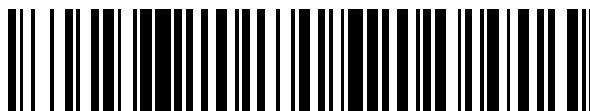


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 541 282**

51 Int. Cl.:

H04N 19/00

(2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.09.2011** **E 11763619 (1)**

97 Fecha y número de publicación de la concesión europea: **15.04.2015** **EP 2649801**

54 Título: **Método y aparato para la valoración objetiva de la calidad de video basado en estimaciones continuas de la visibilidad de la pérdida de paquetes**

30 Prioridad:

10.12.2010 EP 10194580

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

17.07.2015

73 Titular/es:

**DEUTSCHE TELEKOM AG (100.0%)
Friedrich-Ebert-Allee 140
53113 Bonn, DE**

72 Inventor/es:

**ARGYROPOULOS, SAVVAS;
FEITEN, BERNHARD;
GARCIA, MARIE-NEIGE;
LIST, PETER y
RAAKE, ALEXANDER**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 541 282 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

Método y aparato para la valoración objetiva de la calidad de video basado en estimaciones continuas de la visibilidad de la pérdida de paquetes

Campo de la invención

- 5 La presente invención se refiere a un método y a un aparato para la valoración de la calidad de video basada en el análisis de la visibilidad de las pérdidas de paquetes que se pueden producir durante la transmisión de una secuencia de video a través de una red propensa a presentar errores.

Antecedentes de la invención

- 10 La proliferación de contenidos de video ha necesitado durante los últimos años el desarrollo de métodos objetivos de valoración de calidad de video. Es evidente que los parámetros de red que definen la Calidad de Servicio (QoS) no son suficientes para estimar la calidad de servicio percibida por el usuario, típicamente referida como Calidad de Experiencia (QoE). Los métodos de estimación de la calidad soportan comúnmente una estimación diferenciada de la calidad asociada a la codificación (compresión, *Qcod*) de la señal de video y de la calidad debida a la pérdida de paquetes durante la transmisión (*Qtrans*). Es una propiedad inherente de las redes-IP que (sobre todo debido a la
- 15 sobrecarga temporal en algún punto de la red) que los paquetes IP se puedan perder. Algunas de estas pérdidas pueden resultar prácticamente invisibles para el usuario mientras que otras pueden producir una severa degradación de la calidad de video. Incluso en el caso en que se incluyan contramedidas contra estas pérdidas son parte de un sistema de distribución de IPTV, estos medios nunca pueden garantizar un remedio absoluto. Por ejemplo, una petición de retransmisión puede tardar demasiado, o el propio paquete retransmitido puede acabar perdiéndose. Por
- 20 lo tanto, siempre existe una probabilidad distinta de cero, de que se transmitan trenes o flujos de bits fragmentarios hacia el dispositivo del usuario final. Esto a su vez puede causar degradaciones visibles o audibles en el video reconstruido. Las categorías de medición pueden incluir además por tanto valores para expresar la probabilidad de las pérdidas. Este tipo de valores pueden incluir la expresión de una *"tasa de pérdida de paquetes"* y de la *"impulsividad de los eventos de pérdida"*. En el documento de Y. Liang, et al. "Analysis of Packet Loss for
- 25 Compressed Video: Effect of Burst Losses and Correlation between Error Frames." en IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, núm. 7, de Julio 2008, se investigó el impacto de las pérdidas en forma de ráfaga sobre la calidad de video. Se demostró que el patrón de pérdida tiene un efecto significativo sobre la distorsión global y que las pérdidas en forma de ráfaga producen un impacto más severo que las pérdidas individuales. Sin embargo, este estudio no considera la distorsión percibida desde el punto de vista de usuario y se propone principalmente para la estimación de la distorsión en el lado del codificador, donde no se conocen las
- 30 pérdidas de paquete reales.

- Los modelos paramétricos de calidad de video para aplicaciones de IPTV evalúan la distorsión tanto debida a compresión como debida a una transmisión errónea (véase, por ejemplo, el documento de K. Yamagishi y T. Hayashi, "Parametric Packet-Layer Model for Monitoring Video Quality of IPTV Services", en Proc. of IEEE Int. Conf.
- 35 on Communications, 2008, o el documento de M. N. Garcia y A. Raake, "Parametric packet-layer video quality model for IPTV", Int. Conf. en Information Science, Signal Processing and their Applications (ISSPA), Kuala-Lumpur, Mayo 2010).

- Sin embargo, los modelos basados puramente en cabecera no pueden cubrir precisamente el impacto de las alteraciones de la red sobre la calidad visual en términos de las características espacio-temporales de la secuencia
- 40 de video y de las propiedades estadísticas de las pérdidas de paquetes. De ese modo, los modelos de valoración objetiva de la calidad de video deben analizar la relación entre las pérdidas de paquetes y la degradación visual y tener en cuenta el factor de que las pérdidas de paquetes no producen una cantidad igual de degradación percibida. Una revisión de los diferentes tipos de modelos objetivos basados en paquetes, flujo de bits o híbridos, y de la diferente cantidad de información disponible para cada uno de ellos se presenta en el documento de S. Winkler y P.
- 45 Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics", IEEE Trans. en Broadcasting, vol. 54, núm. 3, de septiembre de 2008. Una revisión más detallada de las diferentes capas de información para los algoritmos de supervisión de la calidad de video exentos de referencia se presenta en el documento de A. Takahashi, et al., "Standardisation activities in ITU for a QoE assessment of IPTV", IEEE Communications Magazine, vol. 46, núm. 2, febrero de 2008.

- 50 El problema de predecir la visibilidad de la pérdida de paquetes ha sido abordado en la bibliografía sólo en términos de clasificación de pérdidas de paquetes en un modo binario: visible o invisible. En el documento de Kanumuri, P. C. Cosman, A. R. Reibman, V. A. Vaishampayan, "Modeling packet-loss visibility en video MPEG-2", IEEE Trans. On Multimedia, vol. 8, núm. 2, de abril de 2004, págs. 341 a 355, se extrajo un conjunto de características a partir del flujo de bits MPEG-2 y dos enfoques de modelización, se utilizó un Modelo Lineal Generalizado (GLM) para estimar
- 55 el número relativo de espectadores que han detectado un error, y un clasificador de árbol para determinar si una pérdida de paquete produce una degradación visible. El algoritmo fue extendido a video H.264/AVC en el documento de S. Kanumuri, et al., "Predicting H.264 Packet Loss Visibility using a Generalized Linear Model", en Proc. of ICIP, de octubre de 2006. La clasificación anterior fue extendida para video H.264/AVC en el documento de S. Kanumuri, S. B. Subramanian, P. C. Cosman, A. R. Reibman, "Predicting H.264 packet loss visibility using a generalized linear

model”, en Proc. of IEEE Int. Conf. on Image Processing (ICIP), Atlanta, Georgia, en octubre de 2006, en el cual se examinó el efecto de pérdidas duales de paquetes, y en el documento de T. L. Lin, S. Kanumuri, Y. Zhi, D. Poole, P. C. Cosman y A. R. Reibman, “A versatile model for packet loss visibility and its application to packet prioritization”, IEEE Trans. en Image Processing, vol. 19, núm. 3, págs. 722 a 725, de marzo de 2010, en el que el esquema propuesto fue utilizado para la priorización de paquetes en los encaminadores intermedios de una red.

Adicionalmente, se desarrolló un clasificador de árbol de decisión basado en flujo de bits exento de referencia para secuencias de CIF según el documento de N. Staelens et al., “Viqid: A no-reference bit stream-based visual quality impairment detector”, en IEEE Workshop on Quality of Multimedia Experience, Trondheim, Noruega, en junio de 2010. En ese caso, el efecto del patrón y de la longitud de la pérdida de paquetes sobre la calidad subjetiva es una cuestión no respondida. En el documento de Y. J. Liang, J. G. Apostolopoulos, y B. Girod, “Analysis of packet loss for compressed video: effect of burst losses and correlation between error frames”, IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, núm. 7, págs. 861 a 874, de julio de 2008, se analizó el efecto de las pérdidas en forma de ráfaga sobre la calidad del video reconstruido y se demostró que un patrón de pérdida específico produce una mayor degradación que un número igual de pérdidas aisladas. Asimismo, se tuvo en cuenta la correlación entre las tramas o marcos de error en el modelado de la distorsión inducida. Sin embargo, el algoritmo fue puesto a prueba únicamente sobre secuencias de QCIF y, de ese modo, con un esquema de paquetización en el cual un marco individual está contenido en un paquete. Además, no se comprobó el impacto sobre las valoraciones subjetivas.

