

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 385 324**

51 Int. Cl.:  
**H04B 1/66** (2006.01)  
**H04N 5/14** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **02775897 .8**  
96 Fecha de presentación: **20.09.2002**  
97 Número de publicación de la solicitud: **1430613**  
97 Fecha de publicación de la solicitud: **23.06.2004**

54 Título: **Detección de cortes de escena en un flujo de bitios de vídeo MPEG**

30 Prioridad:  
**26.09.2001 US 324890 P**  
**11.12.2001 US 14963**

45 Fecha de publicación de la mención BOPI:  
**23.07.2012**

45 Fecha de la publicación del folleto de la patente:  
**23.07.2012**

73 Titular/es:  
**Thomson Licensing**  
**1-5, rue Jeanne d'Arc**  
**92130 Issy-les-Moulineaux, FR**

72 Inventor/es:  
**PETERSON, Richard Mateer y**  
**CHALOM, Edmond**

74 Agente/Representante:  
**de Elzaburu Márquez, Alberto**

ES 2 385 324 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**

Detección de cortes de escena en un flujo de bitios de vídeo MPEG.

**Referencias cruzadas con aplicaciones relacionadas**

5 Esta solicitud reivindica los beneficios de la Solicitud de Patente Provisional de EEUU Nº 60/324.890, presentada el 26 de septiembre de 2001.

**Campo del invento**

El presente invento se refiere a un vídeo de tipo MPEG, y más particularmente a la detección de un corte de escena en el flujo de bitios de vídeo.

**Antecedentes del invento**

10 Las secuencias de vídeo basadas en MPEG pueden estar divididas en escenas que pueden variar en longitud desde menos de un segundo hasta varios minutos o más. La detección de estas divisiones o cortes de escena hace posible una amplia variedad de características de valor añadido. Por ejemplo, una trama puede ser seleccionada de cada escena para crear un guión que resume los contenidos de una grabación de vídeo. Además, se puede utilizar un editor de vídeo para manipular la grabación de vídeo en modo de escena a escena, por ejemplo reordenando escenas o mediante la aplicación de técnicas de procesamiento de imágenes a cada trama en una escena.

15 Las secuencias de vídeo MPEG incluyen tres tipos de tramas, Intra-Tramas (I), Inter-Tramas (P), y Tramas Bidireccionales (B). Las tramas I codifican una imagen fija utilizando un método similar a la codificación JPEG. Las tramas P son predichas a partir de una trama I o P previa. Las tramas B son predichas a partir de una trama I o P y de una trama I o P próxima. Estos tres tipos de tramas se codifican usando una Transformación Cosenoidal Discreta (DCT), la cual organiza la redundancia en direcciones espaciales entre tramas. No obstante, para las tramas I la información DCT se obtiene directamente a partir de una muestra de imagen, en tanto que para las tramas P y B la información DCT se obtiene a partir de un error residual después de la predicción.

20 Cada trama se divide en una pluralidad de macrobloques. Cada uno de ellos incluye información relacionada con una pluralidad de bloques de luminancia, por ejemplo Y1, Y2, Y3 e Y4, y una pluralidad de bloques de crominancia, por ejemplo uno U y otro V en un sistema YUV. Cada uno de estos bloques incluye una pluralidad de pels, o elementos de elementos de imagen, por ejemplo un bloque 8x8.

25 Cuando el vídeo ha sido codificado en un flujo de bitios de tipo MPEG la detección del corte de escena puede realizarse sin la decodificación completa del flujo de bitios, lo cual aumenta la velocidad de procesamiento del vídeo. También se puede extraer información adicional tal como tipos de codificación de macrobloques, vectores de movimiento y coeficientes DCT a partir del flujo de bitios sin decodificar totalmente el flujo de bitios.

30 Un método para la detección de cortes de escena se realiza de la siguiente manera:

1. para las tramas I se determinan las diferencias cuadráticas medias entre los coeficientes DCT;
2. para las tramas P el método propuesto determina el número de macrobloques predichos hacia adelante;
3. para las tramas B se cuenta el menor de los números de macrobloques codificados en adelante y del número de macrobloques codificados hacia atrás; y
4. a continuación se determina un mínimo en una representación de estos números con respecto al número de tramas.

