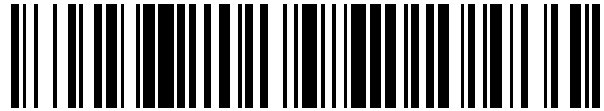


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 575 869**

51 Int. Cl.:

H04N 17/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.07.2010 E 15151228 (2)**

97 Fecha y número de publicación de la concesión europea: **30.03.2016 EP 2866448**

54 Título: **Método y aparato para sincronizar temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.07.2016

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 575 869 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Método y aparato para sincronizar temporalmente el flujo de bits de entrada de un decodificador de vídeo con la secuencia de vídeo procesada descodificada por el decodificador de vídeo

5 Campo de la Invención
La invención se refiere a métodos y aparatos para sincronizar temporalmente el flujo de bits de vídeo de entrada de un decodificador de vídeo con la secuencia de vídeo procesada descodificada por el decodificador de vídeo.

10 Antecedentes de la Invención
La proliferación de aplicaciones de comunicación por vídeo en los últimos años ha hecho necesario el desarrollo de sólidas medidas de calidad de vídeo para evaluar la "calidad de la experiencia" (QoE, por sus siglas en inglés), definida como la calidad del servicio percibida por el usuario. La evaluación de la calidad de vídeo es un aspecto crítico del diseño, planificación y control eficaces de los servicios por parte de los proveedores de contenidos.

15 En la actualidad, los modelos/sistemas híbridos de evaluación de la calidad de vídeo utilizan una combinación de información sobre paquetes, información sobre flujo de bits y la imagen reconstruida descodificada. En general, en un algoritmo híbrido para evaluación de la calidad de vídeo, las características extraídas o calculadas a partir del flujo de bits (por ejemplo, vectores de movimiento, tipos de macrobloques, coeficientes de transformación,
20 parámetros de cuantificación, etc.), y la información extraída de cabeceras de paquete (por ejemplo, tasa de bits, pérdida de paquetes, retardo, etc.) se combinan con las características extraídas de las imágenes reconstruidas de salida, en el campo de los píxeles. Sin embargo, si las anteriores características no corresponden temporalmente a estas últimas debido a la pérdida de alineaciones temporales, la evaluación de la calidad no sería exacta. Por tanto, el primer paso en todos los algoritmos híbridos de evaluación de la calidad de vídeo es la sincronización del flujo de
25 bits de vídeo con las imágenes reconstruidas descodificadas.

En la Figura 1 se representa un diagrama de bloques de un sistema híbrido de evaluación de la calidad de vídeo. En el lado del usuario final, un dispositivo de sondeo captura el flujo de bits entrante y luego lo descompone y
30 analiza, con el fin de extraer y calcular algunas características. Estas características constituyen entradas al módulo responsable de la sincronización temporal del flujo de bits de vídeo con la secuencia de vídeo de salida.

Por otra parte, el dispositivo decodificador, por ejemplo el receptor externo (STB, del inglés "set-top-box"), descodifica el flujo de bits recibido y genera la secuencia de vídeo procesada (PVS, por sus siglas en inglés), que se
35 visualiza mediante el dispositivo de salida. También se envía la PVS como entrada al módulo responsable de la sincronización temporal, de manera que se pueda alinear temporalmente con el flujo de bits de vídeo.

En general, el motivo principal de desalineación temporal entre el flujo de bits y la PVS es el retardo. Cuando se transmite el flujo de vídeo a través de una red "de mejor servicio posible" (en inglés, "best-effort network"), como Internet, el tiempo de llegada de cada paquete no es constante, y puede variar significativamente. La variabilidad en
40 el tiempo de la latencia de los paquetes a través de una red se denomina fluctuación (en inglés, "jitter"). Para asegurar una reproducción regular de la secuencia, sin sacudidas, la mayoría de los sistemas de vídeo emplean una memoria intermedia (en inglés, "buffer") contra la fluctuación. El flujo de bits recibido se escribe en la memoria intermedia de entrada basándose en el tiempo de llegada de cada paquete, y los datos de imagen correspondientes a un fotograma se leen desde la misma hacia el decodificador a intervalos de tiempo predeterminados
45 correspondientes al período entre fotogramas. La temporización de visualización de cada imagen viene determinada por el campo de indicación de tiempo (en inglés, "timestamp") registrado en la cabecera del paquete. Es decir, el valor de indicación de tiempo corresponde al período de tiempo de retardo que transcurre desde la detección del código de comienzo de la imagen hasta el momento de la visualización de la imagen.

50 En el sistema de decodificación de vídeo descrito en lo que antecede, la temporización de visualización de cada fotograma de vídeo se determina en función de los datos que se incluyen en el flujo de bits de vídeo para determinar la temporización de visualización. Puesto que el tiempo de visualización de un fotograma no es fijo, no siempre se puede hacer concordar exactamente la PVS con el flujo de bits original.

55 En la bibliografía, el problema de la alineación temporal entre una fuente y una secuencia de vídeo distorsionada ya se ha estudiado con anterioridad, y se le denomina también "registro de vídeo" (en inglés, "video registration"). En M. Barkowsky, R. Bitto, J. Bialkowski y A. Kaup, "Comparison of matching strategies for temporal frame registration in the perceptual evaluation of video quality", *Proc. of the Second International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, enero de 2006, se presenta y se examina, en términos de
60 prestaciones y complejidad, una comparación entre la puesta en concordancia de bloques y la correlación de fase para el registro de vídeo. Además, en Y.Y. Lee, C.S. Kim y S.U. Lee, "Video frame-matching algorithm using dynamic programming", *Journal of Electronic Imaging*, SPIE, 2009, se ha presentado un algoritmo de concordancia de fotogramas para gestionar eliminaciones, inserciones y alteraciones del orden de fotogramas, así como compresión de datos, basado en la minimización de una función de coste de concordancia utilizando programación
65 dinámica. En J. Lu, "Fast video temporal alignment estimation" (documento US-B1-6 751 360) se ha presentado un método para estimación rápida de alineación temporal, destinado a alinear temporalmente un vídeo distorsionado y

