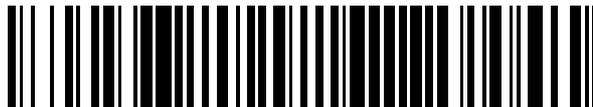


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 622 027**

51 Int. Cl.:

H04N 17/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.08.2013 PCT/CN2013/080882**

87 Fecha y número de publicación internacional: **14.08.2014 WO14121591**

96 Fecha de presentación y número de la solicitud europea: **06.08.2013 E 13874341 (4)**

97 Fecha y número de publicación de la concesión europea: **11.01.2017 EP 2919461**

54 Título: **Método y dispositivo para evaluar calidad de codificación de vídeo**

30 Prioridad:

06.02.2013 CN 201310048015

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

05.07.2017

73 Titular/es:

**HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian
Longgang District , Shenzhen, Guangdong
518129, CN**

72 Inventor/es:

**GAO, SHAN y
SUN, LINA**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 622 027 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 dispositivo para evaluar calidad de codificación de vídeo.

Campo técnico

5 Las realizaciones de la presente invención se refieren al campo de las tecnologías de la comunicación y, más en concreto, a un método y a un dispositivo para evaluar calidad de codificación de vídeo.

Antecedentes

Con el desarrollo de las tecnologías de red, los servicios de vídeo, como ser el vídeo bajo demanda, la televisión web y la videotelefonía, se han transformado en servicios principales de una red de banda ancha, y se volverán servicios principales de una red inalámbrica 3G (la 3ª generación, la tercera generación).

10 Puesto que la cantidad de servicios de vídeo es elevada, la importancia del concepto "tiempo real" es alta, y un usuario es muy sensible a los servicios de vídeo, un operador necesita monitorizar la calidad de los servicios de vídeos transmitidos y adoptar una medida correspondiente de manera oportuna para llevar a cabo un ajuste, con el fin de satisfacer los requerimientos del usuario para experimentar los servicios de vídeo. Diversos factores afectan la calidad de un vídeo como, por ejemplo, la distorsión debido a la compresión, la forma en la que un parámetro en un
15 codificador y descodificador de vídeo se adapta a un canal de transmisión, y la calidad del servicio (como, por ejemplo, límite de ancho de banda, pérdida de paquetes, retraso o inestabilidad horizontal de la imagen) del canal de transmisión.

20 En un método de evaluación de calidad de vídeo objetivo sin referencia existente, un modelo para predecir calidad de codificación (a la calidad de codificación también se la conoce como calidad de referencia o calidad de compresión) sólo tiene en cuenta el impacto de la información de codificación (por ejemplo, una tasa de bits, un Código de formato de codificación, una pausa de vídeo y una pérdida de paquetes de red). Las consideraciones del modelo para predecir calidad de codificación en el método de evaluación de calidad de vídeo objetivo sin referencia (método de evaluación de calidad de vídeo objetivo sin referencia) son incompletas, y el método no puede reflejar de manera precisa los sentimientos subjetivos de la vista humana, y presenta limitaciones en cierta medida. Por lo
25 tanto, la precisión de predecir la calidad de codificación es baja.

La patente europea EP 2252073 describe que se recibe una cadena de bits de un vídeo codificado utilizando una predicción inter-fotograma y DCT con compensación de movimiento y, más en concreto, H.264. Se extrae un parámetro de cuantificación incluido en la cadena de bits recibida, y se calcula su estadística. La calidad subjetiva del vídeo se estima según el valor mínimo del parámetro de cuantificación.

30 La patente europea EP 2296379 describe un método para evaluar la calidad de codificación de vídeo según la complejidad y el parámetro de cuantificación del vídeo.

Compendio

35 Las realizaciones de la presente invención ofrecen un método y un dispositivo para evaluar la calidad de codificación de vídeo, tal y como se define en las reivindicaciones 1 y 13, respectivamente, que pueden mejorar la precisión de predecir la calidad de codificación.

40 De acuerdo con un primer aspecto, se ofrece un método para evaluar la calidad de codificación de vídeo, en el que el método incluye: adquirir un parámetro de cuantificación de un segmento de fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento de fotograma de vídeo de la secuencia de vídeo; determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo, y determinar la complejidad de contenido del vídeo de acuerdo con la cantidad de bytes por píxel del segmento de fotograma de vídeo de la secuencia de vídeo; y predecir la calidad de codificación de vídeo de acuerdo con la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo.

45 Con referencia al primer aspecto, en una primera posible forma de implementación, la implementación específica de la determinación de un parámetro de cuantificación de un vídeo según el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo puede ser: determinar un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación del vídeo.

50 Con referencia a la primera posible forma de implementación del primer aspecto, en un segunda posible forma de implementación, la implementación específica de la determinación de un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación de vídeo puede ser: determinar el parámetro de cuantificación del vídeo de la siguiente manera:

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} i_{\text{segmento_pc}_i}}{N1}$$

o

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} p_i \cdot i_{\text{segmento_pc}_i}}{\sum_{i=1}^{N1} p_i}$$

donde

5 $f_{\text{vídeo_pc}}$ es el parámetro de cuantificación del vídeo, $N1$ es una cantidad de parámetros de cuantificación de segmentos de la secuencia de vídeo, $i_{\text{segmento_pc}_i}$ es un parámetro de cuantificación del $i^{\text{ésimo}}$ segmento entre parámetros de cuantificación de los $N1$ segmentos, $N1$ es un número entero positivo, y p_i es una ponderación correspondiente al $i^{\text{ésimo}}$ segmento.

10 Con referencia al primer aspecto o la primera posible forma de implementación o la segunda posible forma de implementación del primer aspecto, en una tercera posible forma de implementación, la implementación específica de la determinación de la complejidad de contenido del vídeo según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo puede ser: determinar una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm de segmentos del fotograma de vídeo de la siguiente manera:

$$f_{\text{segmento_byte_por_píxel}_k} = \frac{i_{\text{segmento_tamaño}_k}}{i_{\text{segmento_píxel}_k}} \quad \text{donde}$$

15 $i_{\text{segmento_tamaño}_k}$ es una cantidad de bytes del $k^{\text{ésimo}}$ segmento entre el Núm segmentos del fotograma de vídeo, $i_{\text{segmento_píxel}_k}$ es una cantidad de píxeles del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm, y Núm es un número entero positivo;

determinar la complejidad de contenido del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{\text{segmento_contenido_complejidad}_k} = a[i_{\text{segmento_pc}_k}] \cdot f_{\text{segmento_byte_por_píxel}} + b[i_{\text{segmento_pc}_k}]$$

20

donde

tanto $a[i_{\text{segmento_pc}_k}]$ como $b[i_{\text{segmento_pc}_k}]$ son valores correspondientes a un parámetro de cuantificación del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo;

determinar la complejidad de contenido del fotograma de vídeo de la siguiente forma:

$$f_{\text{fotograma_contenido_complejidad}} = \frac{\sum_{k=1}^{Núm} f_{\text{segmento_contenido_complejidad}_k}}{Núm};$$

25

y

determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo.

30 Con referencia a la tercera posible forma de implementación del primer aspecto, en una cuarta posible forma de implementación, la implementación específica de la determinación de complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo puede ser de la siguiente forma:

$$f_{\text{vídeo_contenido_complejidad}} = \frac{\sum_{j=1}^M f_{\text{fotograma_contenido_complejidad}_j}}{M}$$

donde

$f_{\text{vídeo_contenido_complejidad}}$ es la complejidad de contenido del vídeo, la secuencia de vídeo incluye M fotogramas de vídeo, M es un número entero positivo, y j es un número entero positivo y oscila entre 1 y M.

- 5 Con referencia a la cuarta posible forma de implementación del primer aspecto, en una quinta posible forma de implementación, la implementación específica de predicción de la calidad de codificación de vídeo según la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo puede ser de la siguiente forma:

$$d_{\text{compresión_calidad_valor}} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}}$$

donde

$d_{\text{compresión_calidad_valor}}$ es la calidad de codificación de vídeo, y a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 son todas constantes.

- 10 Con referencia a la quinta posible forma de implementación del primer aspecto, en una sexta posible forma de implementación, la implementación específica puede ser que: un valor de a_2 es una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo.

Con referencia a cualquier forma de la quinta posible forma de implementación y la sexta posible forma de implementación del primer aspecto, en una séptima posible forma de implementación, la implementación específica puede ser que: un valor de a_1 es el valor mínimo del intervalo de la calidad de codificación de vídeo.

- 15 Con referencia a cualquier forma de la tercera posible forma de implementación a la séptima posible forma de implementación del primer aspecto, en una octava posible forma de implementación, el método además incluye: ajustar la complejidad de contenido del vídeo de la siguiente forma:

$$f_{\text{vídeo_contenido_complejidad}_n} = \min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right) ; y$$

- 20 usando $f_{\text{vídeo_contenido_complejidad}_n}$ ajustado como la complejidad de contenido del vídeo, donde tanto s como núm1 son todas constantes.

Con referencia a la octava posible forma de implementación del primer aspecto, en una novena posible forma de implementación, la implementación específica puede ser que: un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valor de $f_{\text{vídeo_contenido_complejidad}_n}$ es [0,0, 1,0].

- 25 Con referencia a cualquier forma de la cuarta posible forma de implementación o la novena posible forma de implementación del primer aspecto, en una décima posible forma de implementación, la implementación específica puede ser que: los M fotogramas de vídeo no tienen pérdida de paquetes.

Con referencia a cualquier forma de la cuarta posible forma de implementación o la décima posible forma de implementación del primer aspecto, en una undécima posible forma de implementación, la implementación específica puede ser que: el fotograma de vídeo incluye un fotograma de codificación intra-fotograma y/o un fotograma de codificación inter-fotograma.

- 30

Con referencia a cualquier forma de la tercera posible forma de implementación o la undécima posible forma de implementación del primer aspecto, en una duodécima posible forma de implementación, la implementación específica puede ser que: Núm es una cantidad total de segmentos en el fotograma de vídeo.

5 Con referencia a cualquier forma de la segunda posible forma de implementación o la duodécima posible forma de implementación del primer aspecto, en una decimotercera posible forma de implementación, la implementación específica puede ser que: N1 es una cantidad de parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos, y los parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos indican que los parámetros de cuantificación obtenidos mediante análisis sintáctico y de los segmentos son correctos en un caso en el cual el primer paquete de transmisión hasta un paquete de transmisión en el cual se ubica una diferencia de parámetro de cuantificación de segmentos que se utiliza para adquirir los parámetros de cuantificación de los segmentos, de los segmentos no se pierden.

10 De acuerdo con un segundo aspecto, se ofrece un dispositivo para evaluar la calidad de codificación de vídeo, en el que el dispositivo incluye: una unidad de adquisición, configurada para adquirir un parámetro de cuantificación de un segmento de un fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo; una unidad de determinación, configurada para determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación, adquirido por la unidad de adquisición, del segmento del fotograma de vídeo de la secuencia de vídeo, y determinar la complejidad de contenido del vídeo de acuerdo con la cantidad, adquirida por la unidad de adquisición, de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo; y una unidad de predicción, configurada para predecir la calidad de codificación de vídeo de acuerdo con la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo.