35 En otro método propuesto para la detección de cortes de escena: para las tramas I se utiliza una diferencia entre los histogramas de color formados a partir de los coeficientes DC usados, y se combina con la información sobre la relación entre el número de macrobloques sin compensación de movimiento con el número de ellos con compensación de movimiento. El método propuesto busca un pico en una representación con respecto al número de tramas. Para las tramas B se determina la relación de las predicciones hacia adelante con las predicciones hacia atrás. En todos los casos se aplica una técnica de umbral adaptativa a fin de identificar los picos.

40 Otro método adicional hace uso de histogramas para todas las tramas (I, P y B) formadas a partir de los coeficientes Cosenoidales Discretos (DC) con compensación de movimiento.

45 No obstante, no existe actualmente un sistema o método conocido de detección de corte de escena basado en el examen global de todas las predicciones dentro de un subgrupo de imágenes (GOP).

## Resumen del invento

El invento se refiere a un método de acuerdo con la reivindicación 1. Las realizaciones preferidas se detallan en las reivindicaciones dependientes.

## Breve descripción de los dibujos

- 5 A continuación se describen con detalle las realizaciones preferidas del presente invento haciendo referencia a los dibujos que se acompañan, en los que:
- la Figura 1 ilustra unos cortes de escena en un subgrupo GOP de acuerdo con los principios del presente invento;
- 10 la Figura 2 es un diagrama de flujos que ilustra un primer método para la detección de un corte de escena de acuerdo con los principios del presente invento; y
- la Figura 3 es un diagrama de flujos que ilustra un segundo método para la detección de un corte de escena de acuerdo con los principios del presente invento.

## Descripción detallada

15 De acuerdo con los principios del presente invento la detección del corte de escena se realiza recogiendo datos sobre el número de macrobloques predichos hacia adelante y hacia atrás en cada Inter Trama (P) o Trama Bidireccional (B), analizándolos a un nivel de sub GOP (cada trama de anclaje junto con las tramas B precedentes) sin considerar cada trama de forma independiente. Se sobreentiende que el presente invento puede ser puesto en práctica en diversas formas de soporte físico, soporte lógico, microprogramas, procesadores con fines específicos, o una combinación de ellos. En una realización a modo de ejemplo el presente invento se ha puesto en práctica en un soporte lógico como un programa de aplicación incorporado de forma tangible en un dispositivo de almacenamiento de programas. El programa de la aplicación puede ser cargado en, y ejecutado por, una máquina que comprende cualquier disposición de elementos apropiada. Preferiblemente, la máquina está instalada en una plataforma de ordenador que tiene un soporte físico tal como una o más unidades centrales de procesamiento (CPU), una memoria de acceso aleatorio (RAM), y una interfaz o interfaces de entrada/salida (I/O). La plataforma de ordenador incluye también un sistema operativo y un código de microinstrucción. Los diversos procesos y funciones descritos aquí pueden formar parte bien del código de microinstrucción o del programa de aplicación (o de una combinación de ellos) que se ejecuta a través del sistema operativo. Además, se pueden conectar otros diversos dispositivos periféricos a la plataforma de ordenador tal como un dispositivo o dispositivos de almacenamiento de datos y un dispositivo o dispositivos de impresión.

20 Se ha de entender además que, debido a que algunos de los componentes del sistema y de los pasos del método descritos en las figuras que se acompañan pueden ser puestas en práctica en soporte lógico, las conexiones, interconexiones o relaciones funcionales efectivas entre los componentes del sistema (o los pasos del proceso) pueden diferir dependiendo de la forma en la que se haya programado el presente invento. Las enseñanzas del presente invento aquí proporcionadas así como éstas y otras aplicaciones o configuraciones similares del presente invento podrán ser comprendidas por cualquier persona con una formación normal en la técnica relacionada.

25 Para cada trama P o B se registra el número de macrobloques que utilizan la predicción hacia adelante y el número de macrobloques que utilizan la predicción hacia atrás (encontrada solamente en las tramas B). Se examinan los datos de cada trama de anclaje, intra-trama (I) o P, y todas las tramas B inmediatamente anteriores.