una fuente de vídeo correspondiente, con la finalidad de medir la calidad de vídeo. Se transforma cada secuencia de vídeo en una curva de firma, mediante el cálculo de un punto de datos para cada fotograma, en forma de una correlación cruzada entre dos fotogramas seguidos. A continuación, se determina la desalineación temporal del vídeo distorsionado al hallar el valor máximo de la correlación cruzada normalizada entre las curvas de firma de las secuencias de vídeo examinadas. En H. Cheng, "Video registration based on local prediction errors" (documento US-B2-7 366 361) se describe otro método para identificar la correspondencia espacial, temporal y de histogramas entre dos secuencias de vídeo. Se alinea la PVS con la secuencia de vídeo de referencia mediante la generación de una cartografía (en inglés, "mapping") desde un conjunto seleccionado de uno o más fotogramas originales hasta el conjunto procesado, de modo que cada cartografía minimiza un error de predicción local. En K. Ferguson, "Systems and methods for robust video temporal registration" (documento US-A-2008/0253689) se generan mediciones de destilación de fotograma y subimagen a partir de las secuencias de vídeo de referencia y de prueba. Después se alinean linealmente las mismas utilizando el coeficiente local de correlación cruzada de Pearson entre fotogramas. Por último, en C. Souchard, "Spatial and temporal alignment of video sequences" (documento US-A-2007/0097266) se define una función de movimiento para describir el movimiento de un conjunto de píxeles entre los fotogramas de las secuencias de vídeo de prueba y de referencia, y se utiliza una transformada para alinear las dos imágenes.

Compendio de la Invención

El objeto de la invención es proporcionar un método y aparato para alinear temporalmente el flujo de bits de vídeo de entrada de un descodificador de vídeo, por ejemplo un receptor externo (STB), con la secuencia de vídeo procesada (PVS) y hacer posible su sincronización. Este objeto se consigue con las características de las reivindicaciones.

Según un primer aspecto, un ejemplo proporciona un método para sincronizar temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el método comprende los pasos de:

- a) capturar el flujo de bits de entrada y suministrarlo a un analizador de flujo de bits;
- b) examinar y analizar, mediante el analizador de flujo de bits, el flujo de bits de entrada capturado;
- c) descodificar el flujo de bits de entrada y proporcionar a partir del mismo imágenes reconstruidas N_{rec} ;
- d) almacenar las imágenes reconstruidas N_{rec} en una memoria intermedia; y
- e) comparar una secuencia de vídeo procesada recibida desde el lado de salida del descodificador de vídeo con cada una de las imágenes reconstruidas almacenadas, para encontrar un fotograma concordante n^* del conjunto de imágenes reconstruidas de la memoria intermedia.

El método puede comprender además el paso f) de suministrar el fotograma concordante n^* y su parte correspondiente del flujo de bits a una unidad de evaluación de la calidad de vídeo. La unidad de evaluación de la calidad de vídeo puede estimar una puntuación subjetiva de la calidad basándose en la información obtenida a partir del flujo de bits y la secuencia de vídeo procesada.

El paso b) del método puede comprender, además, extraer para cada imagen, a partir del flujo de bits, los paquetes que se han perdido debido a errores de transmisión y las zonas correspondientes dentro de cada fotograma que han resultado afectadas por la pérdida de paquetes; así como un mapa de propagación de error basado en los vectores de movimiento y tipos de macrobloques, que indique cuáles son los píxeles de cada fotograma que hacen referencia a una zona errónea conforme a su predicción y es posible, por lo tanto, que hayan resultado afectados por la pérdida de paquetes.

Además, el paso e) puede comprender calcular una medida de distancia entre la secuencia de vídeo procesada y la imagen reconstruida de la memoria intermedia, mediante la explotación de la información extraída del flujo de bits y de la PVS. La información extraída del flujo de bits puede comprender la localización espacial y temporal de las degradaciones debidas a la pérdida de paquetes.

Según un segundo aspecto, la invención proporciona un método para sincronizar temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el método comprende los pasos de:

- a) capturar el flujo de bits de entrada y suministrarlo a un analizador de flujo de bits;
- b) analizar, mediante el analizador de flujo de bits, el flujo de bits capturado, a fin de generar mapas de error $I(x,y,t)$ y un conjunto A de píxeles que han resultado afectados por la pérdida de paquetes, en donde $I(x,y,t)$ representa el valor del mapa de error en la posición (x,y) del fotograma t -ésimo, $x = 0, 1, \dots, M$, e $y = 0, 1, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo;
- c) almacenar, en una memoria intermedia, los mapas de error de fotogramas reconstruidos N_{rec} , que están siendo reconstruidos por el analizador de flujo de bits; y
- d) comparar los mapas de error almacenados de los N_{rec} fotogramas reconstruidos, con el mapa de error correspondiente de la secuencia de vídeo procesada en curso, para encontrar un fotograma concordante n^* de entre el conjunto de fotogramas reconstruidos de la memoria intermedia que concuerde con la imagen en curso de la secuencia de vídeo procesada.

El método puede comprender el paso e) de suministrar el fotograma concordante n^* y su correspondiente parte del flujo de bits a una unidad de evaluación de la calidad de vídeo.

5 Se prefiere, además, que entre los pasos c) y d) se lleven a cabo los siguientes pasos:

extraer los bordes de la imagen en curso de la secuencia de vídeo procesada;
 calcular el contorno de los bordes que probablemente pertenezcan a una zona afectada por la pérdida de paquetes;
 10 comparar el contorno de borde de la secuencia de vídeo procesada con el contorno de borde de los mapas de error almacenados en la memoria intermedia.

Según un tercer aspecto, un ejemplo adicional proporciona un método para sincronizar temporalmente el flujo de bits de entrada encriptado de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el método comprende los pasos de:

- 15
- a) suministrar el flujo de bits de entrada encriptado a un módulo de análisis;
 - b) analizar, en el módulo de análisis, la información de cabecera de paquete del flujo de bits de entrada encriptado, y calcular el tamaño y tipo de los fotogramas contenidos en el flujo de bits;
 - 20 c) extraer, basándose en los tamaños y tipos calculados de los fotogramas, características que son representativas del contenido representado en los fotogramas de vídeo;
 - d) extraer, a partir de las imágenes de la secuencia de vídeo procesada, el tipo de contenido y las respectivas características representativas del contenido;
 - e) comparar las características extraídas en el paso d) a partir de la imagen en curso de la secuencia de vídeo procesada con las características extraídas en el paso c) a partir de los fotogramas analizados, para
 - 25 encontrar un fotograma concordante n^* .