20 Con referencia al segundo aspecto, en una primera posible forma de implementación, la unidad de determinación está específicamente configurada para determinar un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación del vídeo.

25 Con referencia a la primera posible forma de implementación del segundo aspecto, en una segunda posible forma de implementación, la unidad de determinación está específicamente configurada para determinar el parámetro de cuantificación del vídeo de la siguiente manera:

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} i_{\text{segmento_pc}_i}}{N1}, \quad \text{o}$$

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} p_i \cdot i_{\text{segmento_pc}_i}}{\sum_{i=1}^{N1} p_i}, \quad \text{donde}$$

30 $f_{\text{vídeo_pc}}$ es el parámetro de cuantificación del vídeo, N1 es una cantidad de parámetros de cuantificación de segmentos de la secuencia de vídeo, $i_{\text{segmento_pc}_i}$ es un parámetro de cuantificación del $i^{\text{ésimo}}$ segmento entre parámetros de cuantificación de los N1 segmentos, N1 es un número entero positivo, y p_i es una ponderación correspondiente al $i^{\text{ésimo}}$ segmento.

Con referencia al segundo aspecto o a cualquier forma de la primera posible forma de implementación y la segunda posible forma de implementación del segundo aspecto, en una tercera posible forma de implementación, la unidad de determinación está específicamente configurada para:

35 determinar una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{\text{segmento_byte_por_píxel}_k} = \frac{i_{\text{segmento_tamaño}_k}}{i_{\text{segmento_píxel}_k}}, \quad \text{donde}$$

40 $i_{\text{segmento_tamaño}_k}$ es una cantidad de bytes del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, $i_{\text{segmento_píxel}_k}$ es una cantidad de píxeles del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm, y Núm es un número entero positivo;

la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{_segmento_contenido_complejidad}_k = a [i_segmento_pc_k] \cdot f_{_segmento_byte_por_píxel}_k + b [i_segmento_pc_k]$$

donde

- 5 tanto $a[i_segmento_pc_k]$ como $b[i_segmento_pc_k]$ son valores correspondientes a un parámetro de cuantificación del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo;

la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del fotograma de vídeo de la siguiente manera:

$$f_{_fotograma_contenido_complejidad} = \frac{\sum_{k=1}^{Núm} f_{_segmento_contenido_complejidad}_k}{Núm} ;$$

- 10 y

la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo.

- 15 Con referencia a la tercera posible forma de implementación del segundo aspecto, en una cuarta posible forma de implementación, la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del vídeo de la siguiente manera:

$$f_{_vídeo_contenido_complejidad} = \frac{\sum_{j=1}^M f_{_fotograma_contenido_complejidad}_j}{M}$$

donde

la secuencia de vídeo incluye M fotogramas de vídeo, M es un número entero positivo, y j es un número entero positivo y oscila entre 1 y M.

- 20 Con referencia a la cuarta posible forma de implementación del segundo aspecto, en una quinta posible forma de implementación, la unidad de predicción está específicamente configurada para predecir la calidad de codificación de vídeo de la siguiente manera:

$$d_{_compresión_calidad_valor} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{_vídeo_pc}}{a_4 - a_5 \cdot f_{_vídeo_contenido_complejidad}} \right)^{a_6}}$$

, donde

$a_1, a_2, a_3, a_4, a_5,$ y a_6 son todas constantes.

- 25 Con referencia a la quinta posible forma de implementación del segundo aspecto, en una sexta posible forma de implementación, un valor de a_2 es una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo.

- 30 Con referencia a cualquier forma de la quinta posible forma de implementación y la sexta posible forma de implementación del segundo aspecto, en una séptima posible forma de implementación, un valor de a_1 es el valor mínimo del intervalo de la calidad de codificación de vídeo.

Con referencia a cualquier forma de la tercera posible forma de implementación a la séptima posible forma de implementación del segundo aspecto, en una octava posible forma de implementación, el dispositivo además incluye: ajustar la complejidad de contenido del vídeo de la siguiente forma:

$$f_{\text{vídeo_contenido_complejidad}_n} = \min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right) ; y$$

- 5 usando $f_{\text{vídeo_contenido_complejidad}_n}$ ajustado como la complejidad de contenido del vídeo, donde tanto s como núm1 son todas constantes.

Con referencia a la octava posible forma de implementación del segundo aspecto, en una novena posible forma de implementación, un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valor de $f_{\text{vídeo_contenido_complejidad}_n}$ es [0,0, 1,0].

- 10 Con referencia a cualquier forma de la cuarta posible forma de implementación o la novena posible forma de implementación del segundo aspecto, en una décima posible forma de implementación, los M fotogramas de vídeo no tienen pérdida de paquetes.

- 15 Con referencia a cualquier forma de la cuarta posible forma de implementación o la décima posible forma de implementación del segundo aspecto, en una undécima posible forma de implementación, el fotograma de vídeo incluye un fotograma de codificación intra-fotograma y/o un fotograma de codificación inter-fotograma.

Con referencia a cualquier forma de la tercera posible forma de implementación o la undécima posible forma de implementación del segundo aspecto, en una duodécima posible forma de implementación, Núm es una cantidad total de segmentos en el fotograma de vídeo.

- 20 Con referencia a cualquier forma de la segunda posible forma de implementación o la duodécima posible forma de implementación del segundo aspecto, en una decimotercera posible forma de implementación, N1 es la cantidad de parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos, y los parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos indican que los parámetros de cuantificación obtenidos mediante análisis sintáctico y de los segmentos son correctos en un caso en el cual el primer paquete de transmisión hasta un paquete de transmisión en el cual se ubica una diferencia de parámetro de cuantificación de segmentos que se utiliza para adquirir los parámetros de cuantificación de los segmentos, de los segmentos no se pierden.
- 25

- 30 En las realizaciones de la presente invención, un parámetro de cuantificación de un vídeo se determina según un parámetro de cuantificación adquirido de un segmento de un fotograma de vídeo de una secuencia de vídeo, la complejidad de contenido del vídeo se determina según una cantidad adquirida de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo, y la calidad de codificación de vídeo se predice utilizando la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo. Por lo tanto, la calidad de codificación predicha por un modelo que se obtiene considerando una característica de contenido del vídeo satisface mejor los sentimientos subjetivos de la vista humana, mejorando, de esta forma, la precisión de la predicción.

Breve descripción de las figuras

- 35 Para describir las soluciones técnicas en las realizaciones de la presente invención de manera más clara, a continuación se describen brevemente los dibujos que acompañan esta memoria necesarios para describir las realizaciones o la técnica anterior. Aparentemente, los dibujos adjuntos de la siguiente descripción simplemente muestran algunas realizaciones de la presente invención, y una persona con experiencia ordinaria en la técnica puede incluso así obtener otros dibujos a partir de los dibujos adjuntos sin esfuerzos creativos.

- 40 La Fig. 1 es un diagrama de flujo de un método para evaluar calidad de codificación de vídeo según una realización de la presente invención;

La Fig. 2 es un diagrama de bloque estructural de un dispositivo para evaluar calidad de codificación de vídeo según una realización de la presente invención; y

- 45 La Fig. 3 es un diagrama de bloque estructural de un dispositivo para evaluar calidad de codificación de vídeo según otra realización de la presente invención.

Descripción de las realizaciones

A continuación se describen de forma clara y completa las soluciones técnicas en las realizaciones de la presente invención haciendo referencia a los dibujos adjuntos en las realizaciones de la presente invención. Aparentemente, las realizaciones descritas son algunas pero no todas las realizaciones de la presente invención. Todas las demás realizaciones obtenidas por una persona con experiencia ordinaria en la técnica a partir de las realizaciones de la presente invención sin esfuerzos creativos estarán comprendidas dentro del alcance de protección de la presente invención.

Se ha de comprender que las soluciones técnicas de las realizaciones de la presente invención pueden aplicarse a diversos sistemas de comunicaciones como, por ejemplo, un sistema GSM, un sistema CDMA (Multiplexación por División de Código), un sistema WCDMA (Acceso Múltiple por División de Código de Banda Ancha), un sistema GPRS (Servicio General de Paquetes vía Radio), un sistema LTE, un sistema LTE FDD (Duplexación por División de Frecuencia), un sistema LTE TDD (Duplexación por División de Tiempo), un sistema UMTS y un sistema de comunicaciones WiMAX (Interoperabilidad Mundial para Acceso por Microondas).

La Fig. 1 es un diagrama de flujo de un método para evaluar calidad de codificación de vídeo según una realización de la presente invención. El método es ejecutado por un dispositivo para evaluar la calidad de codificación de vídeo. El método incluye las siguientes etapas:

101. Adquirir un parámetro de cuantificación de un segmento de un fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo.

102. Determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo, y determinar la complejidad de contenido del vídeo según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo.

103. Predecir la calidad de codificación de vídeo según la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo.

En esta realización de la presente invención, el parámetro de cuantificación (Parámetro de Cuantificación, PC) también se puede expresar de otra manera, es decir, una etapa de cuantificación (etapa C), para predecir la calidad de codificación de vídeo. Para facilitar la descripción, en esta realización de la presente invención, en un modelo para predecir la calidad de codificación de vídeo, el parámetro de cuantificación se utiliza como un ejemplo que se ha de describir. Se ha de comprender que, una forma de expresión de reemplazar el PC con la etapa C también está comprendida dentro del alcance de la presente invención.

En esta realización de la presente invención, un parámetro de cuantificación de un vídeo se determina según un parámetro de cuantificación adquirido de un segmento de un fotograma de vídeo de una secuencia de vídeo, la complejidad de contenido del vídeo se determina según una cantidad adquirida de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo, y la calidad de codificación de vídeo se predice utilizando la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo. Por lo tanto, la calidad de codificación predicha por un modelo que se obtiene considerando una característica de contenido del vídeo satisface mejor los sentimientos subjetivos de la vista humana, mejorando, de esta forma, la precisión de la predicción.

Opcionalmente, en una realización, en la etapa 102, el parámetro de cuantificación del vídeo puede determinarse según el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo, o el parámetro de cuantificación del vídeo puede determinarse según un parámetro de cuantificación de un macrobloque del fotograma de vídeo de la secuencia de vídeo, o el parámetro de cuantificación del vídeo puede determinarse combinando el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo y el parámetro de cuantificación del macrobloque del fotograma de vídeo de la secuencia de vídeo.