30 Si un macrobloque ha sido predicho de forma inapropiada la compensación de movimiento posterior residual puede ser amplia, y se necesitan unos coeficientes DCT amplios con el fin de realizar una compensación por la predicción. Para estas predicciones existe una alta probabilidad (aproximadamente mayor del 50%) de que la predicción sea falsa. Esto es, que un macrobloque de trama de anclaje en el que se base la predicción pueda no tener una asociación real con el macrobloque en la trama predicha, aunque la trama de anclaje pueda ser lo suficientemente similar a para que a partir de ella se pueda realizar una predicción. Debido a esto se ignora tal macrobloque en la recogida de los datos estadísticos. Más particularmente, cuando se cuentan macrobloques con predicción hacia adelante o hacia atrás se examina la energía total (entropía) de los coeficientes DCT. La energía total de los coeficientes DCT es la suma de los cuadrados de los coeficientes no-DC. Si esta energía es superior a un valor umbral no se tiene en cuenta el macrobloque, ni en cuanto al número de predicciones hacia adelante o hacia atrás ( $nFwd$  y  $nBwd$  respectivamente), ni en cuanto al número de macrobloques codificados ( $nCode$ ). Para MPEG-2 se ha encontrado que como ejemplo un valor umbral es aproximadamente 500.

35 Para una trama de anclaje se analiza el patrón de predicciones para cada trama en el sub-GOP. Considerando la Figura 1, en la que se representa un sub-GOP que incluye dos tramas B y una trama P. Los sitios posibles de un corte de escena incluyen, antes de la trama B<sub>1</sub> 102, entre las tramas B<sub>1</sub> 102 y B<sub>2</sub> 103, y entre la trama B<sub>2</sub> 103 y la trama P<sub>3</sub> 104. Además, si las tramas están entrelazadas, se puede producir un corte de escena entre los campos de cualquiera de las tres tramas 105 a 107 respectivamente.

De acuerdo con una realización del presente invento, si se produce un corte de escena en 108 pocos macrobloques predichos hacia adelante estarían presentes en las tramas  $B_1$  102,  $B_2$  103, o  $P_3$  104 debido a que la trama  $I/P_0$  101 no está relacionada con cualquiera de estas tramas. Igualmente, si se produce un corte de escena en 109 en la trama  $B_1$  102 estarían presentes pocos macrobloques predichos hacia atrás debido a que dicha trama  $B_1$  102 no está relacionada con la trama  $P_3$  104 a partir de la cual se han realizado estas predicciones; y esperamos pocos macrobloques predichos hacia adelante en la trama  $B_2$  103 o  $P_3$  104 debido a que la trama  $I/P_0$  101 no está relacionada con estas tramas.

De este modo, una medida de la posibilidad de un corte de escena en 108 es el total del número de macrobloques predichos hacia adelante en las tramas  $B_1$  102,  $B_2$  103, o  $P_3$  104. Una medida de la posibilidad de un corte de escena en 109 es el total del número de macrobloques predichos hacia atrás en la trama  $B_1$  102 más el número de macrobloques predichos hacia adelante en las tramas  $B_2$  103 y  $P_3$  104; y similarmente en cuanto a la posibilidad de un corte de escena en 110. La posibilidad de que haya un corte de escena en el punto correspondiente aumenta a medida que disminuye el número total de macrobloques predichos hacia adelante y hacia atrás.

Con referencia a la Figura 2, después de determinar la imagen ( $p$ ) actual que sea una trama P o una trama B 201 se fija el número total de macrobloques codificados de la trama ( $nCode[p]$ ) en el número de macrobloques que tienen una  $AcEnergy$  menor que una  $energyTresh$  202. El número total de macrobloques predichos hacia adelante de la trama ( $nFwd[p]$ ) se fija como el número de macrobloques que tienen una  $AcEnergy$  menor que una predicción hacia adelante AND  $energyTresh$  203. El número total de macrobloques predichos hacia atrás de la trama ( $nBwd[p]$ ) se fija como el número de macrobloques que tienen una  $AcEnergy$  menor que una predicción hacia atrás AND  $energyTresh$  204.