Las características extraídas en los pasos c) y d) son, preferiblemente, características variables en el tiempo, y estos pasos se pueden realizar en el campo de los píxeles.

30 Según una realización preferida, la característica y la alineación temporal para el fotograma concordante n^* se envían adicionalmente como entrada a un módulo de evaluación de la calidad de vídeo.

El paso b) puede detectar además fotogramas que hayan resultado afectados por la pérdida de paquetes, y puede comprender además los pasos de:

- 35
- f) examinar la imagen en curso de la secuencia de vídeo procesada en busca de artefactos generados debido a las pérdidas de paquetes; y
 - g) poner en concordancia la secuencia de vídeo procesada con el fotograma correspondiente del flujo de bits, en donde, si en el paso f) se ha detectado que la imagen en curso de la secuencia de vídeo procesada
 - 40 contiene artefactos que se insertaron como resultado de una pérdida de paquetes, se pone en concordancia esa imagen con la parte del flujo de bits correspondiente al fotograma en el que se encontró en el paso b) que contenía la pérdida de paquetes.

Un aspecto adicional de un ejemplo se refiere a un aparato para alinear temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el aparato comprende:

45 un analizador de flujo de bits que recibe el flujo de bits de entrada, en donde el analizador de flujo de bits está configurado para examinar y analizar el flujo de bits de entrada, y

50 un módulo de sincronización que está configurado para alinear el flujo de bits analizado con la secuencia de vídeo procesada, en donde el módulo de sincronización comprende:

55 un descodificador que recibe desde el analizador de flujo de bits el flujo de bits analizado, y está configurado para proporcionar a partir del mismo imágenes reconstruidas N_{rec} ,
 una memoria intermedia configurada para almacenar dichas imágenes reconstruidas N_{rec} de dicho descodificador, de manera que la imagen de salida representada por la secuencia de vídeo procesada puede compararse con este número específico de imágenes previamente reconstruidas, y
 una unidad de comparación en el campo de los píxeles que recibe el flujo de bits analizado desde el

60 analizador de flujo de bits, las imágenes reconstruidas desde la memoria intermedia y la secuencia de vídeo procesada, en donde la unidad de comparación está configurada para

65 comparar la secuencia de vídeo procesada con cada una de las imágenes reconstruidas en la memoria intermedia, y
 determinar la imagen más concordante de la memoria intermedia de vídeo.

El aparato comprende preferiblemente un módulo de evaluación de la calidad de vídeo que combina información del flujo de bits y la secuencia de vídeo procesada, con el fin de evaluar la calidad de la secuencia de vídeo. La unidad de comparación puede estar configurada además para calcular una medida de distancia entre la secuencia de vídeo procesada y la imagen reconstruida de la memoria intermedia, examinada, mediante la explotación de información extraída del flujo de bits. Además, la unidad de comparación está configurada para emitir como salida hacia el módulo de evaluación de la calidad de vídeo la imagen más concordante, con el fin de estimar una puntuación subjetiva de la calidad basada en la información del flujo de bits y la secuencia de vídeo procesada.

Se prefiere, además, que el analizador de flujo de bits esté configurado para extraer para cada imagen, a partir del flujo de bits, los paquetes que se han perdido debido a errores de transmisión y las zonas correspondientes dentro de cada fotograma que se han visto afectadas por la pérdida de paquetes; así como un mapa de error basado en los vectores de movimiento y tipos de macrobloques, que indique cuáles son los píxeles de cada fotograma que hacen referencia a una zona errónea conforme a su predicción y es posible, por lo tanto, que hayan resultado afectados por la pérdida de paquetes.

Según un aspecto adicional, la invención proporciona un aparato para alinear temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el aparato comprende:

un analizador de flujo de bits que recibe el flujo de bits de entrada, en donde el analizador de flujo de bits está configurado para analizar el flujo de bits de entrada en busca de información de cabecera de paquetes del flujo de bits encriptado, y calcular el tamaño y tipo de los fotogramas contenidos en el flujo de bits, y un módulo de sincronización que está configurado para alinear el flujo de bits analizado con la secuencia de vídeo procesada, en donde el módulo de sincronización comprende:

un módulo de extracción que recibe desde el analizador de flujo de bits el flujo de bits analizado, y está configurado para generar mapas de error $I(x,y,t)$ y un conjunto A de píxeles que han resultado afectados por la pérdida de paquetes, en donde $I(x,y,t)$ representa el valor del mapa de error en la posición (x,y) del fotograma t -ésimo, $x = 0, 1, \dots, M$, e $y = 0, 1, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo, una memoria intermedia configurada para almacenar los mapas de error de fotogramas reconstruidos N_{rec} , que están siendo reconstruidos por el analizador de flujo de bits, y un módulo de búsqueda de patrones de error que recibe los mapas de error de los fotogramas reconstruidos N_{rec} desde la memoria intermedia, y el mapa de error correspondiente de la secuencia de vídeo procesada, en donde el módulo de búsqueda de patrones de error está configurado para

comparar los mapas de error almacenados de los N_{rec} fotogramas reconstruidos, con el mapa de error correspondiente de la secuencia de vídeo procesada en curso, y para determinar un fotograma concordante n^* a partir del conjunto de fotogramas reconstruidos de la memoria intermedia que concuerda con la imagen en curso de la secuencia de vídeo procesada.

Según un aspecto adicional, un ejemplo proporciona un aparato para alinear temporalmente el flujo encriptado de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el aparato comprende:

un analizador de flujo de bits que recibe el flujo de bits de entrada encriptado, en donde el analizador de flujo de bits está configurado para examinar y analizar el flujo de bits de entrada, y un módulo de sincronización que está configurado para alinear el flujo de bits analizado con la secuencia de vídeo procesada, caracterizado porque el módulo de sincronización comprende:

un primer módulo de extracción que está configurado para extraer, basándose en los tamaños y tipos calculados de los fotogramas, características que son representativas del contenido representado en los fotogramas de vídeo, un segundo módulo de extracción que está configurado para extraer el tipo de contenido y características respectivas, representativas del contenido, a partir de las imágenes de la secuencia de vídeo procesada, y una unidad de comparación conectada a la primera unidad de extracción y a la segunda unidad de extracción, en donde la unidad de comparación está configurada para

comparar las características extraídas de la imagen en curso de la secuencia de vídeo procesada con las características extraídas de los fotogramas analizados, y para determinar el fotograma concordante n^* .