De manera específica, una cantidad total de parámetros de cuantificación de segmentos de la secuencia de vídeo es N, y los parámetros de cuantificación de N1 segmentos pueden seleccionarse de parámetros de cuantificación de N segmentos del vídeo. Preferiblemente, una cantidad de los parámetros de cuantificación de los N1 segmentos puede ser una cantidad total de segmentos de los cuales se obtienen de forma correcta parámetros de cuantificación mediante análisis sintáctico. Utilizando H.264 como ejemplo, los segmentos de los cuales se obtienen parámetros de cuantificación de forma correcta mediante análisis sintáctico indican que en un caso en el que no se pierde un paquete de transmisión en donde se encuentra la cuantificación inicial de imagen menos 26 (imagen_inic_pc_menos26, donde imagen_inic_pc_menos26 es un elemento sintáctico en un conjunto de parámetros de imagen H.264) en un conjunto de parámetros de imagen (imagen_parámetro_conjunto_rbsp, PPS), el primer paquete de transmisión a un paquete de transmisión en donde se encuentra una diferencia de parámetros de cuantificación (segmento_pc_delta, donde segmento_pc_delta es un elemento sintáctico en una cabecera de segmento de H.264) de segmentos utilizada para adquirir los parámetros de cuantificación, de los segmentos no se pierden. En otras palabras, el primer bit a bit que se utiliza para analizar sintácticamente la diferencia de parámetros de cuantificación (segmento_pc_delta, donde segmento_pc_delta es un elemento sintáctico en la cabecera de segmento de H.264) de los segmentos, de los segmentos, no se pierden, donde la diferencia de parámetros de cuantificación de los segmentos se utiliza para adquirir los parámetros de cuantificación de los segmentos. De esta forma, los parámetros de cuantificación de los segmentos se pueden obtener de forma correcta mediante análisis

sintáctico. Una fórmula de cálculo de los parámetros de cuantificación de los segmentos es la siguiente:
 $i_segmento_pc = 26 + imagen_inic_pc_menos26 + segmento_pc_delta$. N es un número entero positivo, N1 es un número entero positivo, y $N1 \leq N$. Asimismo, un valor promedio de los parámetros de cuantificación de los N1 segmentos puede determinarse como el parámetro de cuantificación del vídeo que se utiliza para predecir la calidad de codificación. Un parámetro de cuantificación de un vídeo en la siguiente fórmula (1) es $f_vídeo_pc$:

$$f_vídeo_pc = \frac{\sum_{i=1}^{N1} i_segmento_pc_i}{N1} \quad (1).$$

En la fórmula precedente, $i_segmento_pc_i$ es un parámetro de cuantificación del $i^{ésimo}$ segmento entre los parámetros de cuantificación de los N1 segmentos.

Opcionalmente, el N1 precedente también puede ser una cantidad de parámetros de cuantificación de segmentos cuyos paquetes no se pierden.

Opcionalmente, un valor promedio ponderado de los parámetros de cuantificación de los N1 segmentos también puede determinarse como el parámetro de cuantificación que se utiliza para predecir la calidad de codificación del vídeo, y p_i es una ponderación correspondiente al $i^{ésimo}$ segmento. Es posible asignar distintas ponderaciones a los distintos segmentos según un requisito, tal y como se muestra en la siguiente fórmula:

$$f_vídeo_pc = \frac{\sum_{i=1}^{N1} p_i \cdot i_segmento_pc_i}{\sum_{i=1}^{N1} p_i}$$

Además, es posible realizar un ajuste normalizado en el parámetro de cuantificación $f_vídeo_pc$ del vídeo que se obtiene utilizando las fórmulas precedentes, y se utiliza un parámetro de cuantificación ajustado del vídeo para predecir la calidad de codificación. El ajuste normalizado determina que todos los parámetros de cuantificación estén comprendidos dentro de un intervalo determinado.

Se ha de entender que el ejemplo precedente es meramente un ejemplo, y esta realización de la presente invención no impone limitación alguna sobre cómo obtener un parámetro de cuantificación del vídeo que se utiliza para predecir la calidad de codificación. Un valor promedio o un valor promedio ponderado de todos los parámetros de cuantificación, obtenidos de forma correcta mediante análisis sintáctico, de segmentos de fotogramas de codificación intra-fotograma y fotogramas de codificación inter-fotograma en el vídeo pueden determinarse como el parámetro de cuantificación del vídeo que se utiliza para predecir la calidad de codificación; o un valor promedio o un valor promedio ponderado de todos los parámetros de cuantificación de segmentos y macrobloques en el vídeo pueden determinarse como el parámetro de cuantificación del vídeo que se utiliza para predecir la calidad de codificación; y similares.

Opcionalmente, en un proceso de transmisión de un fotograma de vídeo, también es posible que se pierda un paquete, y es posible realizar el cálculo seleccionando un segmento que tiene pérdida de paquetes o que no tiene pérdida de paquetes. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre esto.

Opcionalmente, el fotograma de vídeo precedente puede ser un fotograma de codificación inter-fotograma y/o un fotograma de codificación intra-fotograma, donde el fotograma de codificación inter-fotograma puede ser un fotograma de codificación inter-fotograma (también conocido como fotograma de referencia inter-fotograma) que sirve como fotograma de referencia, por ejemplo, un fotograma P o un fotograma B, o puede ser un fotograma de codificación inter-fotograma que no sirve como fotograma de referencia, por ejemplo, fotograma b. El fotograma de codificación intra-fotograma puede ser un fotograma I o un fotograma IDR (Refresco de decodificación Instantáneo, por su sigla en inglés), y también puede referirse como un fotograma de referencia intra-fotograma; y similar. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre esto.

Preferiblemente, para determinar la complejidad de contenido del vídeo según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo, la complejidad de contenido del fotograma de vídeo se determina, en primer lugar, según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo, y la complejidad de contenido de vídeo posteriormente se determina según la complejidad de contenido del fotograma de vídeo, específicamente tal y como se muestra en las siguientes fórmulas (2) a (4):

Una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm segmentos del fotograma de vídeo es como sigue:

$$f_{_segmento_byte_por_píxel}_k = \frac{i_{_segmento_tamaño}_k}{i_{_segmento_píxel}_k} \quad (2)$$

, donde

5 $i_{_segmento_tamaño}_k$ es una cantidad de bytes del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, $i_{_segmento_píxel}_k$ es una cantidad de píxeles del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm, y Núm es un número entero positivo.

10 Cabe destacar que, en esta realización de la presente invención, es posible realizar algunas variantes a la cantidad de bytes por píxel, por ejemplo, la cantidad de bytes por píxel puede transformarse en una cantidad de bits por píxel, y la cantidad de bits por píxel se utiliza para realizar una operación correspondiente. Esta variante realiza reemplazos equivalentes a esta realización de la presente invención, sin alejarse del alcance de la presente invención.

La complejidad de contenido del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo es como sigue:

$$f_{_segmento_contenido_complejidad}_k = a [i_{_segmento_pc}_k] \cdot f_{_segmento_byte_por_píxel}_k + b [i_{_segmento_pc}_k] \quad (3)$$

, donde

15 tanto $a [i_{_segmento_pc}_k]$ y $b [i_{_segmento_pc}_k]$ son valores correspondientes a un parámetro de cuantificación del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo.

La complejidad de contenido del fotograma de vídeo es como sigue:

$$f_{_fotograma_contenido_complejidad} = \frac{\sum_{k=1}^{\text{Núm}} f_{_segmento_contenido_complejidad}_k}{\text{Núm}} \quad (4)$$

20 Opcionalmente, Núm puede ser una cantidad (incluida una cantidad de segmentos que tienen una pérdida de paquetes y una cantidad de segmentos que no tienen pérdida de paquetes) de todos los segmentos del fotograma de vídeo; o Núm puede ser una cantidad de segmentos que no tienen pérdida de paquetes y son del fotograma de vídeo, o una cantidad de segmentos que tienen pérdida de paquetes y son del fotograma de vídeo, o puede ser una cantidad de segmentos seleccionados de manera aleatoria de todos los segmentos del fotograma de vídeo; o similar. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre esto.

25 Además, para determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo de la secuencia de vídeo, es posible determinar un valor promedio o un valor promedio ponderado de complejidad de contenido de múltiples fotogramas de vídeo como la complejidad de contenido del vídeo que se utiliza para predecir la calidad de codificación, específicamente tal y como se muestra en la siguiente fórmula (5):

La complejidad de contenido del fotograma de vídeo es como sigue:

$$f_{_vídeo_contenido_complejidad} = \frac{\sum_{j=1}^M f_{_fotograma_contenido_complejidad}_j}{M} \quad (5)$$

, donde

30 la secuencia de vídeo incluye M fotogramas de vídeo, M es un número entero positivo, y j es un número entero positivo y oscila entre 1 y M . Preferiblemente, los M fotogramas de vídeo no tienen pérdida de paquetes.

La complejidad de contenido del vídeo puede incluir una dirección de tiempo y/o una dirección de espacio. En un parámetro de cuantificación fijo de una determinada resolución, una mayor complejidad de contenido del vídeo indica que es menos probable que la vista humana detecte distorsión. La complejidad de contenido del vídeo, obtenida mediante las soluciones técnicas precedentes, satisface mejor los sentimientos subjetivos de la vista humana de lo cual una mayor complejidad de contenido del vídeo indica que es menos probable que la vista humana detecte distorsión, produciéndose así una mejora adicional en términos de precisión a la hora de predecir la calidad de codificación.

Por ejemplo, el fotograma de vídeo en las fórmulas precedentes (2) a (5) es un fotograma de referencia intra-fotograma, es decir, la complejidad de contenido del vídeo puede determinarse mediante una cantidad de bytes por píxel de un segmento del fotograma de referencia intra-fotograma (al menos se utiliza un fotograma de referencia intra-fotograma). En este ejemplo, la complejidad de contenido del vídeo que se determina mediante la cantidad de bytes por píxel del segmento del fotograma de referencia intra-fotograma puede reflejar de manera eficaz la complejidad espacial $f_{\text{vídeo_contenido_complejidad_scc}}$ del vídeo. Opcionalmente, $a[i_segmento_pc_k]$ y $b[i_segmento_pc_k]$ en las fórmulas pueden obtenerse mediante práctica, y distintos formatos de codificación y/o resolución de vídeo y/o formatos de colección de vídeo (es decir, una forma de escaneo entrelazada y una forma de escaneo progresiva) pueden utilizar un mismo valor o valores diferentes. El hecho de que H.264 tenga 52 parámetros de cuantificación (que son 0 a 51 respectivamente) sirve como ejemplo. En consecuencia, $a[i_segmento_pc_k]$ y $b[i_segmento_pc_k]$ en distinta resolución, obtenidos mediante prácticas corresponden, de forma separada, a valores de parámetros de cuantificación de 0 a 51, y se muestran a continuación a modo de ejemplo:

$a[i_segmento_pc_k]$,

resolución de definición estándar:

{24,78954, 24,78954, 25,23854, 25,51193, 25,74990, 25,97533, 26,19479, 26,28303, 26,49158, 26,56645, 26,53197, 26,62563, 26,69239, 26,65409, 26,79309, 26,80578, 26,84816, 27,08741, 27,25370, 27,36097, 27,56078, 27,70162, 27,85621, 28,04059, 28,17621, 28,23445, 28,41471, 28,45078, 28,54265, 28,60014, 28,62930, 28,64529, 28,74102, 28,75523, 28,76358, 28,74681, 28,77488, 28,73642, 28,79531, 28,69430, 28,72766, 28,60666, 28,49484, 28,35642, 28,07614, 27,90134, 27,57123, 27,01405, 26,65987, 26,31439, 25,52575, 25,01169}

resolución de 1280 x 720:

{16,17209, 17,45819, 17,80732, 18,02041, 18,18083, 18,52479, 19,03342, 19,06581, 19,41564, 19,85189, 20,07956, 20,81183, 21,43127, 21,83287, 22,61658, 23,14807, 23,92571, 25,20184, 26,03683, 26,68701, 27,49974, 28,12203, 28,66205, 29,27020, 29,69070, 29,92960, 30,40275, 30,60385, 30,85636, 31,06785, 31,26051, 31,35589, 31,63646, 31,76881, 31,92259, 32,08798, 32,28134, 32,36179, 32,60119, 32,61653, 32,75291, 32,73418, 32,72940, 32,70158, 32,59009, 32,41000, 32,21505, 31,76353, 31,23468, 30,87401, 30,01071, 29,31316}

resolución de alta definición completa (1920 x 1080):

{15,75673, 16,17239, 17,33657, 18,09218, 18,78856, 19,85244, 20,94081, 21,42377, 25,25608, 25,36929, 25,37671, 25,59413, 25,77414, 25,89431, 26,16539, 26,37098, 26,71202, 27,45373, 27,99336, 28,43923, 29,01115, 29,49924, 29,89337, 30,32379, 30,59313, 30,74944, 31,01314, 31,10389, 31,21737, 31,28295, 31,38585, 31,36863, 31,44693, 31,40169, 31,43938, 31,39075, 31,36072, 31,33672, 31,26816, 31,16160, 31,03165, 30,80631, 30,57609, 30,36353, 30,06076, 29,62381, 29,37353, 29,05716, 28,60942, 28,52338, 28,40104, 28,52280}

$b[i_segmento_pc_k]$,

resolución de definición estándar:

{13,39250, 13,39250, 13,97091, 14,53803, 15,25528, 16,13630, 16,99497, 17,66163, 18,80068, 19,89785, 21,20091, 22,86877, 24,44105, 25,98037, 28,04957, 30,07985, 32,07935, 34,30203, 36,32256, 38,18652, 40,93258, 43,77054, 46,53546, 50,53632, 54,36178, 57,82423, 63,29899, 69,18878, 75,07466, 83,80263, 91,47496, 99,18949, 111,47580, 124,34650, 136,49900, 156,17670, 176,23080, 192,16970, 223,83720, 251,77270, 285,92790, 333,53770, 388,41820, 435,09860, 531,05070, 633,24080, 760,16820, 948,15240, 1168,53720, 1361,84570, 1759,43160, 2040,35460}

resolución de 1280 x 720:

{33,81798, 33,05324, 35,11725, 36,95499, 39,10951, 41,62373, 43,87256, 45,95354, 49,32386, 51,87803, 54,92251, 58,42482, 61,62755, 64,56505, 69,19412, 73,35919, 76,10406, 78,96517, 81,95586, 84,59924, 89,05335, 93,59975, 98,31476, 105,41810, 112,34964, 118,73374, 129,00992, 140,01562, 151,12381, 167,62430, 182,02425, 196,08347, 218,72591, 241,16108, 263,35157, 295,99927, 329,06899, 355,66280, 407,64235, 452,09915, 508,72302, 585,36672, 671,43978, 741,49561, 891,18944, 1051,86892, 1246,04333, 1527,50615, 1894,63282, 2204,87735, 2879,95903, 3390,89788}

resolución de alta definición completa (1920 x 1080):

{25,92973, 26,42403, 26,72231, 27,10874, 27,55908, 27,59167, 27,40409, 27,63129, 21,08740, 22,32786, 23,78112, 25,55635, 27,25511, 28,80079, 31,33600, 33,71534, 35,51380, 37,14249, 38,57997, 39,75292, 41,50986, 43,25411, 45,08496, 47,92251, 50,97660, 53,82247, 58,50549, 64,00109, 69,59487, 78,31654, 84,35147, 92,89916, 105,12040, 119,83478, 131,13182, 152,46046, 175,28796, 191,40711, 231,17849, 262,14953, 311,33306, 374,98524, 454,98602, 524,68907, 656,91124, 830,55605, 990,09180, 1196,94617, 1493,32352, 1667,34794, 1966,34090, 2099,62991}

Según un ejemplo adicional, el fotograma de vídeo en las fórmulas precedentes (2) a (5) es un fotograma de referencia inter-fotograma, es decir, la complejidad de contenido del vídeo puede determinarse mediante una cantidad de bytes por píxel de un segmento del fotograma de referencia inter-fotograma (al menos se utiliza un fotograma de referencia inter-fotograma). En este ejemplo, la complejidad de contenido del vídeo que se determina mediante la cantidad de bytes por píxel del segmento del fotograma de referencia inter-fotograma puede reflejar de manera eficaz la complejidad temporal $f_{\text{vídeo_contenido_complejidad_tcc}}$ del vídeo. Opcionalmente, $a_{[i_segmento_pc_k]}$ y $b_{[i_segmento_pc_k]}$ en las fórmulas pueden obtenerse mediante prácticas, y distintos formatos de codificación y/o resolución de vídeo y/o formatos de colección de vídeo (es decir, una forma de escaneo entrelazada y una forma de escaneo progresiva) pueden utilizar un mismo valor o valores diferentes. El hecho de que H.264 tenga 52 parámetros de cuantificación (que son 0 a 51 respectivamente) sirve como ejemplo. En consecuencia, $a_{[i_segmento_pc_k]}$ y $b_{[i_segmento_pc_k]}$ en distinta resolución, obtenidos mediante prácticas corresponden, de forma separada, a valores de parámetros de cuantificación de 0 a 51.

Se ha de comprender que, en los ejemplos precedentes, la selección del fotograma de vídeo, la cantidad de parámetros de cuantificación de segmentos, los valores de $a_{[i_segmento_pc_k]}$ y $b_{[i_segmento_pc_k]}$, y una forma de obtener $a_{[i_segmento_pc_k]}$ y $b_{[i_segmento_pc_k]}$ son meramente ejemplos pero no se pretende con ellos limitar el alcance de la presente invención.

Se halló en un experimento que, cuando existe distorsión de compresión de codificación en el vídeo, la calidad de codificación de vídeo disminuye a la vez que el PC aumenta. Opcionalmente, según el experimento precedente, en la etapa 102, es posible predecir la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ según la complejidad de contenido $f_{\text{vídeo_contenido_complejidad}}$ del vídeo en la fórmula precedente (5) y el parámetro de cuantificación $f_{\text{vídeo_pc}}$ del vídeo en la fórmula precedente (1), tal y como se muestra en la siguiente fórmula (6):

La calidad de codificación de vídeo se predice de la siguiente forma:

$$d_{\text{compresión_calidad_valor}} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}} \quad (6)$$

donde

a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 son todas constantes.

Opcionalmente, en distintos formatos de codificación y/o resolución de vídeo y/o formatos de colección de vídeo (es decir, la forma de escaneo entrelazada y la forma de escaneo progresiva), los valores de a_1 a a_6 pueden ser iguales o diferentes. Los valores de a_1 a a_6 en resolución de definición estándar pueden ser iguales o diferentes respecto de los valores de a_1 a a_6 en resolución de alta definición completa, y se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

Un valor superior de la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ indica mejor calidad de vídeo, y un valor inferior de la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ indica peor calidad de vídeo.

Opcionalmente, un valor de a_2 puede ser una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo. Por ejemplo, en una escala de medición de calidad de codificación de 5 puntos, un valor máximo de la calidad de codificación es 5, y un valor mínimo es 1; por lo tanto, el valor de a_2 es 4. Según otro ejemplo, el valor de a_2 puede ser una diferencia, obtenida mediante prácticas o utilizando medios empíricos, entre el valor máximo y el valor mínimo del intervalo de calidad de codificación de vídeo. Por ejemplo, un valor máximo de la calidad de codificación que se obtiene mediante prácticas o utilizando medios empíricos es 4,5, y el valor mínimo de la calidad de codificación es 1,5; por lo tanto, el valor de a_2 es 3. Se ha de comprender que, en esta realización de la presente invención, una forma de predecir la calidad de codificación puede aplicarse a una escala de medición de calidad de codificación con diferente puntaje, y esta realización de la presente invención no impone limitación alguna sobre ello.

Opcionalmente, un valor de a_1 puede ser el valor mínimo del intervalo de la calidad de codificación de vídeo.

Opcionalmente, también es posible obtener valores de a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 según distintos casos, y ha de comprenderse que esta realización de la presente invención no impone limitación alguna sobre ello. Por ejemplo, también es posible obtener valores de a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 mediante prácticas o utilizando medios empíricos según distintos casos.

5

Opcionalmente, se pueden mostrar a modo de ejemplo, a continuación, valores de a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 :

definición estándar:

$$a_1 = 1,4163,$$

$$a_2 = 2,9116,$$

10 $a_3 = 1,0,$

$$a_4 = 41,5,$$

$$a_5 = 4,7,$$

$$a_6 = 13,0$$

1280 x 720:

15 $a_1 = 1,0519,$

$$a_2 = 3,3876,$$

$$a_3 = 1,0,$$

$$a_4 = 40,0,$$

$$a_5 = 0,75,$$

20 $a_6 = 10,0$

1920 x 1080, escaneo entrelazado (1920 x 1080i):

$$a_1 = 1,2294,$$

$$a_2 = 3,1092,$$

$$a_3 = 1,0,$$

25 $a_4 = 41,5,$

$$a_5 = 0,65,$$

$$a_6 = 10,5$$

1920 x 1080, escaneo progresivo (1920 x 1080p)

$$a_1 = 1,2294,$$

30 $a_2 = 3,1092,$

$$a_3 = 1,0,$$

$$a_4 = 43,0,$$

$$a_5 = 0,85,$$

$$a_6 = 12,0$$

35 Cabe destacar que una forma de utilizar otra fórmula equivalente para expresar la calidad de codificación precedente está comprendida dentro del alcance de la presente invención. Por ejemplo, una calidad de distorsión de codificación de vídeo $d_{\text{compresión_artefacto_valor}}$ se obtiene sustrayendo la calidad de codificación de vídeo precedente del valor máximo de la calidad de codificación de vídeo, tal y como se muestra en la siguiente fórmula (7):

$$d_{\text{compresión_artefacto_valor}} = \frac{a_2 \cdot \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}} \quad (7)$$

En este caso, una calidad de distorsión de codificación de vídeo inferior indica una mejor calidad de vídeo, y una calidad de distorsión de codificación de vídeo superior indica una peor calidad de vídeo.