Adicionalmente, en el documento de F. Yang, et al., “No-reference quality assessment for Networked Video via Primary Analysis of the Bit-Stream”, IEEE Circuits and Systems for Video Technology, vol. 20, núm. 11, págs. 1544 a 1554, de noviembre de 2010, se presenta un algoritmo de supervisión de la calidad de video para evaluar la degradación debida a errores de cuantización y de transmisión. El impacto de la pérdida de paquetes sobre la calidad percibida es ponderada mediante la complejidad temporal de los marcos en los que se produce la pérdida.

Adicionalmente, se presentó aún otro método para la supervisión de la calidad de video en el documento de M. Mu et. al., “A discrete perceptual impact evaluation quality assessment framework for IPTV services”, en Proc. of ICME, de 2010, basado en la suma de degradaciones de video debida a irregularidades de la red.

La diferencia principal de la presente invención es que a la visibilidad de cada evento de pérdida de paquetes se le asigna un “valor de visibilidad” único que es posteriormente explotado en la fórmula de valoración objetiva de la calidad de video para tener en cuenta el impacto de la pérdida de paquetes sobre la calidad percibida. Asimismo, las características utilizadas, que son obtenidas y calculadas a partir del flujo de bits de video, y el algoritmo para la fusión de estas características para el cálculo de la estimación de la visibilidad para cada pérdida son nuevas respecto a los métodos anteriores.

Compendio de la invención

Es el objetivo de la presente invención es facilitar un método y un aparato para la valoración objetiva de la calidad de una secuencia de video. Este objetivo se alcanza mediante las características de las reivindicaciones.

El método y el aparato de la invención facilita la valoración objetiva de la calidad de una secuencia de video en base a por lo menos una o a un conjunto de características extraídas a partir del flujo de bits de video y en la predicción de estimaciones continuas de probabilidad de la visibilidad de las pérdida de paquetes de paquetes que se pueden producir debido a la transmisión de la secuencia de video sobre una red propensa a presentar errores. Aparentemente, el método de la invención es un método de valoración de la calidad de video basado en flujo de bits exento de referencia, debido a que extrae información únicamente a partir del flujo de bits de video recibido sin necesitar una secuencia explícita de referencia.

Breve descripción de los dibujos

Otros aspectos, características y ventajas resultarán evidentes a partir del compendio anterior, así como a partir de la siguiente descripción, incluyendo las figuras y las reivindicaciones.

La Figura 1 representa la arquitectura del modelo de valoración de la calidad de video propuesto;

la Figura 2 representa cómo los marcos son organizados en un “grupo-de-imágenes” (GOP) para su video-codificación;

la Figura 3 representa la partición de macrobloques según la norma H.264/AVC y la obtención de los coeficientes de transformación después de aplicar la transformación de 4x4 números enteros;

la Figura 4 representa unos mapas de propagación de error para cuatro marcos consecutivos en la secuencia de video;

la Figura 5 muestra el diagrama de bloques de la segunda realización de la presente invención para la evaluación de calidad de video para la valoración de la calidad de video; y,

la Figura 6 representa un ejemplo para la estimación de la distorsión visible inducida debida a la(s) pérdida(as) en los marcos de una secuencia de video.

Descripción detallada

En la Figura 1 se muestra la arquitectura del modelo de valoración de la calidad de video según una realización preferida de la invención. La Figura 1 representa el diagrama de bloques del método y del aparato preferidos para la valoración objetiva de la calidad de video basada en la estimación de la visibilidad de eventos de pérdida de paquetes. En particular, se muestra que es un método basado en flujo de bits exento de referencia quien extrae o calcula ocho características a partir del flujo de bits de video recibido para evaluar el impacto perceptual de cada pérdida de paquetes. Posteriormente, se utilizan las características extraídas y el valor pronosticado para la visibilidad de la pérdida de paquetes para evaluar la calidad global del flujo de bits de video.

En el receptor, un dispositivo de sondeo captura el flujo de bits y extrae o calcula un número de características que son utilizadas para la valoración de la calidad de video. Las características son extraídas o bien directamente mediante la descodificación parcial del flujo de bits o bien combinando una información procedente de los píxeles reconstruidos (descodificación completa del flujo de bits). A continuación, las características extraídas son suministradas al módulo responsable de la determinación de la visibilidad de cada evento de pérdida de paquetes. Este módulo asigna una estimación continua de la visibilidad para cada evento de pérdida de paquetes (pérdidas de paquetes aisladas o combinación de eventos) que se producen dentro de un lapso de tiempo específico. Posteriormente, la estimación de la probabilidad del módulo de visibilidad de la pérdida de paquetes es combinada con las características extraídas previamente para evaluar la calidad global de la secuencia. Específicamente, en base a las estimaciones de la probabilidad de la visibilidad de cada pérdida de paquetes, las características extraídas son ponderadas en el algoritmo final para reflejar que cada pérdida de paquetes produce unas degradaciones percibidas desiguales sobre la secuencia de video. Finalmente, el dispositivo de sondeo entrega el valor pronosticado de la calidad de video, Q.

Método para la extracción y el cálculo de características a partir del flujo de bits

A continuación se describen las características extraídas preferiblemente a partir del flujo de bits para determinar la visibilidad de las pérdida de paquetes y la calidad objetiva de la secuencias de video. Nótese que en este ejemplo el análisis está basado en flujos codificados según la norma de codificación de video H.264/AVC (T. Wiegand, G. J. Sullivan, G. Bjontegaard, y A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Trans. on Circuits and Systems for Video Technology, vol. 13, núm. 7, de julio de 2003), pero el método puede ser aplicado a cualquier norma de codificación de video con pequeñas modificaciones del cálculo de las características correspondientes.

Una de las propiedades que se necesita calcular para la estimación de la degradación perceptual causada por una pérdida de paquetes es el "tipo de marco" de todos los marcos y en particular del marco afectado por la pérdida. Los posibles valores de la propiedad del "tipo de marco" incluyen "Intra-marco" o "Marco clave" (denominado a continuación marco-I), "Marco Predecido" (denominado a continuación marco-P) y "Marco bidireccional" (denominado a continuación marco-B). Sólo los marcos-I pueden ser descodificados sin información de referencia de ninguno de los marcos precedentes. Por el contrario, los marcos-P dependen de uno o más marcos predecesores denominados "marcos de referencia", debido a que la información transmitida para un marco-P consiste principalmente en la diferencia entre el marco de video que describe y sus referencias. Por lo tanto, las pérdidas de paquetes dentro de un marco-I o sus consecutivos marcos-P se propagan a los siguientes marcos, incluso si estos siguientes marcos no contienen en ellas mismas ningún paquete perdido. Debido a este mecanismo, un simple error de pérdida de paquetes puede extenderse a través de largas partes de una secuencia de video, hasta que se produce el siguiente marco-I libre de error. Por lo tanto, los errores producidos en marcos-P y particularmente en marcos-I presentarán una visibilidad elevada. El mismo mecanismo de marco de referencia es válido para marcos-B, pero, debido a que los marcos-B no sirven por lo general como referencias por ellos mismos, un error producido en un marco-B sólo causará degradación en ese único marco.