Volviendo al objeto de la invención, en términos generales se puede conseguir, mediante la provisión de un dispositivo (sonda) que, según la presente invención, incluye un analizador de flujo de bits para examinar y descodificar el flujo de bits de entrada, un módulo de sincronización que es responsable de alinear temporalmente el flujo de bits con la PVS.

El dispositivo responsable de la alineación temporal del flujo de bits de vídeo y la PVS comprende una memoria intermedia que almacena las imágenes reconstruidas que constituyen la salida del descodificador. Esta memoria intermedia debe ser capaz de almacenar un número específico de imágenes reconstruidas N_{rec} procedentes del dispositivo descodificador dentro de la sonda, de manera que se pueda comparar la imagen de salida con este número específico de imágenes previamente reconstruidas.

El módulo responsable de la sincronización de la PVS y el flujo de bits de vídeo realiza las siguientes operaciones:

- a) compara la PVS con cada una de las imágenes reconstruidas de la memoria intermedia;
- b) calcula una medida de distancia entre la PVS y la imagen reconstruida, examinada, de la memoria intermedia mediante la explotación de la información extraída del flujo de bits (por ejemplo, la ubicación espacial y temporal de degradaciones debidas a la pérdida de paquetes);
- c) determina la imagen más concordante de la memoria intermedia y la emite como salida al módulo de evaluación de la calidad de vídeo que es responsable de estimar la puntuación subjetiva $M\acute{O}S$ basándose en información de la secuencia de bits y la PVS.

A continuación, se describirá con más detalle la invención.

En primer lugar se describirá, con referencia a la Figura 2, un primer ejemplo de sincronización temporal entre el flujo de bits y la PVS.

La sincronización temporal entre el flujo de bits y la secuencia de vídeo descodificada/procesada consta de dos pasos. Inicialmente, en el primer paso el aparato según la invención, es decir, un dispositivo de sondeo, captura el flujo de bits y lo alimenta al analizador de flujo de bits que, para cada imagen, extrae la siguiente información:

- a. los paquetes que se han perdido debido a errores de transmisión y las zonas correspondientes dentro de cada fotograma que han resultado afectadas por la pérdida de paquetes; y
- b. un mapa de propagación de error basado en los vectores de movimiento y tipos de macrobloques, que indica cuáles son los píxeles de cada fotograma que hacen referencia (para predicción) a una zona errónea conforme a su predicción y es posible, por lo tanto, que hayan resultado afectados por la pérdida de paquetes.

Sea A el conjunto de píxeles que han resultado afectados por la pérdida de paquetes y los píxeles que hacen referencia a píxeles perdidos por su predicción. Además, sea $I(x,y,t)$ el valor del mapa de error en la posición (x,y) del fotograma t -ésimo, $x = 1, 2, \dots, M$, e $y = 1, 2, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo. El valor para los píxeles que pertenecen al conjunto A es uno, en caso contrario su valor es cero. Por tanto:

$$I(x,y,t) = \begin{cases} 1, & (x,y) \notin A \\ 0, & (x,y) \in A \end{cases} \quad \text{Eq. (1)}$$

En otras palabras, el conjunto A contiene todos los píxeles que no van a ser idénticos entre la secuencia de vídeo procesada y la imagen reconstruida del descodificador dentro de la sonda debido a las diferentes técnicas de ocultación de error. Por esta razón, todos los píxeles que pertenecen al conjunto A se marcan como no disponibles para la alineación temporal en el paso siguiente. En la Figura 5 se representa un ejemplo de la obtención de los mapas de propagación de error. La pérdida de paquetes se produce en el primer fotograma (en la Figura 5, la zona oscura a la derecha de la imagen indica los macrobloques que se han deteriorado debido a la pérdida de paquetes) y los píxeles de los fotogramas siguientes que hacen referencia a la zona afectada del primer fotograma para el proceso de predicción se calculan y se marcan como no disponibles para el proceso de sincronización. Esto ayudará a eliminar aquellos píxeles que pueden ser diferentes entre las imágenes reconstruidas y la secuencia de vídeo procesada debido a la aplicación de diferentes técnicas de ocultamiento de error entre el descodificador de vídeo interno de la sonda y el descodificador del dispositivo de descodificación, por ejemplo el STB.

Además, el descodificador separado dentro del dispositivo de sondeo descodifica el flujo de bits recibido, y las imágenes reconstruidas se almacenan en una memoria intermedia. El tamaño de la memoria intermedia debe ser lo suficientemente grande como para tener en cuenta el retardo en que pueda incurrir el STB. Sea d (en segundos) el

retardo máximo esperado y sea f la tasa de fotogramas de la secuencia de vídeo. En consecuencia, la memoria intermedia debería ser capaz de almacenar $N_{rec} = d \cdot f$ fotogramas en total.

En el segundo paso se compara la secuencia de vídeo procesada con cada una de las imágenes reconstruidas de la memoria intermedia. El propósito de esta comparación es encontrar un fotograma concordante n^* a partir del conjunto de imágenes reconstruidas de la memoria intermedia, basándose en la siguiente ecuación:

$$n^* = \arg \min_n \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |F'(x, y, t) \cdot I(x, y, n) - F(x, y, n) \cdot I(x, y, n)| \quad \text{Eq. (2)}$$

en donde $F'(x,y,t)$ es el valor de píxel en la posición (x,y) del fotograma t -ésimo de la imagen de vídeo procesada en examen, y $F(x,y,n)$ es el píxel de la misma posición del fotograma n -ésimo de la memoria intermedia de imágenes reconstruidas, $n = 0, 1, \dots, N_{rec}$. Se aplica la multiplicación de los valores de píxel de cada imagen con el mapa de la propagación de error, con el fin de excluir aquellos píxeles que han resultado afectados por la pérdida de paquetes (ya sea directamente o bien debido a la propagación de error) y asegurar que el proceso de concordancia se limita a aquellos píxeles que es probable que sean más similares entre las dos imágenes. Hay que señalar también que, en lugar de la función específica (que se denomina suma de diferencias absolutas, SDA) utilizada aquí para la minimización, se podría utilizar cualquier función similar, por ejemplo, el error cuadrático medio.