Con referencia ahora a la Figura 3, de acuerdo con el paso 301 se ha determinado si  $p$  es una trama I o P. Si  $p$  no es una trama I o P no existe una detección de corte de escena hasta el final del sub-GOP de acuerdo con el paso 302. De acuerdo con el paso 303, se ha determinado si  $p$  es una trama I y si la imagen anterior era una trama I o P. Si se ha determinado que  $p$  es una trama I y que la imagen anterior era una trama I o P, entonces no se ha detectado un corte de escena en este sub-GOP de acuerdo con el paso 304. Tras la determinación de que  $p$  es de tipo P o que la imagen anterior es de tipo B entonces se recogen datos con respecto al sub-GOP actual en los pasos 305-311. El recuento de las predicciones que cruzan el corte de escena potencial entre el campo  $f$  y el campo anterior ( $predMBs[f]$ ) y el recuento de los macrobloques cuyas predicciones podrían cruzar este límite ( $codedMBs[f]$ ) se inician en 0 para cada campo  $f$  en el sub-GOP, de acuerdo con el paso 305.

Para cada campo  $f$  del sub-GOP se determinado si dicho campo  $f$  está en una imagen I, de acuerdo con el paso 307. Si el campo  $f$  está en una imagen I, entonces continúa el proceso en el siguiente campo de acuerdo con el paso 308. Después de determinar que el campo  $f$  no está en una imagen I, entonces el proceso hace un bucle sobre los sitios potenciales de un corte de escena en el sub-GOP, esto es, cada sitio  $f'$  entre dos campos. Se aumenta el valor de  $codedMBs[f']$  en  $nCode[f']$  para cada corte de escena potencial  $f'$  en el sub-GOP, de acuerdo con el paso 309. Además, se aumenta el valor de  $predMBs[f']$  en  $nFwd[f']$  para cada corte de escena potencial  $f'$  antes o igual a  $f$ , de acuerdo con el paso 310.

De acuerdo con el paso 311 se ha determinado si se ha alcanzado el último campo, y si no se ha alcanzado, el proceso continúa con el último campo, de acuerdo con el paso 308. Después de haber determinado que se ha alcanzado el último campo se determina un campo  $f$  para el cual la relación de  $predMBs[f]$  con  $codedMBs[f]$  (por ejemplo,  $predMBs[f]$  con  $codedMBs[f]$ ) sea un mínimo, de acuerdo con el paso 312. Para el campo  $f$  determinado que tiene la relación mínima, si  $predMBs[f]$  dividido por  $codedMBs[f]$  es menor que una  $predRatioTresh$  de acuerdo con el paso 313, el proceso detecta e informa de un corte de escena entre este campo y el campo anterior, de acuerdo con el paso 314, o bien no se selecciona corte de escena alguno en el sub-GOP, de acuerdo con el paso 315.

Obsérvese que para cada campo se determinan el número de macrobloques predichos ( $predMBs$ ) y de macrobloques codificados de campos ( $codedMBs$ ) incluso si las imágenes son imágenes de trama, ya que se puede producir un corte de escena entre campos de una imagen de trama (si la imagen de trama está entrelazada). Los macrobloques de una imagen de trama se aplican a cada campo de la trama al calcular el cálculo de  $predMBs$  y  $codedMBs$ . Esto es,  $nCode[f]$  es el número total de macrobloques codificados en la trama cuando  $f'$  se refiere a cualquier campo de la trama. Igualmente,  $nFwd[f]$  y  $nBwd[f]$  representan el número de macrobloques de predicción hacia adelante y de macrobloques de predicción hacia atrás en el campo  $f'$ , respectivamente. Cuando  $f'$  se encuentra en la misma imagen de trama que  $f$ ,  $predMBs[f]$  se aumenta en  $nBwd[f]$ , y  $codedMBs[f]$  se aumenta en  $nCode[f]$ , solamente cuando  $f$  es el primer campo en la trama. Las matrices del segundo campo no se aumentan debido a que se supone las predicciones hacia adelante que las predicciones anteriores se aplican solamente al primer campo, y las predicciones hacia atrás solamente se aplican al segundo campo. Si no se hiciera esto quedaría enmascarado un corte de escena entre campos de una misma imagen por predicciones a lo largo del límite de la escena.