La cadena de marcos de video comprendidos entre dos marcos-I sucesivos se denomina "Grupo de Imágenes" (GoP). En la mayoría de los casos los marcos-P y los marcos-B en un GoP siguen un patrón más o menos estricto tal como el típico patrón-GoP: "I, B, B, P, B, B, P ...". Sea N la denotación de la longitud-GoP, es decir, la distancia entre dos marcos-I sucesivos. Entonces, si la primera pérdida de paquetes se produce t marcos después del primer marco-I de un GoP específico, la primera característica que se extrae es el número de marcos que están degradados debido a la pérdida de paquetes, denominada **ErrDur**. Esta propiedad se calcula tal como (véase la Figura 2):

$$ErrDur = \begin{cases} N - t + 1 & , \text{ si el tipo de marco es I ó P} \\ 1 & , \text{ si el tipo de marco es B} \end{cases} \quad \text{Eq. (1)}$$

De ese modo, si la pérdida de paquetes se produce en el primer marco (que es un marco-I), entonces $t = 1$, y el número de marcos dañados por la pérdida de paquetes es N , que es igual al número de marcos en el GoP.

Entonces se calcula, preferiblemente, la magnitud promedio de los vectores de movimiento, **AvgMv**, que refleja al vector de movimiento promedio en las direcciones horizontal y vertical de todos los macrobloques (mb) afectados por la pérdida. Sea $v_n = [v_{n,x}, v_{n,y}]$ la denotación vector de movimiento del macrobloque n -ésimo en el marco en la que se ha producido la pérdida de paquetes. A continuación, el término **AvgMv** se calcula como:

$$AvgMv = \frac{1}{L} \sqrt{\left(\sum_{i=1}^L v_{i,x}\right)^2 + \left(\sum_{i=1}^L v_{i,y}\right)^2} \quad \text{Eq. (2)}$$

donde L es la cardinalidad del conjunto de los macrobloques que se han perdido en el marco n debido a la pérdida de paquetes. Aparentemente, debido a que no se puede recuperar la información del vector de movimiento para los macrobloques perdidos, se estima esta información a partir de sus vecinos temporales. Es decir, para cada macrobloque perdido se recupera la información del vector de movimiento a partir del macrobloque co-ubicado en el marco correcto recibido previamente.

Adicionalmente, se puede calcular la diferencia promedio del vector de movimiento, **AvgMvDiff**, a partir del flujo de bits. En H.264/AVC, para explotar la redundancia entre los vectores de movimiento, sólo se incluye en el flujo de bits la diferencia entre el vector de movimiento de un bloque y su vector de movimiento pronosticado a partir de los macrobloques vecinos. Esta característica es extremadamente significativa en secuencias con un movimiento predecible (por ejemplo, enfoque o "panning"), donde la movimiento promedio puede ser grande, pero donde las diferencias de vector de movimiento codificadas en el flujo de bits son pequeñas. Sea $vd_n = [vd_{n,x}, vd_{n,y}]$ la denotación de la diferencia del vector de movimiento del macrobloque n -ésimo en el marco en el que se ha producido la pérdida de paquetes. Entonces, el término **AvgMvDiff** se calcula como:

$$AvgMv = \frac{1}{L} \sqrt{\left(\sum_{i=1}^L vd_{i,x}\right)^2 + \left(\sum_{i=1}^L vd_{i,y}\right)^2} \quad \text{Eq. (3)}$$

donde L es la cardinalidad del conjunto de los macrobloques que se han perdido en el marco n debido a la pérdida de paquetes. De modo similar al caso anterior, debido a que no se puede recuperar la información del vector de movimiento para los macrobloques perdidos, se estima esta información a partir de sus vecinos temporales. Es decir, para cada macrobloque perdido se recupera la información del vector de movimiento a partir del macrobloque co-ubicado en el marco correcto recibido previamente.

Otra característica que preferiblemente se extrae a partir del flujo de video recibido es la energía del residuo, **ResEnergy**, que está contenida dentro de los macrobloques perdidos y que se calcula a partir de los coeficientes de transformación del error de predicción. Sea $c_{b,i,j}$, $b = 0, \dots, 3$, $i = 0, \dots, 3$ y $j = 0, \dots, 3$ la denotación del coeficiente de transformación para la fila i -ésima y la columna j -ésima del bloque 4×4 b -ésimo de un macrobloque (véase la Figura 3). Entonces, la energía residual se calcula como:

$$ResEnergy = \frac{1}{L} \sum_{i=1}^L \sum_{b=1}^4 \sum_{i=0}^3 \sum_{j=0}^3 c_{b,i,j}^2 \quad \text{Eq. (4)}$$

donde i y L están definidas como anteriormente.

Otra característica que puede ser extraída es el número máximo de particiones, denominado **MaxPartNr**, en el marco en el que se produce la pérdida. En H.264/AVC, cada macrobloque 16×16 puede ser dividido adicionalmente en bloques más pequeños para el proceso de estimación de la movimiento, es decir, bloques de tamaño 16×8 , 8×16 u 8×8 . Si se escoge el tamaño de partición 8×8 , cada partición 8×8 puede ser adicionalmente dividida en sub-particiones de tamaño 8×8 , 8×4 , 4×8 o transformación de 4×4 muestras de brillantez (luma). De ese modo, el parámetro **MaxPartNr** es igual al número máximo de particiones de los macrobloques correctamente recibidos en el marco en el que se produce la pérdida. Si todos los macrobloques son borrados (es decir, si se pierde todo el marco), entonces se obtiene el parámetro a partir del marco recibido previamente.

La sexta característica que se extrae a partir del flujo de bits de video recibido es el número de macrobloques que no son descodificables debido a la pérdida de paquetes y que necesitan ser ocultados, denominada **LostMbs** a continuación.

Adicionalmente, en base a la información del vector de movimiento y al tipo de macrobloques, se puede generar un mapa binario de propagación de error para cada imagen, el cual denota a los píxeles en cada imagen que reciben el impacto de la(s) pérdida(s) de paquetes ya sea debido al error de innovación o debido a la propagación de error hacia los marcos posteriores del mismo GoP debida a la codificación predictiva. Sea $I(x,y,k)$ la denotación del valor del mapa binario de error en la ubicación (x,y) del marco k -ésimo, $x = 1, 2, \dots, H$, e $y = 1, 2, \dots, W$ donde H, W son el alto y el ancho de cada marco de la secuencia de video, respectivamente. Asimismo, $k = 1, \dots, K$, donde K es el

número de marcos que contienen píxeles dañados debido a la pérdida de paquetes. El valor para esos píxeles que están, o bien dañados debido a la pérdida de paquetes o bien están asociados con esas regiones y tienen probabilidad de ser erróneos, es establecido a uno, en caso contrario el valor se establece a cero. De ese modo:

$$I(x, y, k) = \begin{cases} 1, & \text{si el píxel en la posición } (x, y) \text{ del marco } k \text{ está dañado} \\ 0, & \text{si el píxel en la posición } (x, y) \text{ del marco } k \text{ no está dañado} \end{cases} \quad \text{Eq. (5)}$$

5 En otras palabras, el valor en la ordenación de mapas de propagación es establecido a uno para todos los píxeles que no son idénticos entre la secuencia de video original y la imagen reconstruida del decodificador dentro de la sonda debido al error causado por la pérdida de paquetes y su propagación. En la Figura 4 se representa un ejemplo de la deriva de los mapas de propagación de error.

10 La Figura 4 representa los mapas de propagación de error para cuatro marcos consecutivos en la secuencia de video para demostrar cómo el error inicial causado por una pérdida de paquetes se extiende a los siguientes marcos que están asociados al marco inicial para su predicción. En particular, en la Figura 4a, la región oscura corresponde a la región que no puede ser decodificada debido a que la información para estos píxeles estaba contenida en un paquete que fue borrado. Como resultado, el decodificador necesita ocultar (utilizando normalmente vecinos temporales o espaciales) esta región para reemplazar los píxeles perdidos. Debido a que la ocultación no puede recuperar exactamente la información original, se introduce cierta degradación en esos marcos. Las Figuras 4b, 4c y 15 4d muestran cómo este error se propaga hacia los siguientes marcos. Las regiones oscuras corresponden a píxeles asociados a una información relacionada a la región inicialmente pérdida (y ocultada) y son, por lo tanto, propensos a ser también erróneos.

20 La pérdida de paquetes se produce en el primer marco (la región oscura en la figura denota los macrobloques que han sido borrados debido a la pérdida de paquetes) y se propaga a los marcos posteriores que están asociados a la región afectada del primer marco para el proceso de predicción. Nótese que en realizaciones alternativas el mapa de propagación de error puede ser estimado a nivel de macrobloque, significando que la propagación es rastreada para cada macrobloque en lugar de para cada píxel.