Después de terminado el proceso de puesta en concordancia, se elimina de la memoria intermedia de las imágenes reconstruidas el fotograma concordante n^* y, por ejemplo, se pueden usar además el fotograma puesto en concordancia, junto con su correspondiente parte del flujo de bits, como entrada para el módulo responsable de la evaluación de la calidad del vídeo híbrido.

A continuación, se describirá una segunda realización de la sincronización entre el flujo de bits de vídeo y la secuencia de vídeo descodificada.

En la segunda realización de la presente invención, la sincronización entre el flujo de bits y la secuencia de vídeo procesada se lleva a cabo analizando el flujo de bits, extrayendo información acerca de la zona de cada fotograma que ha resultado afectada por la pérdida de paquetes, y examinando después si aparece el mismo patrón de errores en la secuencia de vídeo procesada. En la Figura 3 se representa el diagrama de bloques del procedimiento general.

En el primer paso, el analizador de flujo de bits interno del dispositivo de sondeo procesa y analiza el flujo de bits a fin de generar (para cada fotograma) el mapa del error $I(\dots)$ y, de este modo, el conjunto A de píxeles que han resultado afectados por la pérdida de paquetes, como se ha definido más arriba en el contexto de la primera realización. Los mapas de error de N_{rec} fotogramas se almacenan en una memoria intermedia de modo que más tarde se pueden poner en concordancia con los mapas de error correspondientes de la secuencia de vídeo procesada, como se explica a continuación.

En el paso siguiente, se estiman los macrobloques de la imagen en curso de la secuencia de vídeo procesada que se han deteriorado debido a una pérdida de paquetes y se generan el mapa de error $I_{PVS}(\dots)$ y el conjunto A_{PVS} , como se ha definido más arriba en el contexto de la primera realización. A continuación se estima el contorno de los píxeles del mapa de error $I_{PVS}(\dots)$. El método para estimar los macrobloques deteriorados de la imagen en curso y el contorno de los píxeles pertenecientes a estos macrobloques está fuera del alcance de esta invención, y por lo tanto se podría utilizar cualquier método conocido por un experto en la materia, por ejemplo el método propuesto en G. Valenzise *et al.*, "Estimating channel-induced distortion in H.264/AVC vídeo without bitstream information", QoMEX, 2010.

El borde del contorno de la secuencia de vídeo procesada debe compararse con el borde del contorno de los mapas de error que se han almacenado previamente en la memoria intermedia. Si la concordancia de los contornos se ve limitada en la zona definida por el mapa de error que está constituida por los píxeles que pertenecen al conjunto A , no se toman en consideración los bordes del contorno detectado de la secuencia de vídeo procesada. Asimismo, dado que típicamente se producen leves errores de interpretación de la zona detectada como errónea y la zona correspondiente que se detecta como errónea de la secuencia de vídeo procesada, se incrementa en k píxeles el área de búsqueda para asegurar que los bordes del contorno detectados por el descodificador no son menores que los detectados en la secuencia de vídeo procesada. En esta realización, se elige que k sea igual a 5. Sin embargo, se podría elegir cualquier otro valor.

Así se elabora el conjunto A_{PVS} de píxeles que pertenecen a la zona en donde se han detectado los bordes del contorno debido a la pérdida de paquetes (incrementada como se ha señalado más arriba), y se genera el mapa de error I_{PVS} . Sea $I_{PVS}(x,y,t)$ el valor de mapa de error de la t -ésima imagen de secuencia de vídeo procesada en la posición (x,y) , $x = 1, 2, \dots, M$, e $y = 1, 2, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo. Entonces:

5

$$I_{PVS}(x, y, t) = \begin{cases} 1, & (x, y) \notin A_{PVS} \\ 0, & (x, y) \in A_{PVS} \end{cases} \quad \text{Eq. (3)}$$

15

En el tercer paso, se compara el mapa de error de la imagen en curso de la secuencia de vídeo procesada con cada uno de los mapas de error (que se han generado analizando el flujo de bits en el primer paso) que se almacenan en la memoria intermedia. El propósito de esta comparación es encontrar un fotograma concordante n^* , a partir del conjunto de imágenes reconstruidas de la memoria intermedia, que concuerde con la imagen en curso de la secuencia de vídeo procesada en base a la siguiente ecuación:

20

$$n^* = \arg \min_n f(I(x, y, n), I_{PVS}(x, y, t)) \quad \text{Eq. (4)}$$

30

donde $I_{PVS}(x, y, t)$ es el valor en la posición (x, y) del t -ésimo fotograma del mapa de error de la PVS en examen e $I(x, y, n)$ es el píxel en la misma posición del mapa de error correspondiente al n -ésimo fotograma, $n = 1, 2, \dots, N_{rec}$.

35

Puede estar presente por último, como se muestra en la Figura 3, un módulo híbrido de evaluación de la calidad de vídeo que combine información del flujo de bits y la secuencia de vídeo procesada para evaluar la calidad de la secuencia de vídeo.

40

La función $f(\dots)$ es una función que se puede utilizar para determinar la similitud entre los dos mapas de error basándose en la concordancia de patrones de contorno. El algoritmo de concordancia de patrones basado en contorno está fuera del alcance de la presente invención, y se podría utilizar para este fin cualquier método relevante, por ejemplo el método de T. Adamek y N. O'Connor, "Efficient contour-based shape representation and matching", Multimedia Information Retrieval, 2003.

45

A continuación, se describirá la sincronización temporal entre el flujo de bits y la imagen descodificada de una secuencia de vídeo encriptada según una realización adicional de la invención.

El problema de la sincronización de fotogramas se hace más evidente cuando la secuencia de vídeo está encriptada y la única información disponible para el flujo de bits se puede extraer de las cabeceras de paquete. Los parámetros que se pueden extraer del flujo de bits basándose sólo en la inspección de las cabeceras de paquete son el tamaño y el tipo de cada fotograma.

50

Alineación basada en concordancia basada en fotogramas

El primer ejemplo del método para sincronización entre el flujo de bits de vídeo y la imagen descodificada de una secuencia de vídeo encriptada aborda el caso en donde no existen pérdidas de paquetes en el flujo de bits, y se muestra en la Figura 4.