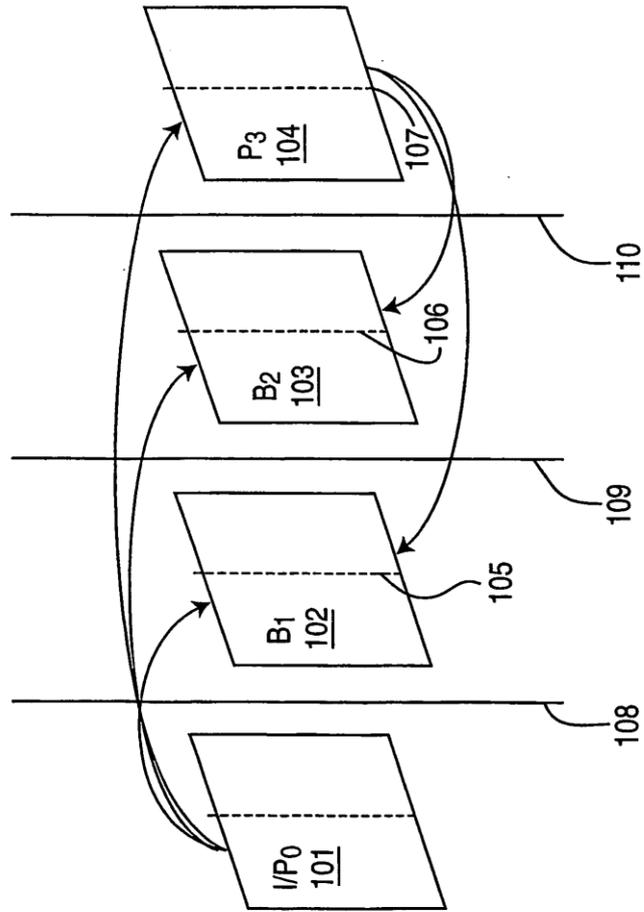
Se ha encontrado mediante experimentación que un buen valor para la predicción del umbral de la relación de predicción  $predRatioTresh$  es 0,05. No obstante, otras experimentaciones y/o aplicaciones pueden determinar que algún valor distinto de este valor, o que un valor umbral puede ser más apropiado en una aplicación específica.

Además, el valor específico escogido para el umbral de energía AC, *AcTresh*, puede igualmente ser distinto del valor escogido aquí como ejemplo. Sería evidente para los expertos en la técnica evaluar y modificar estos y otros valores para maximizar la eficiencia y precisión del invento, incluyendo tal modificación los principios del presente invento aquí descritos.