25 En base a los mapas de propagación de error, el número total de píxeles dañados debido a la pérdida de paquetes, denominado **ErrProp**, se calcula como:

$$ErrProp = \sum_{k=1}^K \sum_{x=1}^H \sum_{y=1}^W I(x, y, k) \quad \text{Eq. (6)}$$

30 Finalmente, se puede extraer otra característica basándose en los vectores de moción, los tipos de macrobloques y los residuos, denominada **EstErr**, para cuantificar la magnitud de la distorsión inducida debido a la pérdida de paquetes y debido a la propagación del error que permite la estimación del error cuadrático medio (MSE) inducido de un modo exento de referencia. El método para estimar la distorsión inducida está fuera del alcance de esta invención, pudiéndose usar por tanto cualquier método conocido por cualquier experto en la técnica, por ejemplo el método propuesto en M. Naccari, M. Tagliasacchi, y S. Tubaro, "No-reference video quality monitoring for H.264/AVC coded video", IEEE Trans. on Multimedia, vol. 11, núm. 5, págs. 932 a 946, de agosto de 2009.

35 En el ejemplo mostrado, las ocho características mencionadas anteriormente son extraídas para cada pérdida de paquetes que se produce en un GoP. Cuando aparece más de una pérdida de paquetes en el mismo GoP, entonces, en la realización preferida, se escoge el valor máximo para **ErrorDur**, y **MaxPartNr**, mientras que para las otras características se considera su suma como características finales. Sin embargo, en otras realizaciones, las características extraídas a partir de cada pérdida de paquetes dentro de un GoP se pueden combinar de manera diferente, por ejemplo mediante el calculo de la media, la mediana o cualquier otra función. De ese modo, dentro de 40 cada GoP, se extrae el siguiente vector de característica, el cual se utiliza para la estimación de la visibilidad de eventos de pérdida de paquetes:

$$f = [ErrDur, AvgMV, AvgMvDiff, ResEnergy, MaxPartNr, LostMbs, ErrProp, EstErr] \quad \text{Eq. (7)}$$

45 Se ha de observar que, en otras realizaciones de la presente invención, se puede utilizar cualquier número y combinación de las características mencionadas anteriormente. Esto es especialmente significativo cuando no se desea calcular ciertas características (por ejemplo, debido a razones de complejidad o a costes de implementación).

Estimación continua de la visibilidad de la pérdida de paquetes utilizando regresión de vectores de soporte

50 Para la clasificación de la visibilidad de la pérdida de paquetes en base a las características extraídas de cada evento de pérdida de paquetes, se utiliza una técnica basada en Regresión de Vectores de Soporte (SVR). Con este método, el vector de características de entrada es mapeado sobre un espacio de características de mayor dimensión

utilizando una función no lineal de mapeo y a continuación se construye un modelo lineal en este espacio de características. El algoritmo funciona en dos etapas: la etapa de entrenamiento, en la que se utilizan los datos de entrenamiento para ajustar los parámetros del modelo, y la etapa de evaluación, en la que el modelo produce el valor pronosticado para el vector de características de entrada. Estas dos etapas son descritas en detalle a continuación.

En la etapa de entrenamiento, un conjunto de n datos de entrenamiento, denotados como $\{(f_1, y_1), (f_2, y_2), \dots, (f_n, y_n)\}$, $f_i \in \mathcal{R}^8$, $y_i \in \mathcal{R}$, $i = 1, \dots, n$, donde f_i es un vector de características descrito en la Eq. (7), e y_i es el valor objetivo de salida, son utilizados para estimar los parámetros del modelo que maximizan su precisión de predicción. El método para entrenar el algoritmo SRV está fuera del alcance de la presente invención, de modo que se puede utilizar cualquier método conocido por el experto en la técnica. El resultado del algoritmo de entrenamiento en conjunto con los datos de entrenamiento de entrada da lugar al cálculo de un conjunto de parámetros α_i , $i = 1, \dots, n$, β_i , $i = 1, \dots, n$, y $b \in \mathcal{R}$, que son utilizados durante la etapa de evaluación. En la presente invención, se escoge el valor de b igual a 1,27. Sin embargo, se puede utilizar cualquier otro valor.

En la etapa de evaluación, el algoritmo recibe un vector de características de entrada f (tal como se describe en la Eq. (7)) y el valor pronosticado de la visibilidad de cada pérdida, V , viene dado por:

$$V(f) = \sum_{i=1}^n (\alpha_i - \beta_i) K(f, f_i) + b \quad \text{Eq. (8)}$$

donde $K(\dots)$ es la función esencial y f_i , $i = 1, \dots, n$ son los vectores de características utilizados como datos de entrenamiento. En la presente invención, se ha elegido la siguiente función como función esencial, la cual también se conoce como Función de Base Radial (RBF):

$$K(f, f_i) = e^{-\frac{\|f - f_i\|^2}{\gamma}} \quad \text{Eq. (9)}$$

En la presente invención, el valor de γ se escoge igual a 2. Además, se puede utilizar, cualquier otra función esencial en otras realizaciones. Adicionalmente, se puede convertir el valor pronosticado V de la visibilidad de una pérdida de paquetes en un valor binario, cuando sea deseable clasificar una pérdida de paquetes en uno de las siguientes dos opciones: visible o invisible. En ese caso, el valor pronosticado binario de una pérdida de paquetes, denotado como V_B se puede calcular como:

$$V_B = \begin{cases} 0 & , \text{ si } V < T \\ 1 & , \text{ si } V \geq T \end{cases} \quad \text{Eq. (10)}$$

donde T es un valor umbral para clasificar la estimación continua de la visibilidad de la pérdida de paquetes en un valor binario. Dentro de la presente invención, el valor sugerido es 0,7, pero se puede utilizar cualquier otro valor.

Cuando el algoritmo funciona solamente en la etapa de evaluación y no necesita entrenamiento sobre datos nuevos, se utilizan los parámetros almacenados y se predice la visibilidad V de cada evento de pérdida de paquetes tal como se ha definido en la Eq. (8). Las características extraídas que han sido mencionadas anteriormente son refinadas de manera que reflejan el valor pronosticado de la visibilidad de cada pérdida de paquetes y son ponderadas de forma correspondiente, tal como se analiza en la siguiente sección.

Método de valoración de la calidad de video

El modelo propuesto de valoración de la calidad de video basado en flujo de bits es una combinación de la degradación producida por la compresión de las secuencias de video, Q_{cod} , y de la degradación de la calidad debida a la pérdida de paquetes durante una transmisión (Q_{trans}). De ese modo, la calidad global viene dada por:

$$Q = Q_0 - Q_{cod} - Q_{trans} \quad \text{Eq. (11)}$$

donde Q denota la calidad global de la secuencia y Q_0 denota la calidad de la fuente del video insertada en la cadena de transmisión. El término Q_{cod} se calcula mediante:

$$Q_{cod} = a_1 \cdot e^{a_2 B} + a_3 \quad \text{Eq. (12)}$$

donde B es la tasa de bit de la secuencia de video, y a_1 , a_2 , y a_3 son constantes. En el ejemplo presente, los valores de a_1 , a_2 , y a_3 se escogen iguales a 89,33, -1,21 y 11,47.

El término de la degradación visual inducida por el canal refleja la distorsión producida por la pérdida de paquetes, la extensión de su propagación y la intensidad en la percepción visual humana. Por este motivo, se utiliza la siguiente fórmula para Q_{trans} :

$$Q_{trans} = f(V_i, EstErr_i, ErrProp_i) \quad \text{Eq. (13)}$$

En este caso, el índice i se utiliza para denotar un único evento de pérdida de paquetes particular que se produce durante la secuencia y $f(\dots)$ representa cualquier función adecuada. En otras palabras, la degradación global debida a los errores de transmisión es una función de la visibilidad de cada pérdida de paquetes, del error que se induce sobre el marco cuando se produce la pérdida de paquetes y la propagación del error a los marcos posteriores.

A continuación, se presentan dos realizaciones de la presente invención para representar la función $f(\dots)$ mencionada anteriormente con fórmulas específicas.