- 5 Opcionalmente, es posible ajustar la complejidad de contenido del vídeo en las fórmulas precedentes (5) a (7), y puede utilizarse $f_{\text{vídeo_contenido_complejidad_n}}$ ajustada como la complejidad de contenido del vídeo, donde una forma de ajuste se muestra en la siguiente fórmula (8):

$$f_{\text{vídeo_contenido_complejidad_n}} = \min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right) \quad (8), \text{ donde}$$

s y núm1 son constantes. Por ejemplo, cuando se lleva a cabo un ajuste normalizado, un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valores de $f_{\text{vídeo_contenido_complejidad_n}}$ es [0,0, 1,0].

- 10 Se ha de comprender que la complejidad de contenido del vídeo puede ser complejidad temporal (por ejemplo, solo se considera el fotograma de referencia inter-fotograma para el fotograma de vídeo), y se expresa como $f_{\text{vídeo_contenido_complejidad}}=f_{\text{vídeo_contenido_complejidad_tcc}}$ (en la fórmula precedente (5)) o $f_{\text{vídeo_contenido_complejidad_n}}=f_{\text{vídeo_contenido_complejidad_ntcc}}$ (complejidad temporal ajustada en la fórmula precedente (8)); o puede ser complejidad espacial (por ejemplo, solo se considera el fotograma de referencia intra-fotograma para el fotograma de vídeo), y expresado como $f_{\text{vídeo_contenido_complejidad}}=f_{\text{vídeo_contenido_complejidad_scc}}$ (en la fórmula precedente (5)) o $f_{\text{vídeo_contenido_complejidad_n}}=f_{\text{vídeo_contenido_complejidad_nscc}}$ (complejidad espacial ajustada en la fórmula precedente (8)); o puede ser una combinación de la complejidad temporal y de la complejidad espacial, y expresada como $f_{\text{vídeo_contenido_complejidad}}=\text{func1}(f_{\text{vídeo_contenido_complejidad_tcc}}, f_{\text{vídeo_contenido_complejidad_scc}})$ o $f_{\text{vídeo_contenido_complejidad_n}}=\text{func2}(f_{\text{vídeo_contenido_complejidad_ntcc}}, f_{\text{vídeo_contenido_complejidad_nscc}})$. Esta realización de la presente invención no impone limitación alguna sobre ello.

- 25 Cabe destacar que, en una realización que utiliza la fórmula (6) o la fórmula (7) para predecir la calidad de codificación o la calidad de distorsión de codificación de vídeo, es posible utilizar la complejidad de contenido del vídeo obtenida ajustando la fórmula (8), es decir:

Es posible predecir la calidad de codificación de vídeo de la siguiente forma:

$$d_{\text{compresión_calidad_valor}} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad_n}}} \right)^{a_6}} \quad (9)$$

donde

- 30 $a_1, a_2, a_3, a_4, a_5,$ y a_6 son todas constantes.

También es posible predecir la calidad de codificación de vídeo de la siguiente forma:

$$d_compresión_artefacto_valor = \frac{a_2 \cdot \left(\frac{f_vídeo_pc}{a_4 - a_5 \cdot f_vídeo_contenido_complejidad_n} \right)^{a_6}}{a_3 + \left(\frac{f_vídeo_pc}{a_4 - a_5 \cdot f_vídeo_contenido_complejidad_n} \right)^{a_6}} \quad (10).$$

Se ha de comprender que la complejidad ajustada de contenido del vídeo $f_vídeo_contenido_complejidad_n$ puede ser complejidad temporal del vídeo $f_vídeo_contenido_complejidad_ntcc$ que se obtiene después de un ajuste normalizado, puede ser complejidad espacial del vídeo $f_vídeo_contenido_complejidad_nsc$ que se obtiene después de un ajuste normalizado, o puede ser una combinación de complejidad temporal $f_vídeo_contenido_complejidad_ntcc$ y complejidad espacial $f_vídeo_contenido_complejidad_nsc$. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

Por ejemplo, la calidad de codificación de vídeo $d_compresión_calidad_valor$ se predice mediante la combinación de complejidad temporal $f_vídeo_contenido_complejidad_ntcc$ y la complejidad espacial $f_vídeo_contenido_complejidad_nsc$, y los parámetros de cuantificación $f_vídeo_pc$ del vídeo, es decir, en la fórmula (6), la complejidad de contenido del vídeo utiliza la complejidad ajustada de contenido del vídeo $f_vídeo_contenido_complejidad_n$, y $a_5 \cdot f_vídeo_contenido_complejidad_n$ puede ser reemplazada por $a_{51} \cdot f_vídeo_contenido_complejidad_ntcc + a_{52} \cdot f_vídeo_contenido_complejidad_nsc$, donde a_{51} y a_{52} pueden ser iguales o diferentes.

Una forma de predecir calidad de codificación según esta realización de la presente invención satisface mejor los sentimientos subjetivos de la vista humana, con lo cual mejora la precisión al predecir la calidad de codificación.

Ciertamente, la complejidad de contenido del vídeo puede determinarse según otra información de codificación del fotograma de vídeo, por ejemplo, una cantidad de bytes de codificación de segmentos del fotograma de vídeo, una diferencia de vector de movimiento, un valor de vector de movimiento, un modo de codificación de macrobloque, un coeficiente DCT (transformada de coseno discreta). Por ejemplo, la complejidad de contenido del vídeo se expresa como un valor promedio de coeficientes de CA (corriente alterna) de coeficientes de transformada de coseno discreta DCT de todos los fotogramas de codificación intra-fotograma (I-fotograma/IDR fotograma), un valor promedio de coeficientes de CD (corriente directa) de coeficientes de transformada de coseno discreta DCT de todos los fotogramas de codificación intra-fotograma (I-fotograma/IDR fotograma), un valor promedio de coeficientes de transformada de coseno discreta DCT de todos los fotogramas de codificación intra-fotograma (I-fotograma/IDR fotograma), un valor promedio de coeficientes de CA de coeficientes de transformada de coseno discreta DCT de fotogramas de codificación inter-fotograma, un valor promedio de coeficientes de CD de coeficientes de transformada de coseno discreta DCT de fotogramas de codificación inter-fotograma, un valor promedio de coeficientes de transformada de coseno discreta DCT de fotogramas de codificación inter-fotograma, o un valor promedio de coeficientes de transformada de coseno discreta DCT de fotogramas de codificación intra-fotograma y fotogramas de codificación inter-fotograma. Según otro ejemplo, la complejidad de contenido del vídeo se describe mediante un modo de codificación de macrobloque (por ejemplo, 16 x 16, 16 x 8, 8 x 16), donde más modos de codificación de macrobloque indican mayor complejidad de contenido del vídeo. Según un ejemplo adicional, la complejidad de contenido del vídeo se describe mediante un vector de movimiento, donde un mayor vector de movimiento indica mayor complejidad de contenido del vídeo. Según otro ejemplo adicional, la complejidad de contenido del vídeo se expresa mediante una proporción de bloques de codificación inter-fotograma respecto de bloques de codificación intra-fotograma, donde una mayor proporción indica mayor complejidad de contenido del vídeo. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

La Fig. 2 es un diagrama de bloque estructural de un dispositivo para evaluar calidad de codificación de vídeo según una realización de la presente invención. El dispositivo 200 para evaluar calidad de codificación de vídeo incluye una unidad de adquisición 201, una unidad de determinación 202, y una unidad de predicción 203.

La unidad de adquisición 201 está configurada para adquirir un parámetro de cuantificación de un segmento de un fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo.

La unidad de determinación 202 está configurada para determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación, adquirido por la unidad de adquisición 201, del segmento del fotograma de vídeo de la secuencia de vídeo, y para determinar la complejidad de contenido del vídeo según la cantidad, adquirida por la unidad de adquisición 201, de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo.

La unidad de predicción 203 está configurada para predecir la calidad de codificación de vídeo según la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo determinados mediante la unidad de determinación 202.

5 En esta realización de la presente invención, un parámetro de cuantificación también se puede expresar de otra manera, es decir, una etapa de cuantificación (etapa C), para predecir la calidad de codificación de vídeo. Para facilitar la descripción, en esta realización de la presente invención, en un modelo para predecir la calidad de codificación de vídeo, el parámetro de cuantificación se utiliza como un ejemplo que se ha de describir. Se ha de comprender que, una forma de expresión de reemplazar el PC con la etapa C también está comprendida dentro del alcance de la presente invención.

10 En esta realización de la presente invención, un parámetro de cuantificación de un vídeo se determina según un parámetro de cuantificación adquirido de un segmento de un fotograma de vídeo de una secuencia de vídeo, la complejidad de contenido del vídeo se determina según una cantidad adquirida de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo, y la calidad de codificación de vídeo se predice utilizando la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo. Por lo tanto, la calidad de codificación predicha por un modelo que se obtiene considerando una característica de contenido del vídeo satisface mejor los sentimientos subjetivos de la vista humana, mejorando, de esta forma, la precisión de la predicción.

El dispositivo 200 puede implementar etapas del dispositivo para evaluar calidad de codificación de vídeo que están incluidas en el método representado en la Fig. 1, y con el fin de evitar repeticiones, no se vuelven a describir detalles en la presente memoria.

20 Opcionalmente, en una realización, la unidad de determinación 202 está específicamente configurada para determinar un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación del vídeo. Específicamente, la unidad de determinación 202 puede configurarse para:

determinar $f_{\text{vídeo_pc}}$ como el parámetro de cuantificación del vídeo:

25
$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} i_{\text{segmento_pc}_i}}{N1},$$

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} p_i \cdot i_{\text{segmento_pc}_i}}{\sum_{i=1}^{N1} p_i},$$

donde

$N1$ es una cantidad de parámetros de cuantificación de segmentos de la secuencia de vídeo, $i_{\text{segmento_pc}_i}$ es un parámetro de cuantificación del $i^{\text{ésimo}}$ segmento entre parámetros de cuantificación de los $N1$ segmentos, $N1$ es un número entero positivo, y p_i es una ponderación correspondiente al $i^{\text{ésimo}}$ segmento.

30 Preferiblemente, $N1$ es una cantidad total de parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos, y los parámetros de cuantificación, obtenidos correctamente mediante análisis sintáctico, de segmentos indican que los parámetros de cuantificación obtenidos mediante análisis sintáctico y de los segmentos son correctos en un caso en el cual el primer paquete de transmisión a un paquete de transmisión en el cual se ubica una diferencia de parámetro de cuantificación de segmentos que se utiliza para adquirir los parámetros de cuantificación de los segmentos, de los segmentos no se pierden.

Además, la unidad de determinación 202 puede adicionalmente realizar un ajuste normalizado en el parámetro de cuantificación $f_{\text{vídeo_pc}}$ del vídeo que se obtiene utilizando la fórmula precedente, y utilizar un parámetro de cuantificación ajustado del vídeo para predecir la calidad de codificación.