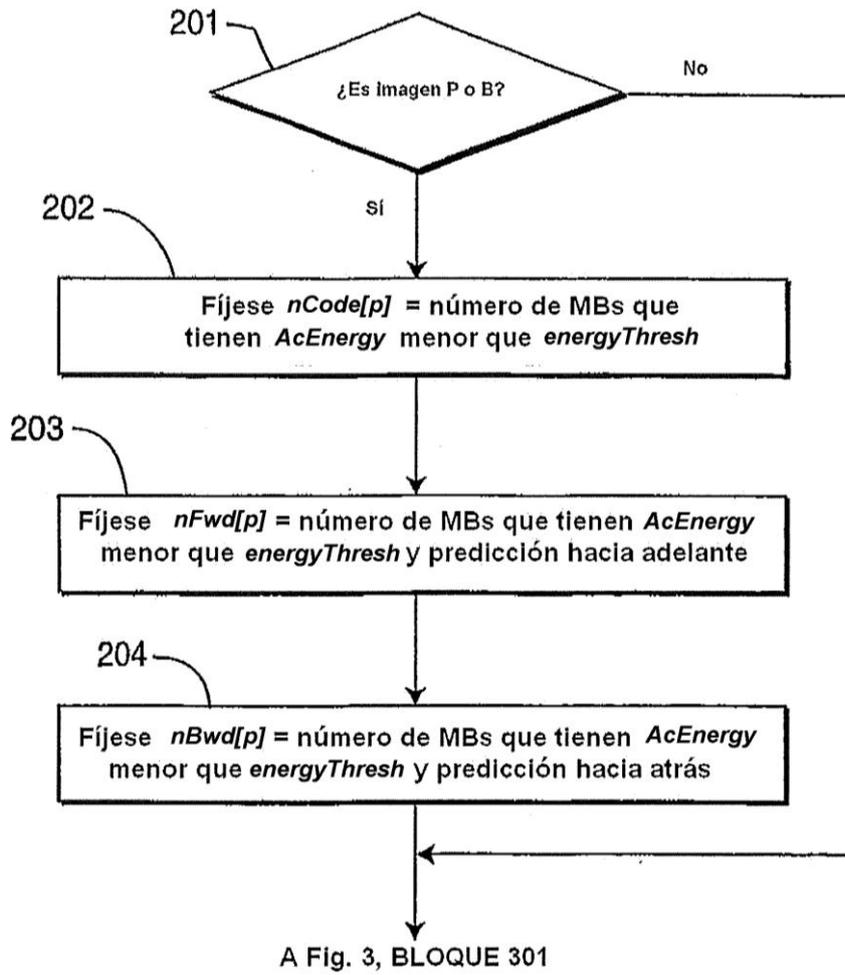
- 5 El método específico antes descrito es el que ha sido puesto en práctica y comprobado. No tiene un carácter limitativo el alcance de esta exposición o de las reivindicaciones en ella contenidas. Un principio subyacente del presente invento es la detección precisa de un corte de escena súbito mediante el examen global de todas las predicciones dentro de un sub-GOP. Buscando la coherencia de toda la información disponible en las predicciones, se pueden reducir las positivas falsas (informes no exactos de los cortes de escena).
- 10 Los principios del presente invento pueden ser integrados con, por ejemplo, enfoques basados en histogramas, a fin de tratar los casos de trama-I solamente, y de una imagen P seguida por una imagen I, en cuyo caso este método no puede detectar un corte de escena. Se pueden también poner en práctica ampliaciones en las que intervienen histogramas para detectar transiciones graduales de escenas tales como desvanecimientos de memoria, fundidos y barridos.
- 15 Habiendo descrito realizaciones para la exacta detección de un corte de escena mediante el examen global de todas las predicciones dentro de un sub-GOP se ha advertido que se pueden realizar modificaciones y variaciones por personas expertas en la técnica a la luz de las anteriores enseñanzas.

**REIVINDICACIONES**

1. Un método para detectar cortes de escena en un flujo de bitios en un vídeo de tipo MPEG, que comprende los pasos de:
- 5 (a) determinar en una pluralidad de tramas cada trama que tiene unos campos primero y segundo, un número de macrobloques codificados para cada campo en cada trama primera y segunda;
- (b) fijar un número de macrobloques predichos igual a un número de macrobloques predichos hacia adelante después de la determinación de un campo correspondiente que ha de ser el primer campo en una trama; y
- 10 (c) fijar el número de macrobloques predichos igual a un número de macrobloques predichos hacia atrás después de la determinación del campo correspondiente que ha de ser el segundo campo en la trama;
- (d) determinar una relación entre el número de macrobloques predichos con el número de macrobloques codificados de cada uno;
- (e) seleccionar un campo que tenga una relación más baja entre los campos primero y segundo;
- (f) comparar la relación más baja con un valor umbral; y
- 15 (g) detectar un corte de escena después de determinar que la relación más baja es menor que el valor umbral, en donde el corte de escena corresponde al campo seleccionado.
2. El método de la reivindicación 1, en el que la pluralidad de tramas es un grupo de tramas, en el que cada una de ellas incluye dos campos, en donde el grupo de tramas comprende un subgrupo de imágenes y el flujo de bitios de vídeo incluye una pluralidad de subgrupos de imágenes.
- 20 3. El método de la reivindicación 2, en el que antes del paso (a) están comprendidos los pasos de:
- determinar si una primera imagen de un subgrupo de imágenes es una imagen I;
- determinar si una imagen anterior era una imagen I o P; y
- avanzar el procesamiento hasta un subgrupo de imágenes siguiente, en respuesta a una determinación de que la primera imagen que ha de ser una imagen I y que la imagen previa ha de ser una imagen I o P.
- 25 4. El método de la reivindicación 1, que además comprende los pasos de:
- iniciar el número de macrobloques predichos en cero para cada campo; e
- iniciar el número de macrobloques codificados en cero para cada campo.
5. El método de la reivindicación 1, en el que al menos uno de la pluralidad de potenciales cortes de escena existe entre campos en entrelazados de una trama.
- 30