55

En el primer paso, el flujo de bits de vídeo se alimenta a un módulo que es responsable del análisis de la información de cabecera de paquete y del cálculo del tamaño y tipo de los fotogramas (intra, predictivos o bipredictivos). El método para extraer el tipo de fotogramas del vídeo encriptado basándose en su tamaño está fuera de alcance de esta patente. Por ejemplo, se podría utilizar el método propuesto en la solicitud de patente europea n.º 10 17 1431.9 del presente solicitante, presentada el 30 de julio de 2010 (incorporada a la presente memoria por referencia).

60

En el segundo paso, la información sobre los tamaños y tipos de fotogramas se utiliza para extraer (por ejemplo, de manera variable en el tiempo) características que son representativas del contenido representado en los fotogramas de vídeo. Es evidente que puesto que no está disponible, debido a la encriptación, la información de carga útil, solamente se puede estimar una aproximación de la clase de contenido. El método para estimar la clase de contenido basándose en tamaños de fotogramas y tipos de fotogramas está fuera del alcance de la presente invención. Se podría utilizar cualquier método disponible para este propósito, por ejemplo el de A. Clark, "Method and system for content estimation of packet video streams", documento WO 2009/012297 A1.

70

Análogamente, de las imágenes de la secuencia de vídeo procesada se extraen el tipo de contenido y características respectivas (por ejemplo, variables con el tiempo) representativas del contenido. Puesto que la presente invención se centra sólo en los aspectos de sincronización entre el flujo de bits y la secuencia de vídeo procesada, el método para la extracción de características de la imagen (en el campo de los píxeles) y la detección de tipo de contenido no está dentro de su alcance.

75

En el tercer paso, las características extraídas de la imagen en curso de la secuencia de vídeo procesada se ponen en concordancia con las características extraídas a partir de los fotogramas que fueron analizados por el módulo de

análisis de cabeceras de paquete. Nótese que se pueden utilizar la característica correspondiente y la alineación temporal como entradas al módulo responsable de la evaluación de la calidad de vídeo. Tal como se muestra en la Figura 4, puede estar presente un módulo híbrido de evaluación de la calidad de vídeo que combine información procedente del flujo de bits y la secuencia de vídeo procesada con el fin de evaluar la calidad de la secuencia de vídeo.

Alineación temporal basada en el desencadenamiento de evento de pérdida de paquete

En este ejemplo se supone que el flujo de bits y la secuencia de vídeo procesada están sincronizados cuando no se producen errores (por ejemplo, debido a pérdidas de paquete), basándose en el método descrito en la subsección precedente, y se considera el caso en donde se requiere sincronización cuando se detecta una pérdida de paquete.

En el primer paso, se alimenta el flujo de bits de vídeo al módulo responsable del análisis de la información de cabecera de paquete, del cálculo de los tipos de fotograma y de la detección de los fotogramas que han resultado afectados por la pérdida de paquetes. Se detecta el tipo de todos los fotogramas y así se determina si una pérdida de paquetes se propaga o no, debido al proceso de predicción, a sus fotogramas subsiguientes.

En el siguiente paso, se examina la imagen en curso de la secuencia de vídeo procesada en búsqueda de distorsiones que pudieran haber sido generadas por pérdidas de paquetes. Hay que advertir que este algoritmo debe ser capaz de distinguir entre las distorsiones causadas por pérdidas de paquetes y las distorsiones producidas como consecuencia de la compresión, con pérdida, de la secuencia de vídeo original. El algoritmo para la detección y clasificación de las distorsiones está fuera del alcance de la presente invención. Por ejemplo, se podría utilizar para este propósito el método de G. Valenzise *et al.*, "Estimating channel-induced distortion in H.264/AVC video without bitstream information", QoMEX, 2010.

En el tercer paso, se realiza la puesta en concordancia de la secuencia de vídeo procesada con el fotograma correspondiente del flujo de bits. Si se ha detectado que la imagen en curso de la secuencia de vídeo procesada contiene distorsiones que se han originado como consecuencia de una pérdida de paquetes, se pone en concordancia esa imagen con la parte del flujo de bits correspondiente al fotograma que ha resultado (en el paso inicial) contener la pérdida de paquetes.

Por último, se calcula la información de sincronización (es decir, el fotograma en curso de la PVS junto con su parte correspondiente del flujo de bits de vídeo). Se puede enviar esta información, como entrada, al módulo responsable de la evaluación de la calidad de vídeo.

La técnica anterior mencionada en la parte introductoria precedente considera que el problema de la alineación temporal entre dos secuencias de vídeo de debe al salto, repetición, supresión, alteración del orden y compresión de datos de los fotogramas que pueden producirse a causa de la transmisión de la secuencia de vídeo a través de una red errónea o por el procesamiento de la secuencia. Por el contrario, la presente invención sólo tiene en cuenta la sincronización de la secuencia de vídeo procesada con el flujo de bits en el lado del usuario final y, por lo tanto, la única fuente de desalineación es el retardo introducido por el dispositivo descodificador y el analizador de flujo de bits y las diferentes técnicas de ocultamiento de error aplicadas por el descodificador (STB) y el descodificador dentro del dispositivo de sondeo.

La Figura 1 muestra un dispositivo de sondeo convencional para la alineación temporal de la secuencia de vídeo procesada con el flujo de bits y los datos de imagen descodificada;
 la Figura 2 muestra un diagrama de bloques para la sincronización del flujo de bits con la imagen de salida, según un primer ejemplo;
 la Figura 3 muestra un diagrama de bloques de una segunda realización preferida de la presente invención para la sincronización entre el flujo de bits de vídeo y la secuencia de vídeo procesada;
 la Figura 4 muestra la sincronización entre el flujo de bits y la imagen descodificada de una secuencia de vídeo encriptada, según un ejemplo adicional; y
 la Figura 5 un ejemplo de mapas de propagación de error para fotogramas subsiguientes.

Aunque se ha ilustrado y descrito con detalle la invención en los dibujos y la descripción precedente, dicha ilustración y descripción han de considerarse ilustrativas o ejemplares y no restrictivas. Se entenderá que los expertos pueden realizar cambios y modificaciones dentro del alcance de las reivindicaciones que siguen.