En la primera realización, la Eq. (13) se reescribe como:

$$Q_{trans} = a_4 \cdot \sum_i V_i \cdot EstErr_i + a_5 \cdot \sum_i V_i \cdot ErrProp_i + a_6 \cdot \sum_i V_i \cdot EstErr_i \cdot ErrProp_i \quad \text{Eq. (14)}$$

donde a_4 , a_5 , y a_6 son constantes determinadas mediante regresión, i es un índice correspondiente a cada pérdida de paquetes individual, $EstErr_i$ y $ErrProp_i$ son características extraídas tal como se ha mencionado anteriormente, asociadas con cada pérdida de paquetes. De ese modo, el modelo tiene en cuenta la importancia de cada pérdida de paquetes para la determinación de la degradación visual y utiliza la salida de la estimación de la visibilidad para ponderar la distorsión inducida de cada pérdida de paquetes. En la presente realización, los valores de a_4 , a_5 , y a_6 se escogen iguales a 0,023, 0,000176 y 0,0000465.

En la segunda realización, representada en la Figura 5, se utiliza otra implementación para la función $f(\dots)$ de la Eq. (14). Sea i la denotación del índice de una pérdida de paquetes individual que tiene lugar dentro de un GoP. Asimismo, sea $EstErr_i$ la distorsión inducida en el marco donde ha tenido lugar la pérdida de paquetes i -ésima, estimada tal como se ha explicado anteriormente en la sección de extracción de características. Entonces, la degradación visible asociada con esta pérdida de paquetes viene dada por:

$$VisEstErr_0^i = V_i \cdot EstErr_i \quad \text{Eq. (15)}$$

Asimismo, sea $NrImpPx_k^i$ la denotación del número de píxeles dañados en el marco k debido a la pérdida de paquetes i -ésima, o de modo equivalente

$$NrImpPx_k^i = \sum_{x=1}^H \sum_{y=1}^W I(x, y, k) \quad \text{Eq. (16)}$$

donde $I(x,y,k)$ denota el valor del mapa binario de error en la ubicación (x,y) del marco k -ésimo, $x = 1, 2, \dots, H$, y $y = 1, 2, \dots, W$ donde H , W son el alto y el ancho de cada marco de la secuencia de video, respectivamente. Por tanto, la degradación percibida en el marco posterior k , $k > 0$, debido a la propagación del error de la pérdida de paquetes i -ésima viene dada por:

$$VisEstErr_k^i = VisEstErr_0^i \cdot \frac{NrImpPx_k^i}{NrImpPx_0^i} \quad \text{Eq. (17)}$$

En este caso, $NrImpPx_0^i$ es el número de píxeles dañados a causa del error de innovación (el error en el primer marco que ha sido afectado por la pérdida de paquetes bajo consideración) debido al evento de pérdida de paquetes i , al cual se le asigna el índice-de-marco relativo $k = 0$. En la Figura 6 se representa un ejemplo del cálculo del término $VisEstErr_k^i$ para los marcos posteriores con índice k . En esta figura, la región oscura en el marco con índice t corresponde al término $NrImpPx_0^i$, mientras que la región oscura en los marcos posteriores (índices $t+1, \dots, t+m$) corresponde a los términos $NrImpPx_k^i$, $k = 1, \dots, m$.

Por tanto, la degradación visual global causada por la pérdida de paquetes i -ésima, denotada $VisEstErr^i$, puede ser considerada como la suma de las degradaciones causadas a todos los marcos K_i que fueron dañados debido a esta pérdida de paquetes, de manera que:

$$VisEstErr^i = \sum_{k=0}^{K_i} VisEstErr_k^i \quad \text{Eq. (18)}$$

Nótese que el sumatorio comienza desde el índice cero para tener en cuenta la distorsión inicial tal como se ha descrito en la Eq. (15). Se calcula la distorsión total dentro de un GoP específico como la suma de las distorsiones generadas por todas las pérdidas de paquetes que se producen dentro de este GoP, de modo que:

$$VisEstErr_{GOP} = \sum_i VisEstErr^i \quad \text{Eq. (19)}$$

Finalmente, sea G la denotación del número total de GoPs dentro de una secuencia, y T la denotación de la duración total de la secuencia. Entonces, la distorsión global debida a los errores de transmisión viene dada por:

$$VisErr_{SEQ} = \frac{1}{T} \sum_{g=1}^G VisEstErr_{GOP} \quad \text{Eq. (20)}$$

$$Q_{trans} = f(VisErr_{SEQ}) \quad \text{Eq. (21)}$$

En la realización preferida de este segundo tipo, se establece que

$$Q_{trans} = \theta \cdot \log(\xi \cdot VisErr_{SEQ} + 1) \quad \text{Eq. (22)}$$

Donde θ y ξ son constantes que son, por ejemplo, determinadas en un procedimiento de regresión que utiliza valoraciones de la calidad a modo de valores objetivo.

- 15 Aunque la invención ha sido ilustrada y descrita en detalle en las figuras y en la descripción anterior, dichas ilustración y descripción deben ser consideradas ilustrativas o ejemplares y no restrictivas. Debe entenderse que los expertos en la técnica pueden hacer cambios y modificaciones dentro del alcance de las siguientes reivindicaciones. En particular, la presente invención cubre realizaciones adicionales con cualquier combinación de características de las diferentes realizaciones descritas anteriormente y a continuación.
- 20 Adicionalmente, en las reivindicaciones las palabras “que comprende” no excluyen otros elementos o pasos, y el artículo indefinido “un” o “uno” no excluye una pluralidad. Una unidad simple puede satisfacer las funciones de varias características descritas en las reivindicaciones. Los términos “esencialmente”, “más o menos”, “aproximadamente” y similares en relación con un atributo o un valor, también definen particularmente de forma exacta al atributo o de forma exacta al valor, respectivamente. Cualquier signo de referencia en las reivindicaciones no debe ser
- 25 interpretado como limitativo del alcance.

REIVINDICACIONES

1. Método para evaluar la calidad de una señal de video transmitida en el lado del receptor, comprendiendo el método los pasos de:

a) capturar el flujo de bits de entrada y suministrarlo a un analizador de flujo de bits de video;

5 b) extraer al menos una característica o un conjunto de características a partir del flujo bits de video de entrada capturado mediante el analizador de flujo de bits;

c) entregar la característica extraída o el conjunto de características a un módulo de estimación de la visibilidad de la pérdida de paquetes;

10 d) determinar, mediante el módulo de estimación de la visibilidad de la pérdida de paquetes, utilizando las características extraídas suministradas del flujo de bits de video, la probabilidad continua de la visibilidad para cada evento de pérdida de paquetes que ocurra dentro de un intervalo de tiempo específico;

15 e) utilizar la probabilidad continua de la visibilidad de la pérdida de paquetes, determinada por el módulo de estimación de la visibilidad de la pérdida de paquetes, como un factor de ponderación de al menos una característica o conjunto de características extraídas a partir del flujo de bits de video para calcular una estimación de la calidad global, Q, de la secuencia de video transmitida;

20 en el que el paso (d) utiliza al menos una característica del flujo de bits del grupo que comprende: tipo de marco, magnitud promedio de los vectores de movimiento (AvgMv), diferencia promedio del vector de movimiento (AvgMvDiff), energía del residuo (ResEnergy), número máximo de particiones (MaxPartNr), número de macrobloques no-descodificables (LostMbs), información del vector de movimiento (mv), tipo de macrobloques (tipo mb); y,

en el que el paso (e) combina la estimación de la visibilidad (V) de la pérdida de paquetes con la magnitud determinada de la distorsión (EstErr) y el número total calculado de píxeles dañados debido a la pérdida de paquetes (ErrProp).

25 2. Método según la reivindicación 1, en el que la extracción de característica según el paso (b) se realiza directamente mediante la descodificación parcial del flujo de bits de video, donde "parcialmente" significa sin descodificar el flujo de bits a nivel de píxel.

3. Método según la reivindicación 1, en el que la extracción de característica según el paso (b) se realiza mediante la descodificación completa del flujo de bits de video y mediante la combinación de la información procedente de los píxeles reconstruidos de la señal de video.