Opcionalmente, la unidad de determinación 202 puede configurarse, de manera específica, para:

40 determinar una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_segmento_byte_por_píxel_k = \frac{i_segmento_tamaño_k}{i_segmento_píxel_k}$$

, donde

5 $i_segmento_tamaño_k$ es una cantidad de bytes del $k^{ésimo}$ segmento entre los Núm segmentos del fotograma de vídeo, $i_segmento_píxel_k$ es una cantidad de píxeles del $k^{ésimo}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm, y Núm es un número entero positivo.

La unidad de determinación 202 está específicamente configurada para determinar la complejidad de contenido del $k^{ésimo}$ segmento entre los Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_segmento_contenido_complejidad_k = a [i_segmento_pc_k] \cdot f_segmento_byte_por_píxel_k + b [i_segmento_pc_k]$$

, donde

10 tanto $a[i_segmento_pc_k]$ como $b[i_segmento_pc_k]$ son valores correspondientes a un parámetro de cuantificación del $k^{ésimo}$ segmento entre los Núm segmentos del fotograma de vídeo.

La unidad de determinación 202 está específicamente configurada para determinar la complejidad de contenido del fotograma de vídeo de la siguiente manera:

$$f_fotograma_contenido_complejidad = \frac{\sum_{k=1}^{Núm} f_segmento_contenido_complejidad_k}{Núm}$$

15 La unidad de determinación 202 está específicamente configurada para determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo.

Opcionalmente, el fotograma de vídeo precedente puede ser un fotograma de codificación inter-fotograma y/o un fotograma de codificación intra-fotograma, donde el fotograma de codificación inter-fotograma puede ser un fotograma de codificación inter-fotograma (también conocido como fotograma de referencia inter-fotograma) que sirve como fotograma de referencia, por ejemplo, un fotograma P o un fotograma B, o puede ser un fotograma de codificación inter-fotograma que no sirve como fotograma de referencia, por ejemplo, fotograma b; el fotograma de codificación intra-fotograma puede ser un I-fotograma o un IDR (Refresco de Descodificación Instantánea)-fotograma, y también es conocido como fotograma de referencia intra-fotograma; y fotogramas similares. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

25 Además, la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del vídeo de la siguiente manera:

$$f_vídeo_contenido_complejidad = \frac{\sum_{j=1}^M f_fotograma_contenido_complejidad_j}{M}$$

, donde

la secuencia de vídeo incluye M fotogramas de vídeo, M es un número entero positivo, y j es un número entero positivo y oscila entre 1 y M.

30 Opcionalmente, Núm puede ser una cantidad (incluida una cantidad de segmentos que tienen una pérdida de paquetes y una cantidad de segmentos que no tienen pérdida de paquetes) de todos los segmentos del fotograma de vídeo; o Núm puede ser una cantidad de segmentos que no tienen pérdida de paquetes y son del fotograma de vídeo, o una cantidad de segmentos que tienen pérdida de paquetes y son del fotograma de vídeo, o puede ser una cantidad de segmentos seleccionados de manera aleatoria de todos los segmentos del fotograma de vídeo; o similar. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

35 Opcionalmente, los M fotogramas de vídeo no tienen pérdida de paquetes.

Opcionalmente, el fotograma de vídeo puede incluir un fotograma de codificación intra-fotograma y/o un fotograma de codificación inter-fotograma.

La complejidad de contenido del vídeo puede incluir una dirección de tiempo y/o una dirección de espacio. En un parámetro de cuantificación fijo de una determinada resolución, una mayor complejidad de contenido del vídeo indica que es menos probable que la vista humana detecte distorsión. La complejidad de contenido del vídeo, obtenida mediante las soluciones técnicas precedentes, satisface mejor los sentimientos subjetivos de la vista humana de lo cual una mayor complejidad de contenido del vídeo indica que es menos probable que la vista humana detecte distorsión, produciéndose así una mejora adicional en términos de precisión a la hora de predecir la calidad de codificación.

Opcionalmente, la unidad de predicción 203 puede estar específicamente configurada para: predecir la calidad de codificación de vídeo de la siguiente manera:

$$d_{\text{compresión_calidad_valor}} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}}$$

, donde

a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 son todas constantes.

Un valor superior de la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ indica mejor calidad de vídeo, y un valor inferior de la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ indica peor calidad de vídeo.

Opcionalmente, un valor de a_2 puede ser una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo. Por ejemplo, en una escala de medición de calidad de codificación de 5 puntos, un valor máximo de la calidad de codificación es 5, y un valor mínimo es 1; por lo tanto, el valor de a_2 es 4. Según otro ejemplo, el valor de a_2 puede ser una diferencia, obtenida mediante prácticas o utilizando medios empíricos, entre el valor máximo y el valor mínimo del intervalo de calidad de codificación de vídeo. Por ejemplo, un valor máximo de la calidad de codificación que se obtiene mediante prácticas o utilizando medios empíricos es 4,5, y el valor mínimo de la calidad de codificación es 1,5; por lo tanto, el valor de a_2 es 3. Se ha de comprender que, en esta realización de la presente invención, una forma de predecir la calidad de codificación puede aplicarse a una escala de medición de calidad de codificación con diferente puntaje, y esta realización de la presente invención no impone limitación alguna sobre ello.

Opcionalmente, un valor de a_1 puede ser el valor mínimo del intervalo de la calidad de codificación de vídeo.

Opcionalmente, también es posible obtener valores de a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 según distintos casos, y ha de comprenderse que esta realización de la presente invención no impone limitación alguna sobre ello. Por ejemplo, también es posible obtener valores de a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 mediante prácticas o utilizando medios empíricos según distintos casos.

Cabe destacar que una forma de utilizar otra fórmula equivalente para expresar la calidad de codificación precedente está comprendida dentro del alcance de la presente invención. Por ejemplo, una calidad de distorsión de codificación de vídeo $d_{\text{compresión_artefacto_valor}}$ se obtiene sustrayendo la calidad de codificación de vídeo precedente del valor máximo de la calidad de codificación de vídeo, tal y como se muestra en la fórmula precedente (7). En este caso, una calidad de distorsión de codificación de vídeo inferior indica una mejor calidad de vídeo, y una calidad de distorsión de codificación de vídeo superior indica una peor calidad de vídeo.

Opcionalmente, la unidad de determinación 202 puede estar específicamente configurada para ajustar la complejidad de contenido del vídeo de la siguiente manera:

$$f_{\text{vídeo_contenido_complejidad_n}} = \min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right)$$

, donde

s y núm1 son constantes. Por ejemplo, cuando se lleva a cabo un ajuste normalizado, un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valores de $f_{\text{vídeo_contenido_complejidad_n}}$ es [0,0, 1,0].

Se ha de comprender que la complejidad de contenido del vídeo puede ser complejidad temporal (por ejemplo, solo se considera el fotograma de referencia inter-fotograma para el fotograma de vídeo), y se expresa como

$f_{\text{vídeo_contenido_complejidad}} = f_{\text{vídeo_contenido_complejidad_tcc}}$ (en la fórmula precedente (5)) o $f_{\text{vídeo_contenido_complejidad_n}} = f_{\text{vídeo_contenido_complejidad_ntcc}}$ (complejidad temporal ajustada en la fórmula precedente (8)); o puede ser complejidad espacial (por ejemplo, solo se considera el fotograma de referencia intra-fotograma para el fotograma de vídeo), y expresado como
 5 $f_{\text{vídeo_contenido_complejidad}} = f_{\text{vídeo_contenido_complejidad_scc}}$ (en la fórmula precedente (5)) o $f_{\text{vídeo_contenido_complejidad_n}} = f_{\text{vídeo_contenido_complejidad_nscc}}$ (complejidad espacial ajustada en la fórmula precedente (8)); o puede ser una combinación de la complejidad temporal y de la complejidad espacial, y expresada como $f_{\text{vídeo_contenido_complejidad}} = \text{func1}(f_{\text{vídeo_contenido_complejidad_tcc}}, f_{\text{vídeo_contenido_complejidad_scc}})$ o
 10 $f_{\text{vídeo_contenido_complejidad_n}} = \text{func2}(f_{\text{vídeo_contenido_complejidad_ntcc}}, f_{\text{vídeo_contenido_complejidad_nscc}}$). Esta realización de la presente invención no impone limitación alguna sobre ello.

Cabe destacar que, en una realización que utiliza la fórmula (6) o la fórmula (7) para predecir la calidad de codificación o la calidad de distorsión de codificación de vídeo, es posible utilizar la complejidad de contenido del vídeo obtenida mediante ajuste en la fórmula (8): La complejidad ajustada de contenido del vídeo $f_{\text{vídeo_contenido_complejidad_n}}$ puede ser complejidad temporal del vídeo $f_{\text{vídeo_contenido_complejidad_ntcc}}$ que se obtiene después de un ajuste normalizado, puede ser complejidad espacial del vídeo $f_{\text{vídeo_contenido_complejidad_nscc}}$ que se obtiene después de un ajuste normalizado, o puede ser una combinación de complejidad temporal $f_{\text{vídeo_contenido_complejidad_ntcc}}$ y complejidad espacial $f_{\text{vídeo_contenido_complejidad_nscc}}$. Se ha de comprender que esta realización de la presente invención no impone limitación alguna sobre ello.

Por ejemplo, la calidad de codificación de vídeo $d_{\text{compresión_calidad_valor}}$ se predice mediante la combinación de complejidad temporal $f_{\text{vídeo_contenido_complejidad_ntcc}}$ y la complejidad espacial $f_{\text{vídeo_contenido_complejidad_nscc}}$, y los parámetros de cuantificación $f_{\text{vídeo_pc}}$ del vídeo, es decir, en la fórmula (6), la complejidad de contenido del vídeo utiliza la complejidad ajustada de contenido del vídeo $f_{\text{vídeo_contenido_complejidad_n}}$, y $a_5 * f_{\text{vídeo_contenido_complejidad_n}}$ puede ser reemplazada por $a_{51} * f_{\text{vídeo_contenido_complejidad_ntcc}} + a_{52} * f_{\text{vídeo_contenido_complejidad_nscc}}$, donde a_{51} y a_{52} pueden ser iguales o diferentes.

Una forma de predecir calidad de codificación según esta realización de la presente invención satisface mejor los sentimientos subjetivos de la vista humana, produciéndose así una mejora adicional en términos de precisión a la hora de predecir la calidad de codificación.

La Fig. 3 es un diagrama de bloque estructural de un dispositivo para evaluar calidad de codificación de vídeo según otra realización de la presente invención.