**FIG. 1**



**FIG. 2**

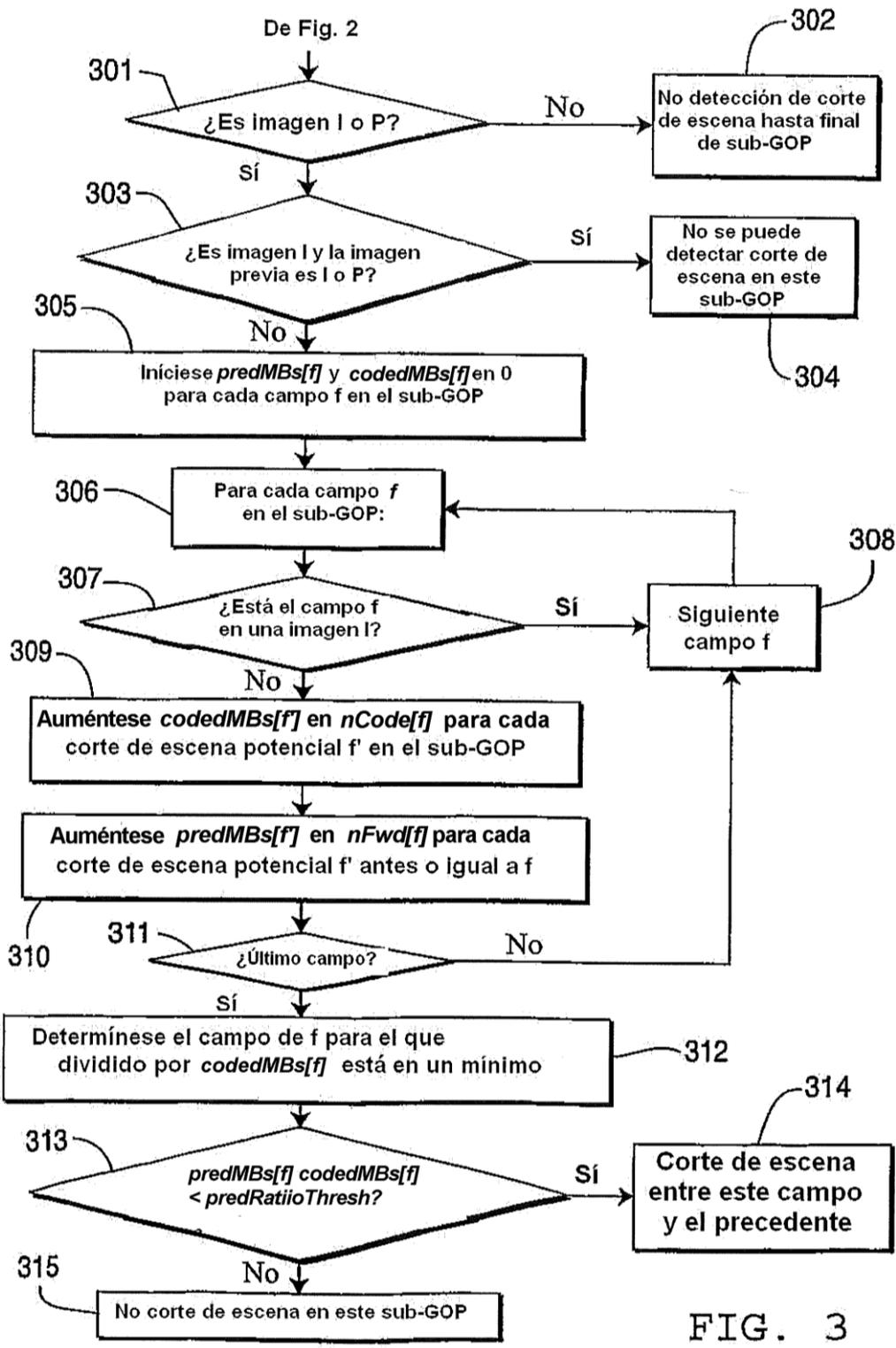


FIG. 3