Además, en las reivindicaciones la palabra "comprende" no excluye otros elementos o pasos, y el artículo indefinido "un" o "una" no excluye una pluralidad. Una sola unidad puede cumplir las funciones de varias características mencionadas en las reivindicaciones. Los términos "esencialmente", "alrededor de", "aproximadamente" y similares, referidos a un atributo o un valor en particular también definen exactamente el atributo o exactamente el valor, respectivamente. Ningún signo de referencia en las reivindicaciones debe interpretarse como limitante del alcance.

REIVINDICACIONES

1. Método para sincronizar temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el método comprende los pasos de:

- 5 a) capturar el flujo de bits de entrada y suministrarlo a un analizador de flujo de bits;
 b) analizar, mediante el analizador de flujo de bits, el flujo de bits capturado, a fin de generar mapas de error $I(x,y,t)$ y un conjunto A de píxeles que han resultado afectados por la pérdida de paquetes, en donde $I(x,y,t)$ representa el valor del mapa de error en la posición (x,y) del fotograma t -ésimo, $x = 0, 1, \dots, M$, e $y = 0, 1, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo;
 10 c) almacenar, en una memoria intermedia, los mapas de error de N_{rec} fotogramas, generados internamente por el analizador de flujo de bits; y
 d) comparar los mapas de error almacenados de los N_{rec} fotogramas reconstruidos con el mapa de error correspondiente de la secuencia de vídeo procesada en curso, con el fin de encontrar un fotograma concordante n^* de entre el conjunto de fotogramas reconstruidos de la memoria intermedia que concuerde con la imagen en curso de la secuencia de vídeo procesada.
 15

2. El método según la reivindicación 1, que comprende además el paso e) de suministrar el fotograma concordante n^* y su correspondiente parte del flujo de bits a una unidad de evaluación de la calidad de vídeo.
 20

3. El método según la reivindicación 1, en donde entre los pasos c) y d) se llevan a cabo los siguientes pasos:

- 25 estimar los macrobloques deteriorados de la imagen en curso de la secuencia de vídeo procesada;
 calcular el contorno de los píxeles que pertenecen a una zona afectada por la pérdida de paquetes;
 comparar el contorno de la secuencia de vídeo procesada con el contorno de los mapas de error almacenados en la memoria intermedia.

4. Aparato para sincronizar temporalmente el flujo de bits de entrada de un descodificador de vídeo con la secuencia de vídeo procesada descodificada por el descodificador de vídeo, en donde el aparato comprende:

- 30 un analizador de flujo de bits que recibe el flujo de bits de entrada, en donde el analizador de flujo de bits está configurado para analizar el flujo de bits de entrada en busca de información de cabecera de paquetes del flujo de bits encriptado, y calcular el tamaño y tipo de los fotogramas contenidos en el flujo de bits, y
 un módulo de sincronización que está configurado para sincronizar el flujo de bits analizado con la secuencia de vídeo procesada,
 en donde el módulo de sincronización comprende:

- 40 un módulo de extracción que recibe desde el analizador de flujo de bits el flujo de bits analizado, y está configurado para generar mapas de error $I(x,y,t)$ y un conjunto A de píxeles que han resultado afectados por la pérdida de paquetes, en donde $I(x,y,t)$ representa el valor del mapa de error en la posición (x,y) del fotograma t -ésimo, $x = 0, 1, \dots, M$, e $y = 0, 1, \dots, N$, donde M, N son, respectivamente, las dimensiones horizontales y verticales de cada fotograma de la secuencia de vídeo,
 una memoria intermedia configurada para almacenar los mapas de error de N_{rec} fotogramas, generados internamente por el analizador de flujo de bits, y
 45 un módulo de búsqueda de patrones de error que recibe los mapas de error de los fotogramas reconstruidos N_{rec} , desde la memoria intermedia, y el mapa de error correspondiente de la secuencia de vídeo procesada, en donde el módulo de búsqueda de patrones de error está configurado para

- 50 comparar los mapas de error almacenados de los N_{rec} fotogramas reconstruidos, con el mapa de error correspondiente de la secuencia de vídeo procesada en curso, y para determinar un fotograma concordante n^* a partir del conjunto de fotogramas reconstruidos de la memoria intermedia que concuerda con la imagen en curso de la secuencia de vídeo procesada.
 55

