

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 546 775**

51 Int. Cl.:

G06T 7/20

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.04.2012 E 12164975 (0)**

97 Fecha y número de publicación de la concesión europea: **26.08.2015 EP 2525324**

54 Título: **Método y aparato para la generación de un mapa de profundidad y un vídeo 3D**

30 Prioridad:

20.05.2011 TR 201104918

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

28.09.2015

73 Titular/es:

**VESTEL ELEKTRONIK SANAYI VE TICARET A.S.
(100.0%)**

**Organize Sanayi Bölgesi
45030 Manisa, TR**

72 Inventor/es:

**BASTUG, AHMET;
OZDEMIR, HUSEYIN y
OZER, NURI**

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 546 775 T3

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

DESCRIPCIÓN

Método y aparato para la generación de un mapa de profundidad y un vídeo 3D

La presente invención se refiere a un método y aparato para la generación de un mapa de profundidad a partir de dos cuadros en una secuencia de vídeo 2D y un método y aparato para la generación de un vídeo 3D.

5 Las realizaciones preferidas de la presente invención pueden encontrar una aplicación en 3D compatible con televisores de pantalla plana de LCDs, LED y plasma los cuales requieren la reproducción de contenido (3D) desde una fuente de vídeo o imagen (2D) mono. Se pueden encontrar otros posibles usos en los dispositivos tales como monitores, proyectores, decodificadores. Las realizaciones preferidas de la presente invención pueden encontrar aplicaciones en la industria de fabricación de televisores, el campo de gráficas por ordenador, el sector de cine y aplicaciones médicas (tales como generación de imágenes médicas).

10 El principal problema de los productores de televisión en 3D han encontrado y continuarán encontrando por algún tiempo por venir, es la insuficiencia de los contenidos 3D debido a que es muy costoso rodar una película 3D y preparar la difusión 3D (vídeo estéreo). La solución de este problema es crear un contenido 3D desde contenidos de vídeo 2D (mono) calculando el mapa de profundidad de los cuadros. Se reproduce una segunda secuencia de vídeo desde la primera utilizando un perfil de profundidad estimado en la primera secuencia de vídeo. Las secuencias de vídeo estéreo generadas pueden ser visualizadas en una pantalla 3D una después de la otra en una forma de tiempo compartido o con dos vistas diferentes de una forma intercalada espacialmente. De esta forma, las diferentes vistas de una escena son alimentadas separadamente a los ojos izquierdo y derecho. El problema básico de calcular el mapa de profundidad es separar el fondo y el primer plano (segmentación).

15 Un ejemplo de la estimación de profundidad es dado en el documento técnico "On Creating Depth maps from Monoscopic Vídeo using Structure from Motion": Ping Li, Dirk Farin, Rene Klein Gunnewiek, Peter H. N. de With en el 27 Simposio de Teoría de la Información en Benelux, vol. p. 508-515, Junio 2016. Este documento divulga una técnica llamada Estructura desde el Movimiento (SFM) se discute como una de las dos partes importantes de su esquema para la estimación de la profundidad. Este método utiliza un Flujo Óptico (OF) para la Estimación del Movimiento (ME). El campo del vector del movimiento generado desde es OF es escaso. Por lo tanto este obtiene un campo denso utilizando una técnica de triangulación Delaunay sofisticada. El campo del vector obtenido es utilizado para estimar un número de parámetros de cámara y la estructura de la escena desde la cual se estima un mapa de profundidad multinivel.

20 Como también se reporta en la técnica anterior, las técnicas SFM tienen problemas en las escenas no estáticas con objetos en movimiento o deformables. Por esta razón para fortalecer el método también se utiliza otra técnica involucrando el uso de señales monoculares. El método Delaunay reporta también que falla en grandes áreas suaves y en los límites alrededor de los objetos. Además, el método SFM es complejo computacionalmente. La razón de su complejidad es que es muy agresivo en la obtención de un mapa de profundidad continuo desde la estimación de toda la estructura en el cuadro. Lo que es necesario, es una técnica para estimar la profundidad la cual es robusta y menos compleja y costosa en términos de esfuerzo computacional.