Tal y como se representa en la Fig. 3, como regla general se incluyen al menos un procesador 410 como, por ejemplo, un CPU, al menos un puerto 420, una memoria 430, y al menos un bus de comunicaciones 440. El bus de comunicaciones 440 está configurado para implementar conexiones y comunicación entre estos aparatos. El procesador 410 está configurado para ejecutar módulos ejecutables almacenados en la memoria 430, por ejemplo, programas informáticos. Opcionalmente, el dispositivo 300 puede incluir una interfaz de usuario 450, donde la interfaz de usuario 450 incluye pero no está limitada a una pantalla, un teclado, y un dispositivo para hacer clic, por ejemplo, un ratón, una bola de seguimiento (*trackball*), un panel táctil, o una pantalla táctil. La memoria 430 puede incluir una memoria RAM de alta velocidad, o puede incluir una memoria no volátil (memoria no volátil), por ejemplo al menos una memoria de disco.

En algunas formas de implementación, la memoria 430 almacena los siguientes elementos, una estructura de datos o módulo ejecutable, o un subconjunto de los mismos, o un conjunto de extensión de los mismos:

45 un sistema operativo 432, que incluye diversos software de sistema, y configurado para implementar diversos servicios básicos y procesar tareas basadas en hardware; y

un módulo de aplicación 434, que incluye diversas aplicaciones, y configurado para implementar diversos servicios de aplicación.

El módulo de aplicación 434 incluye pero no está limitado a una unidad de adquisición 201, una unidad de determinación 202, y una unidad de predicción 203.

Para la implementación específica de unidades en el módulo de aplicación 434, será preciso referirse a unidades correspondientes en la realización representada en la Fig. 2, y no se vuelven a describir detalles en la presente memoria.

El dispositivo precedente para evaluar calidad de codificación de vídeo puede ser un terminal, por ejemplo, un aparato móvil portátil, de bolsillo, de mano, integrado en un ordenador o montado en un vehículo, o el dispositivo puede ser un servidor o similar.

5 Una persona con experiencia ordinaria en la técnica puede ser consciente de que, en combinación con los ejemplos descritos en las realizaciones descritas en esta memoria descriptiva, es posible implementar unidades y etapas de algoritmos mediante hardware electrónico o una combinación de software de ordenador y hardware electrónico. Si las funciones se realizan mediante un hardware o software depende de las aplicaciones particulares y las condiciones de limitaciones de diseño de las soluciones técnicas. Una persona experta en la técnica puede utilizar diferentes métodos para implementar las funciones descritas para cada aplicación particular, pero no se debe considerar que la implementación excede el alcance de la presente invención.

10 Un experto en la técnica puede comprender claramente que, a los fines de una descripción conveniente y breve, para un proceso de trabajo detallado del sistema, aparato y unidad anteriores, se debe hacer referencia a un proceso correspondiente en las realizaciones de método anteriores, y los detalles no se vuelven a describir en la presente memoria.

15 En las diversas realizaciones proporcionadas en la presente solicitud, se ha de comprender que el sistema, aparato y método descritos pueden implementarse de otras maneras. Por ejemplo, la realización de aparato descrita es meramente un ejemplo. Por ejemplo, la división de unidad es meramente una división de función lógica y en la implementación real la división puede ser otra. Por ejemplo, se pueden combinar o integrar en otro sistema múltiples unidades o componentes, o algunas características se pueden ignorar o no llevar a cabo. Además, los acoplamientos mutuos representados o descritos o los acoplamientos directos o conexiones de comunicaciones se pueden implementar a través de algunas interfaces. Los acoplamientos indirectos o conexiones de comunicación entre los aparatos o unidades se pueden implementar de forma electrónica, mecánica u otras.

20 Las unidades descritas como partes separadas pueden o no estar físicamente separadas, y las partes representadas como unidades pueden o no ser unidades físicas, estar ubicadas en una posición, o pueden estar distribuidas en múltiples unidades de red. Algunas o todas las unidades se pueden seleccionar según las necesidades reales para lograr los objetivos de las soluciones de las realizaciones.

25 Además, las unidades funcionales en las realizaciones de la presente invención pueden estar integradas en una unidad de procesamiento, o cada una de las unidades puede estar sola físicamente, o dos o más unidades pueden estar integradas en una unidad.

30 Cuando las funciones se implementan en la forma de una unidad funcional de software y se venden o usan como un producto independiente, las funciones se pueden almacenar en un medio de almacenamiento legible por ordenador. Según este acuerdo, las soluciones técnicas de la presente invención esencialmente, o la parte que contribuye a la técnica anterior, o algunas de las soluciones técnicas, se pueden implementar en una forma de un producto de software. El producto de software está almacenado en un medio de almacenamiento, e incluye diversas instrucciones para indicar a un dispositivo (que puede ser un ordenador, un servidor o un dispositivo de red) que realice alguna o todas las etapas de los métodos descritos en las realizaciones de la presente invención. El medio de almacenamiento anterior incluye: cualquier medio que pueda almacenar un código de programa, tal y como una
35 unidad flash USB, un disco duro removible, una memoria de solo lectura (ROM, memoria de sólo lectura), una memoria de acceso aleatorio (RAM, memoria de acceso aleatorio), un disco magnético, o un disco óptico.

REIVINDICACIONES

1. Un método para evaluar calidad de codificación de vídeo, que comprende:

adquirir un parámetro de cuantificación de un segmento de un fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo;

5 determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo, y determinar la complejidad de contenido del vídeo según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo;

en donde determinar la complejidad de contenido del vídeo según la cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo comprende:

10 determinar una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{_segmento_byte_por_píxel}_k = \frac{i_{_segmento_tamaño}_k}{i_{_segmento_píxel}_k}, \text{ donde}$$

15 $i_{_segmento_tamaño}_k$ es una cantidad de bytes del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, $i_{_segmento_píxel}_k$ es una cantidad de píxeles del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm, y Núm es un número entero positivo;

determinar la complejidad de contenido del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{_segmento_contenido_complejidad}_k = a [i_{_segmento_pc}_k] \cdot f_{_segmento_byte_por_píxel}_k + b [i_{_segmento_pc}_k], \text{ donde}$$

20 tanto $a[i_{_segmento_pc}_k]$ como $b[i_{_segmento_pc}_k]$ son valores correspondientes a un parámetro de cuantificación del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo;

determinar la complejidad de contenido del fotograma de vídeo de la siguiente forma:

$$f_{_fotograma_contenido_complejidad} = \frac{\sum_{k=1}^{Núm} f_{_segmento_contenido_complejidad}_k}{Núm};$$

y

25 determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo;

y

predecir la calidad de codificación de vídeo según la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo.

30 2. El método según la reivindicación 1, en donde determinar un parámetro de cuantificación de un vídeo de acuerdo con el parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo comprende:

determinar un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación del vídeo.

35 3. El método según la reivindicación 2, en donde determinar un valor promedio o un valor promedio ponderado del parámetro de cuantificación del segmento del fotograma de vídeo de la secuencia de vídeo como el parámetro de cuantificación del vídeo comprende:

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} i_{\text{segmento_pc}_i}}{N1}, 0$$

$$f_{\text{vídeo_pc}} = \frac{\sum_{i=1}^{N1} p_i \cdot i_{\text{segmento_pc}_i}}{\sum_{i=1}^{N1} p_i}, \text{ donde}$$

5 $f_{\text{vídeo_pc}}$ es el parámetro de cuantificación del vídeo, N1 es una cantidad de parámetros de cuantificación de segmentos de la secuencia de vídeo, $i_{\text{segmento_pc}_i}$ es un parámetro de cuantificación del $i^{\text{ésimo}}$ segmento entre parámetros de cuantificación de los N1 segmentos, N1 es un número entero positivo, y p_i es una ponderación correspondiente al $i^{\text{ésimo}}$ segmento.

4. El método según la reivindicación 1, en donde determinar la complejidad de contenido del vídeo según la complejidad de contenido del fotograma de vídeo comprende:

$$f_{\text{vídeo_contenido_complejidad}} = \frac{\sum_{j=1}^M f_{\text{fotograma_contenido_complejidad}_j}}{M}$$

10 donde

$f_{\text{vídeo_contenido_complejidad}}$ es la complejidad de contenido del vídeo, la secuencia de vídeo incluye M fotogramas de vídeo, M es un número entero positivo, y j es un número entero positivo y oscila entre 1 y M.

5. El método según la reivindicación 4, en donde predecir la calidad de codificación de vídeo según la complejidad de contenido del vídeo y el parámetro de cuantificación del vídeo comprende:

$$d_{\text{compresión_calidad_valor}} = a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}}$$

15 , donde

$d_{\text{compresión_calidad_valor}}$ es la calidad de codificación de vídeo, y a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 son todas constantes.

6. El método según la reivindicación 5, en donde un valor de a_2 es:

una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo.

7. El método según la reivindicación 5 o 6, en donde un valor de a_1 es:

20 el valor mínimo del intervalo de la calidad de codificación de vídeo.

8. El método según cualquiera de las reivindicaciones 3 a 7, que además comprende: ajustar la complejidad de contenido del vídeo de la siguiente manera:

$$f_{\text{vídeo_contenido_complejidad_n}} = \min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right), y$$

usando $f_{\text{vídeo_contenido_complejidad_n}}$ ajustado como la complejidad de contenido del vídeo, donde tanto s como núm1 son constantes.

5 9. El método según la reivindicación 8, en donde un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valores de $f_{\text{vídeo_contenido_complejidad_n}}$ es [0,0, 1,0].

10. El método según cualquiera de las reivindicaciones 4 a 9, en donde los M fotogramas de vídeo no tienen pérdida de paquetes.

11. El método según cualquiera de las reivindicaciones 4 a 10, en donde el fotograma de vídeo comprende un fotograma de codificación intra-fotograma y/o un fotograma de codificación inter-fotograma.

10 12. El método según cualquiera de las reivindicaciones 1 a 11, en donde Núm es una cantidad total de segmentos en el fotograma de vídeo.

13. Un dispositivo para evaluar calidad de codificación de vídeo, que comprende:

15 una unidad de adquisición, configurada para adquirir un parámetro de cuantificación de un segmento de un fotograma de vídeo de una secuencia de vídeo y una cantidad de bytes por píxel del segmento del fotograma de vídeo de la secuencia de vídeo;

una unidad de determinación, configurada para determinar un parámetro de cuantificación de un vídeo según el parámetro de cuantificación, adquirido mediante la unidad de adquisición, del segmento del fotograma de vídeo de la secuencia de vídeo; la unidad de determinación está configurada para:

20 determinar una cantidad de bytes por píxel del $k^{\text{ésimo}}$ segmento entre Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{\text{segmento_byte_por_píxel}_k} = \frac{i_{\text{segmento_tamaño}_k}}{i_{\text{segmento_píxel}_k}}, \text{ donde}$$

25 $i_{\text{segmento_tamaño}_k}$ es una cantidad de bytes del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, $i_{\text{segmento_píxel}_k}$ es una cantidad de píxeles del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo, y k es un número entero positivo, donde en el fotograma de vídeo, k oscila entre 1 y Núm , y Núm es un número entero positivo;

la unidad de determinación está específicamente configurada para determinar la complejidad de contenido del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo de la siguiente forma:

$$f_{\text{segmento_contenido_complejidad}_k} = a [i_{\text{segmento_pc}_k}] \cdot f_{\text{segmento_byte_por_píxel}_k} + b [i_{\text{segmento_pc}_k}]$$

donde

30 tanto $a [i_{\text{segmento_pc}_k}]$ como $b [i_{\text{segmento_pc}_k}]$ son valores correspondientes a un parámetro de cuantificación del $k^{\text{ésimo}}$ segmento entre los Núm segmentos del fotograma de vídeo;

la unidad de determinación está configurada para determinar la complejidad de contenido del fotograma de vídeo de la siguiente manera:

$$f_{_fotograma_contenido_complejidad} = \frac{\sum_{k=1}^{M_{\text{v\u00eda}}} f_{_segmento_contenido_complejidad}_k}{\text{N\u00fam}} ;$$

y

la unidad de determinaci\u00f3n est\u00e1 configurada para determinar la complejidad de contenido del v\u00eddeo seg\u00fan la complejidad de contenido del fotograma de v\u00eddeo;

5 y

una unidad de predicci\u00f3n, configurada para predecir la calidad de codificaci\u00f3n de v\u00eddeo seg\u00fan la complejidad de contenido del v\u00eddeo y el par\u00e1metro de cuantificaci\u00f3n del v\u00eddeo determinados mediante la unidad de determinaci\u00f3n.

