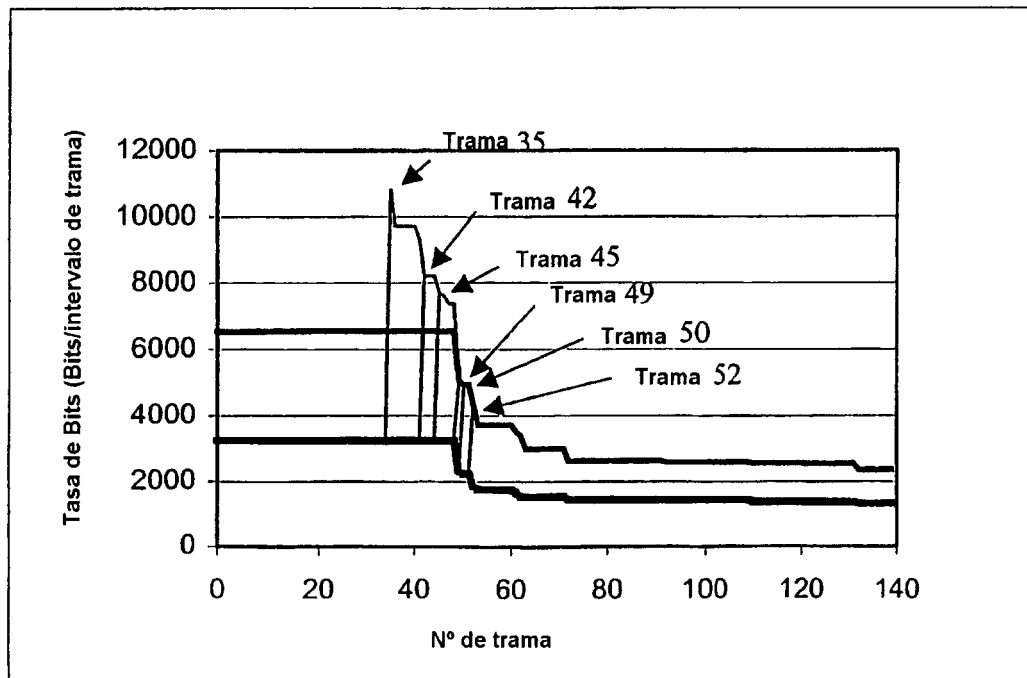


Figura 9

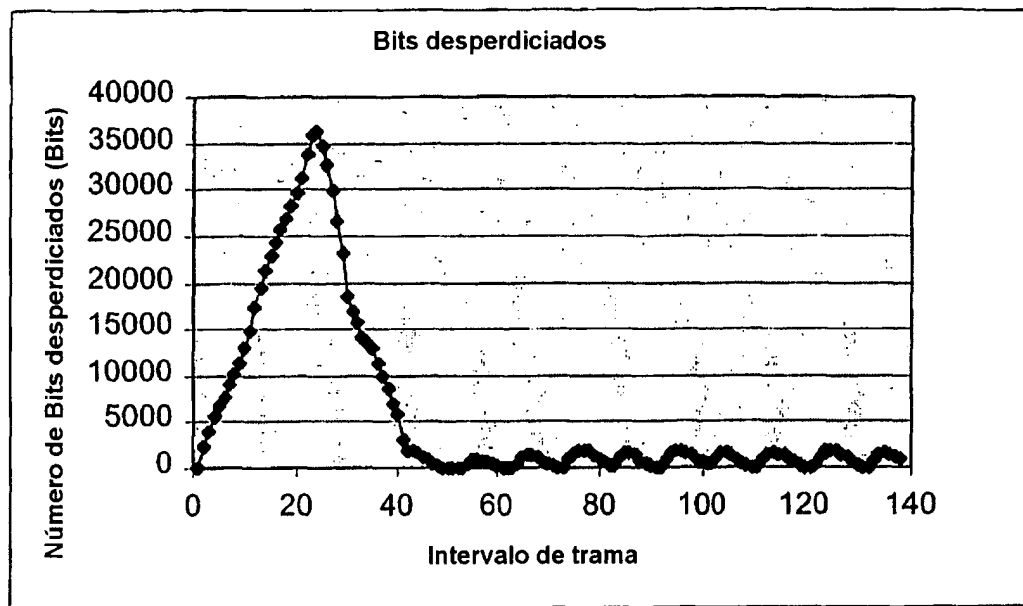


Figura 10