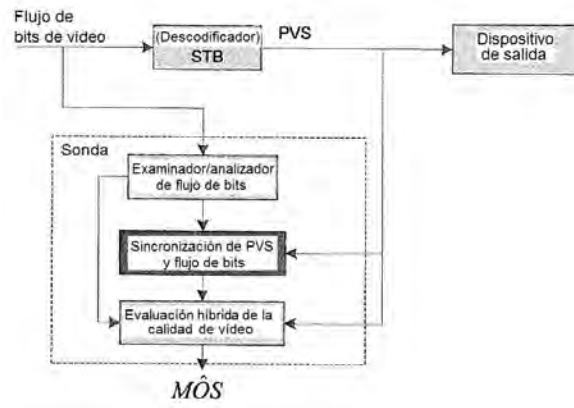


Fig. 1

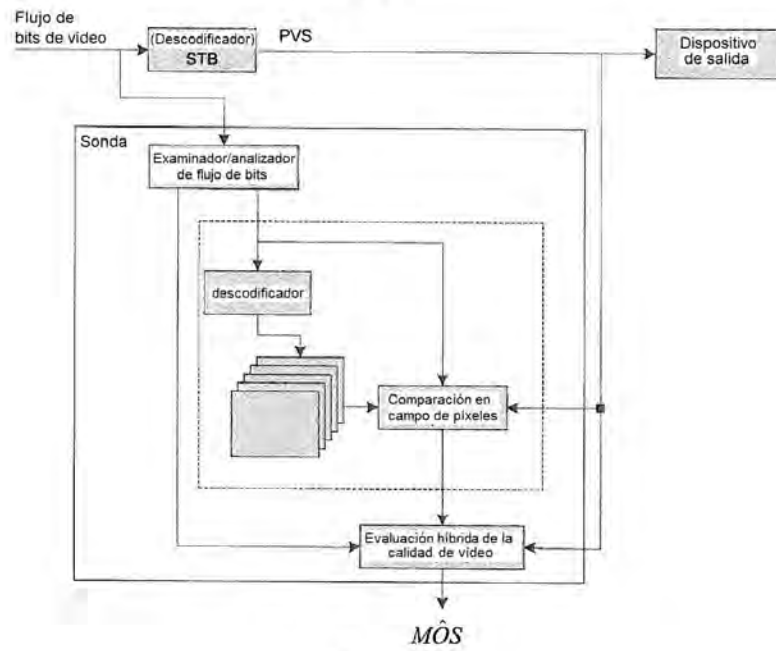


Fig. 2

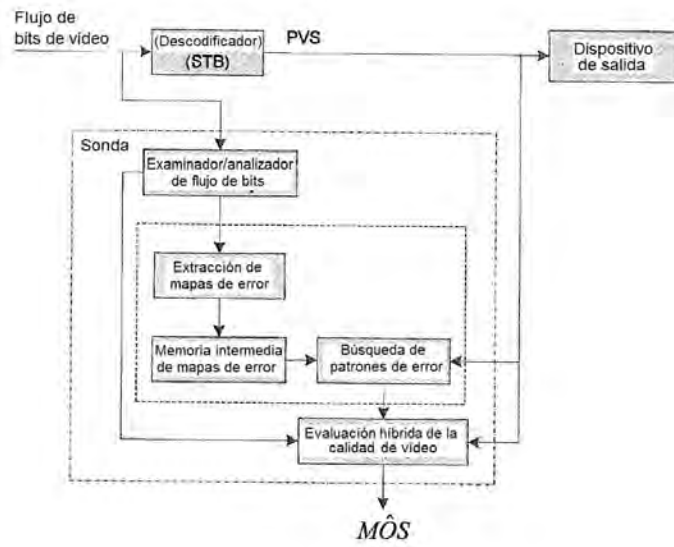


Fig. 3

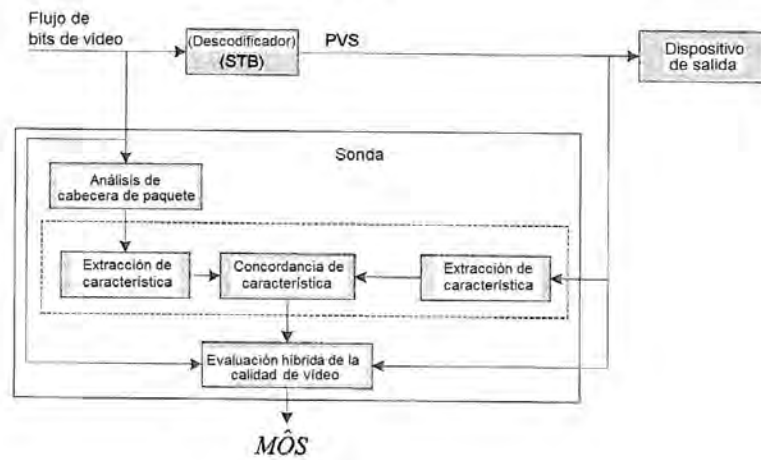


Fig. 4

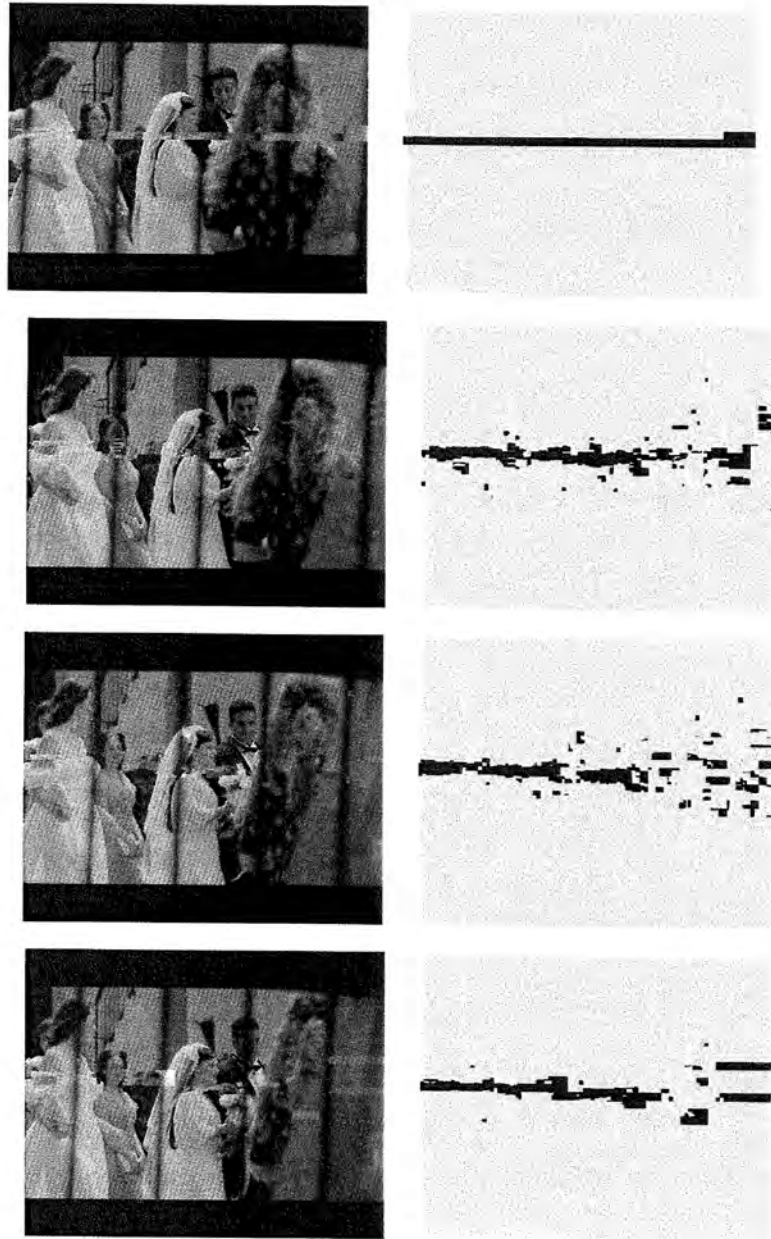


Fig. 5