14. El dispositivo seg\u00fan la reivindicaci\u00f3n 13, en el que:

10 la unidad de determinaci\u00f3n est\u00e1 espec\u00edficamente configurada para determinar un valor promedio o un valor promedio ponderado del par\u00e1metro de cuantificaci\u00f3n del segmento del fotograma de v\u00eddeo de la secuencia de v\u00eddeo como el par\u00e1metro de cuantificaci\u00f3n del v\u00eddeo.

15. El dispositivo seg\u00fan la reivindicaci\u00f3n 14, en el que:

la unidad de determinaci\u00f3n est\u00e1 espec\u00edficamente configurada para determinar el par\u00e1metro de cuantificaci\u00f3n del v\u00eddeo de la siguiente manera:

$$f_{_v\u00eddeo_pc} = \frac{\sum_{i=1}^{N1} i_{_segmento_pc}_i}{N1} , \text{ o}$$

$$f_{_v\u00eddeo_pc} = \frac{\sum_{i=1}^{N1} p_i \cdot i_{_segmento_pc}_i}{\sum_{i=1}^{N1} p_i} , \text{ donde}$$

20 $f_{_v\u00eddeo_pc}$ es el par\u00e1metro de cuantificaci\u00f3n del v\u00eddeo, $N1$ es una cantidad de par\u00e1metros de cuantificaci\u00f3n de segmentos de la secuencia de v\u00eddeo, $i_{_segmento_pc}_i$ es un par\u00e1metro de cuantificaci\u00f3n del $i^{\text{\u00e9simo}}$ segmento entre par\u00e1metros de cuantificaci\u00f3n de los $N1$ segmentos, $N1$ es un n\u00famero entero positivo, y p_i es una ponderaci\u00f3n correspondiente al $i^{\text{\u00e9simo}}$ segmento.

16. El dispositivo seg\u00fan la reivindicaci\u00f3n 13, en el que:

la unidad de determinaci\u00f3n est\u00e1 espec\u00edficamente configurada para determinar la complejidad de contenido del v\u00eddeo de la siguiente manera:

$$f_{_v\u00eddeo_contenido_complejidad} = \frac{\sum_{j=1}^M f_{_fotograma_contenido_complejidad}_j}{M} ;$$

25 donde

la secuencia de v\u00eddeo incluye M fotogramas de v\u00eddeo, M es un n\u00famero entero positivo, y j es un n\u00famero entero positivo y oscila entre 1 y M .

17. El dispositivo seg\u00fan la reivindicaci\u00f3n 16, en el que:

30 la unidad de predicci\u00f3n est\u00e1 espec\u00edficamente configurada para predecir la calidad de codificaci\u00f3n de v\u00eddeo de la siguiente manera:

$d_{\text{compresión_calidad_valor}} =$

$$a_1 + \frac{a_2}{a_3 + \left(\frac{f_{\text{vídeo_pc}}}{a_4 - a_5 \cdot f_{\text{vídeo_contenido_complejidad}}} \right)^{a_6}}$$

, donde

a_1 , a_2 , a_3 , a_4 , a_5 , y a_6 son todas constantes.

18. El dispositivo según la reivindicación 17, en donde un valor de a_2 es:

una diferencia entre un valor máximo y un valor mínimo de un intervalo de la calidad de codificación de vídeo.

5 19. El dispositivo según la reivindicación 17 o 18, en donde un valor de a_1 es:

el valor mínimo del intervalo de la calidad de codificación de vídeo.

20. El dispositivo según cualquiera de las reivindicaciones 13 a 19, que además comprende: ajustar la complejidad de contenido del vídeo de la siguiente manera:

$f_{\text{vídeo_contenido_complejidad_n}} =$

$$\min\left(s, \sqrt{\frac{f_{\text{vídeo_contenido_complejidad}}}{\text{núm1}}}\right)$$

, y

10 usando $f_{\text{vídeo_contenido_complejidad_n}}$ ajustado como la complejidad de contenido del vídeo, donde tanto s como núm1 son todas constantes.

21. El dispositivo según la reivindicación 20, en donde un valor de s es 1,0, un valor de núm1 es 60,0, y un intervalo de valores de $f_{\text{vídeo_contenido_complejidad_n}}$ es [0,0, 1,0].

15 22. El dispositivo según cualquiera de las reivindicaciones 16 a 21, en donde los M fotogramas de vídeo no tienen pérdida de paquetes.

23. El dispositivo según cualquiera de las reivindicaciones 16 a 22, en donde el fotograma de vídeo comprende un fotograma de codificación intra-fotograma y/o un fotograma de codificación inter-fotograma.

24. El dispositivo según cualquiera de las reivindicaciones 13 a 23, en donde Núm es una cantidad total de segmentos en el fotograma de vídeo.

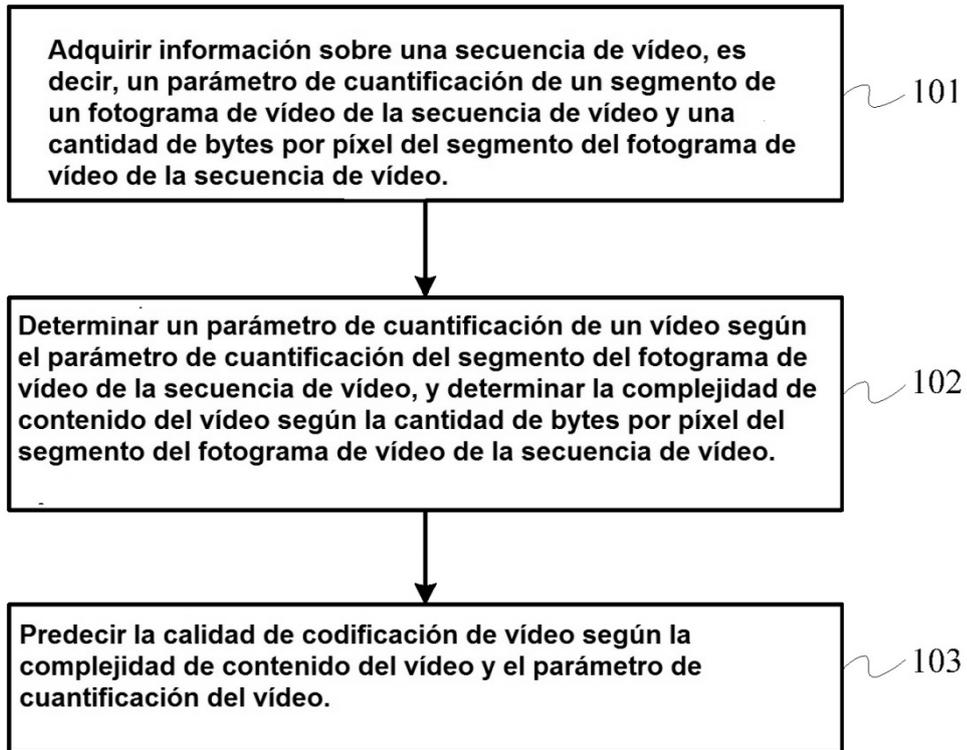


FIG. 1

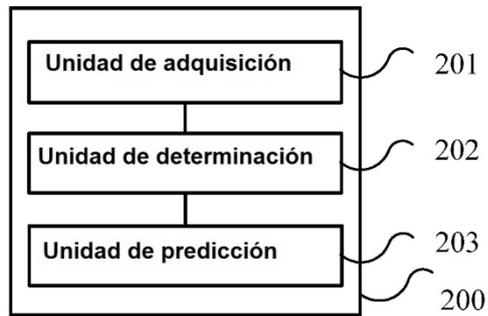


FIG. 2

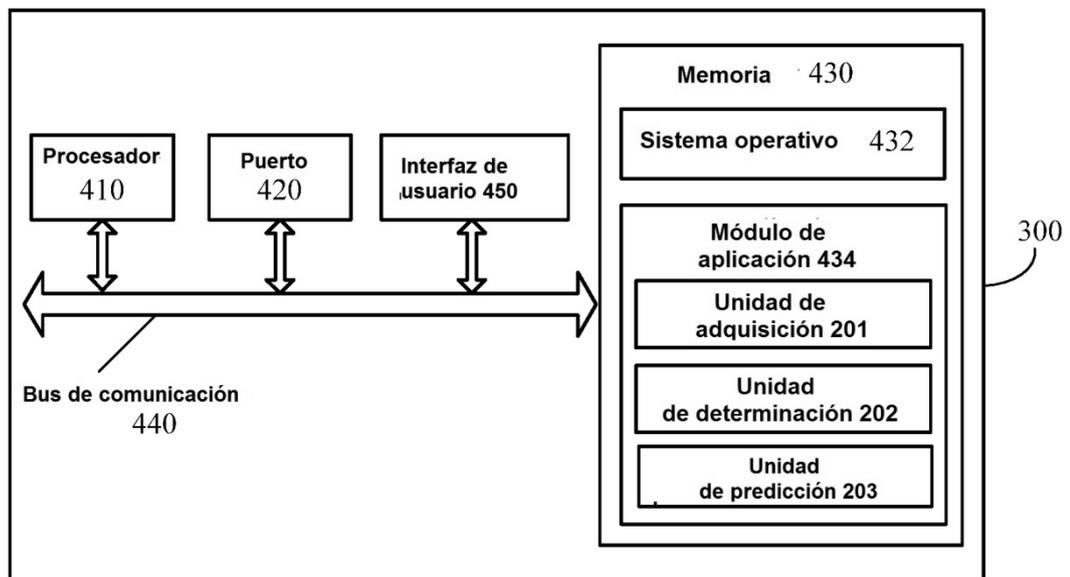


FIG. 3