30 4. El método de cualquiera de las reivindicaciones 1 a 3, en el que, en base a la característica extraída "tipo de marco", se determina el número de marcos (ErrDur) degradados por una pérdida de paquetes y se utiliza para la predicción de la probabilidad de la visibilidad de la pérdida de paquetes, así como para combinarlo con la probabilidad de la visibilidad para estimar el impacto de una pérdida de paquetes sobre la calidad visual.

35 5. El método de cualquiera de las reivindicaciones 1 a 4, en el que se genera un mapa binario de propagación de error para cada imagen de la secuencia de señal de video en base a la información del vector de movimiento y al tipo de macrobloques.

6. Método según la reivindicación 5, en el que se determina el número de marcos (ErrDur) degradados por una pérdida de paquetes en base al mapa de propagación de error y a la información del tipo de marco.

40 7. Método según la reivindicación 5 ó 6, en el que se calcula el número de total de píxeles dañados debido a la pérdida de paquetes (ErrProp) en base al mapa de propagación de error.

8. Método según cualquiera de las reivindicaciones 1 a 7, en el que se determina la magnitud de la distorsión (EstErr) inducida, debida a la pérdida de paquetes y a la propagación del error, en base a los vectores de movimiento, los tipos de macrobloque y los residuos.

45 9. Método según cualquiera de las reivindicaciones precedentes, en el que el paso (d) se lleva a cabo utilizando una técnica de regresión de vectores de soporte mapeando el vector de características de entrada, que comprende las características mencionadas anteriormente, sobre un espacio de características de gran dimensión utilizando una función no-lineal de mapeo y construyendo un modelo lineal en este espacio de características.

50 10. Método según cualquiera de las reivindicaciones precedentes, que comprende además, entre los pasos d) y e), el paso de estimar la distorsión visible global generada mediante la encuesta temporal de todas las pérdidas de paquetes que se producen dentro de la secuencia de señales de video.

11. Aparato para valorar la calidad de una secuencia de señal de video transmitida en el lado receptor,

comprendiendo el aparato:

- un analizador de flujo o flujo de bits de video que recibe la flujo de bits de video de entrada capturado, y que está configurado para extraer al menos una característica o un conjunto de características a partir del flujo bits de video de entrada capturado;
- 5 - un módulo de estimación de la visibilidad de la pérdida de paquetes que recibe la característica o el conjunto de características extraídas, estando el módulo de estimación de la visibilidad de la pérdida de paquetes configurado para determinar la visibilidad de una pérdida de paquetes ocurrida durante la transmisión de una señal de video mediante la predicción de la probabilidad continua de la visibilidad para cada evento de pérdida de paquetes que se produzca dentro de un intervalo de tiempo específico;
- 10 - un combinador para combinar la probabilidad continua de la visibilidad de la pérdida de paquetes, determinada por el módulo de visibilidad de la pérdida de paquetes, a modo de factor de ponderación, de al menos una característica o conjunto de características extraídas a partir del flujo de bits de video para calcular una estimación de la calidad global, Q, de la secuencia de señales de video transmitida;
- 15 en la que el módulo de estimación de la visibilidad de la pérdida de paquetes utiliza al menos una característica del flujo de bits del grupo que comprende: tipo de marco, magnitud promedio de los vectores de movimiento (AvgMv), diferencia promedio del vector de movimiento (AvgMvDiff), energía del residuo (ResEnergy), número máximo de particiones (MaxPartNr), número de macrobloques no-descodificables (LostMbs), información del vector de movimiento (mv), tipo de macrobloques (tipo mb); y,
- 20 en el que el combinador combina la estimación de la visibilidad (V) de la pérdida de paquetes con la magnitud determinada de la distorsión (EstErr) y el número total calculado de píxeles dañados debido a la pérdida de paquetes (ErrProp).

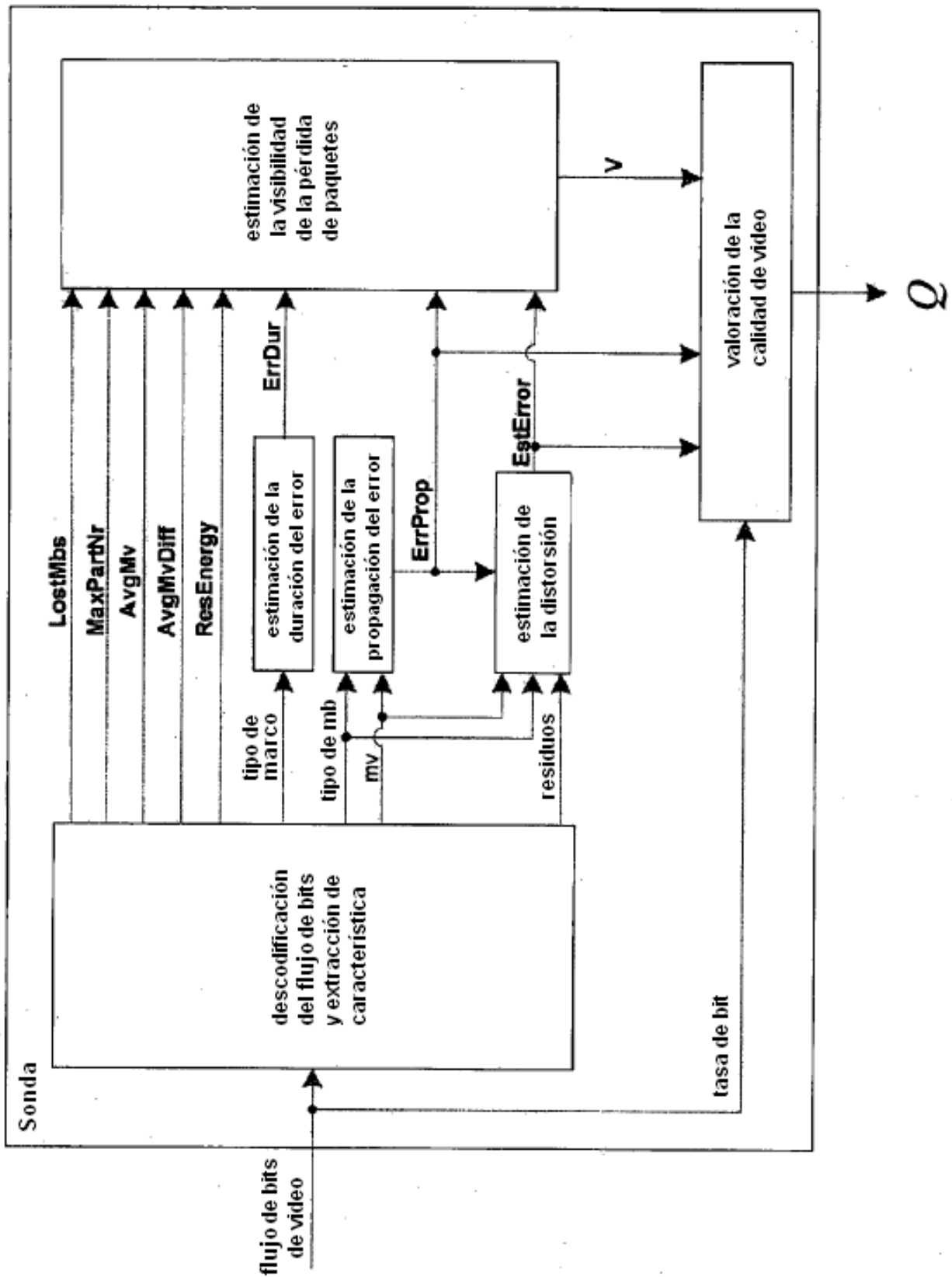


Fig. 1

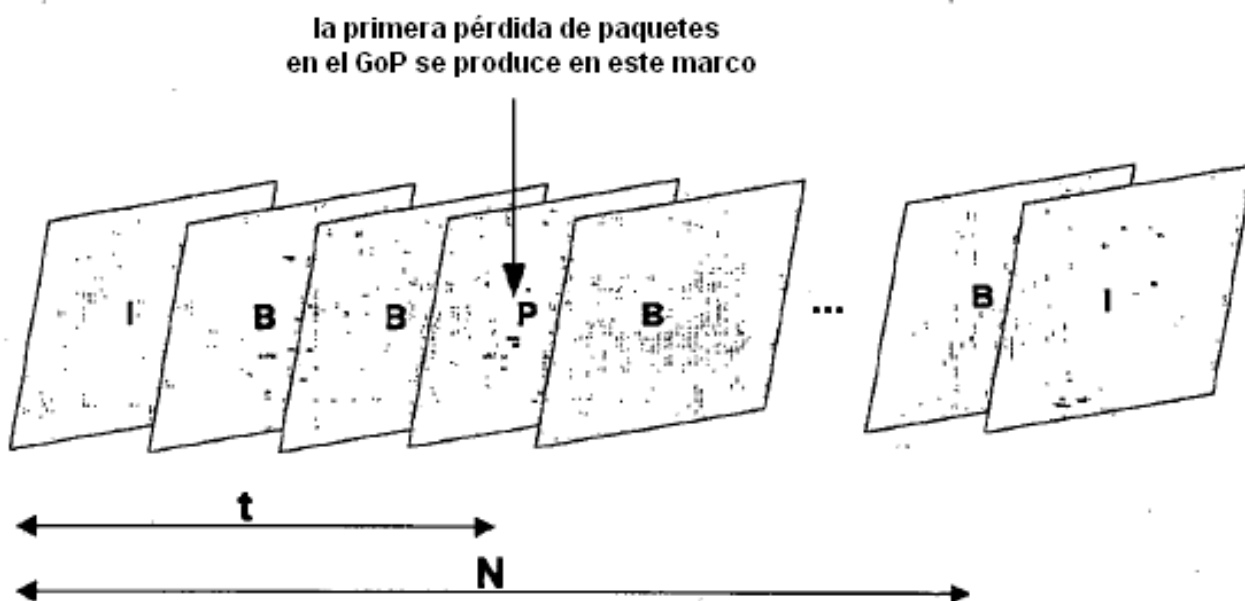


Fig. 2

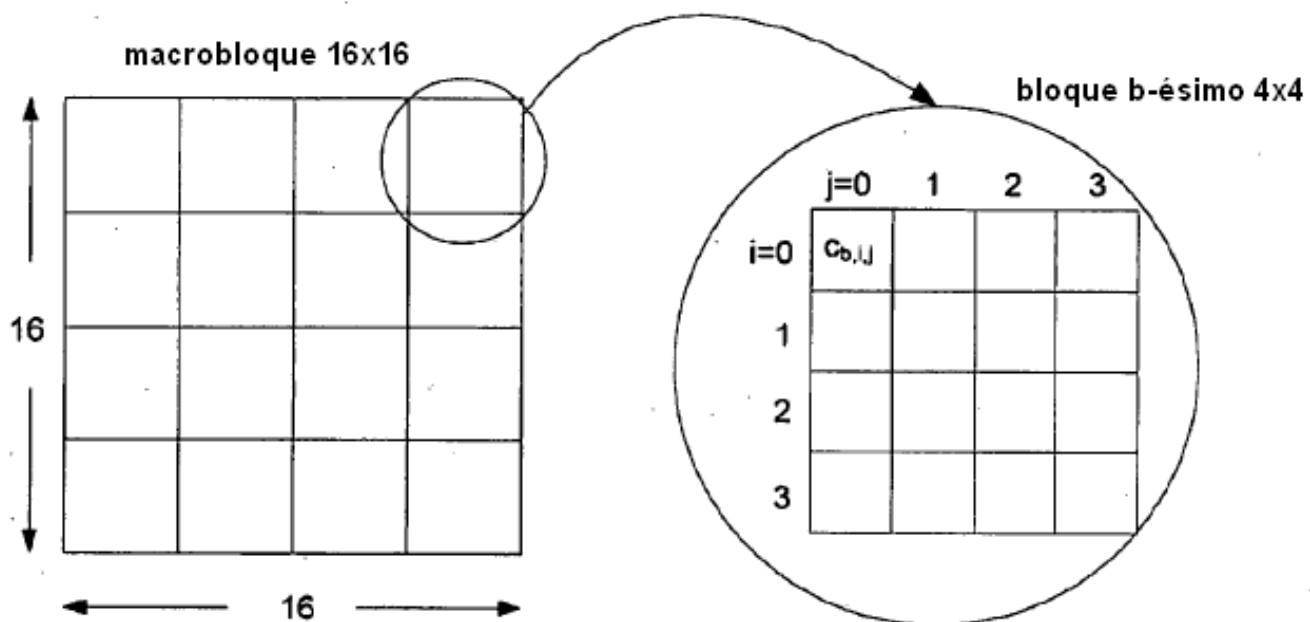


Fig. 3

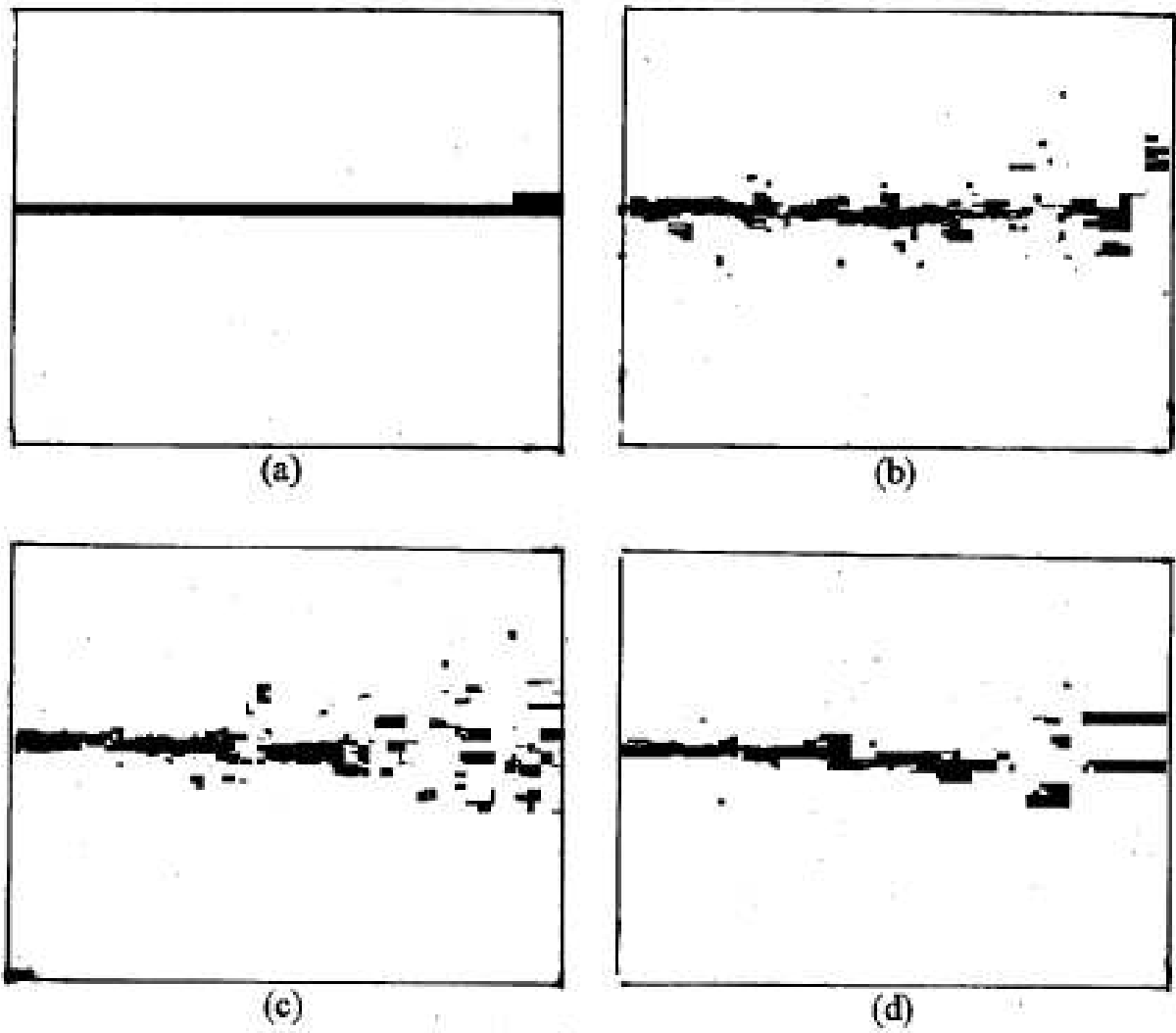


Fig. 4

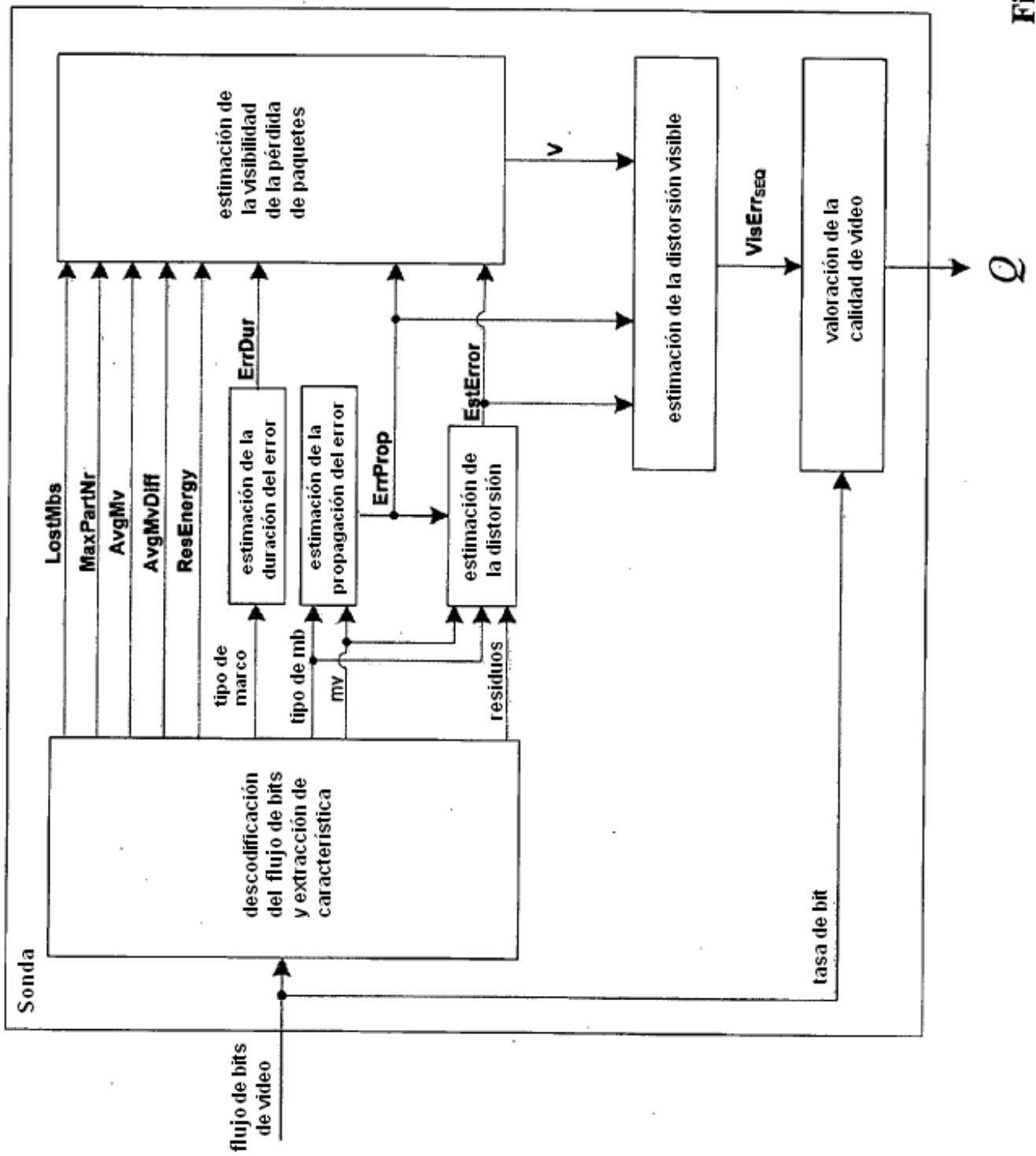


Fig. 5

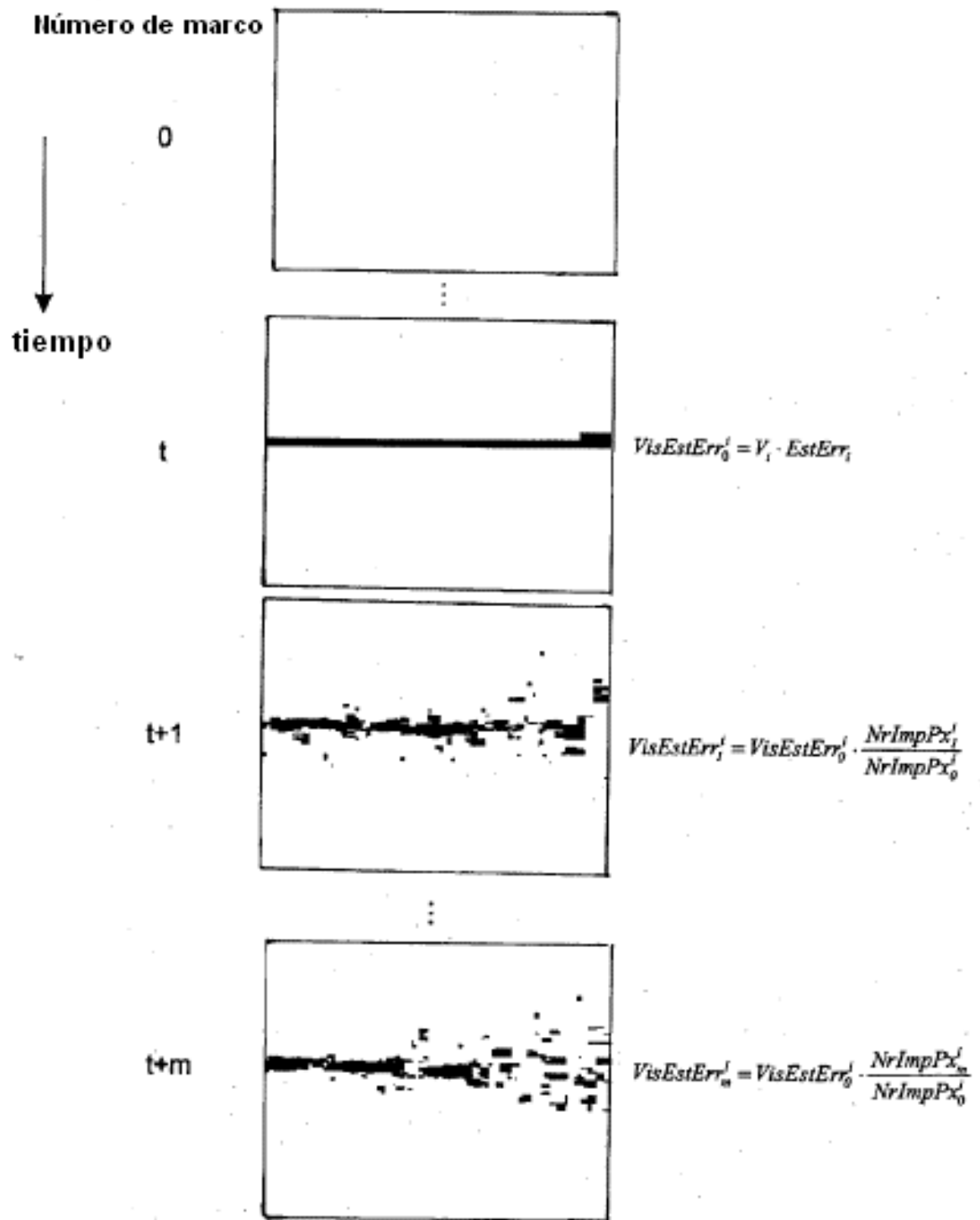


Fig. 6