25 Se conocen otros esquemas para la creación de vídeo 3D desde un vídeo 2D. Por ejemplo, la US-A-2007/024614 enseña la generación de mapas de profundidad desde una fuente de imagen 2D y suavizarlos asimétricamente. Cada mapa de profundidad contiene valores de profundidad enfocados en los bordes de las regiones locales en la fuente de la imagen donde el valor de profundidad asignado a cada píxel en el mapa indica la profundidad del borde correspondiente. Una aproximación al proceso de análisis del borde con base en la estimación de los niveles de desenfoque (opuesto a nitidez) en las regiones locales de la fuente de la imagen 2D, y utilizando el principio de si los bordes y las líneas son gruesas son consideradas desenfocadas, y nítidas si estas son delgadas. Alternativamente, el mapa de profundidad puede ser generado desde la fuente de la imagen estimando la ubicación de los bordes y los límites de los objetos utilizando técnicas de detección de bordes/líneas. En una alternativa adicional, el mapa de profundidad es generado con base en la estimación de la distribución de intensidad de la luminancia en cada región local. El mapa de profundidad es utilizado para procesar la fuente de la imagen por una imagen de profundidad con base en el algoritmo de reproducción para crear al menos una imagen desviada, la cual forma con la fuente de la imagen un grupo de imágenes 2D. El mapa de profundidad es utilizado para formar imágenes 3D. Así, esta técnica utiliza una sola imagen 2D para generar una imagen 3D.

30 La US-B-7161614 divulga un dispositivo y métodos para convertir un vídeo bidimensional en un vídeo tridimensional calculando un mapa de profundidad. Se divulgan diversos métodos para calcular el mapa de profundidad. En un primer

método, una pantalla está dividida en varias áreas. Se obtiene la información relacionada a las frecuencias altas y los contrastes de la luminancia, y el componente de crominancia de cada una de estas áreas. Se encuentra entonces una estimación de profundidad para cada una de las áreas desde esta información. En un segundo método, se genera un vector de movimiento comparando una imagen en el campo actual con una imagen en el campo anterior, para extraer como un vector de movimiento la cantidad de movimiento y la dirección del movimiento de un sujeto en la imagen. La imagen es segmentada en áreas de características y estas áreas están agrupadas con base en la información de crominancia de las áreas. A continuación se obtienen diversos mapas de profundidad considerando para cada área el movimiento detectado, la frecuencia alta y la información de luminancia y contraste, y el componente de crominancia, así como la forma en la cual están agrupadas las áreas. A continuación se obtiene un mapa de profundidad compuesto desde los diferentes mapas. Un tercer método calcula un vector de movimiento entre las dos imágenes en un vídeo y utiliza este, junto con otros factores, en la creación de un mapa de profundidad. Estos métodos generalmente confían en el uso de las "señales" visuales para determinar la profundidad.

La US-B-6996267 enseña un sistema de imagen visual el cual opera en una señal de imagen visual tridimensional para controlar el grado tridimensional de la imagen estereoscópica para intentar reducir la fatiga ocular del usuario.

La US-B-6377625 se relaciona con la generación de imágenes estereoscópicas desde un vídeo MPEG (2D). Un generador de vector de movimiento genera vectores de movimiento individualmente correspondientes a los macrobloques en la imagen. Se analiza el campo del vector de movimiento para determinar el tipo de movimiento de la imagen actual y utiliza esta información en la generación de la imagen estereoscópica. En particular, este incluye la determinación si la actual imagen es una imagen fija, una imagen de movimiento horizontal, una imagen de movimiento no horizontal, o una imagen de movimiento rápido. Esta determinación es utilizada en la generación de una imagen estereoscópica apropiada al tipo de movimiento.

La US-A-2011/007136 se relaciona con el procesamiento de un vídeo 3D para interpolar los cuadros o para mejorar la calidad de la imagen. Se calculan los vectores de movimiento X-Y desde los cuadros consecutivos en el canal izquierdo y en el canal derecho. Estos son procesados para entregar un vector de movimiento en la dirección Z. Este vector de movimiento Z es utilizado en el procesamiento de los cuadros de vídeo 3D para interpolar o mejorar la calidad de las imágenes.

Se hace referencia a THOMAS MEIER ET AL: "Automatic Segmentation of Moving Objects for Vídeo Object Plane Generation", TRANSACCIONES EN CIRCUITOS Y SISTEMAS PARA TECNOLOGIA DE VÍDEO IEEE, CENTRO DE SERVICIO IEEE, PISCATAWAY, NJ, US, vol. 8, no. 5, Septiembre 1 de 1998 (1998-09-01), XP011014498, ISSN: 1051-8215. Esta publicación discute un algoritmo de segmentación de secuencia de vídeo automático que extrae los objetos en movimiento. Un modelo binario 2D del objeto es emparejado contra los cuadros subsecuentes utilizando la distancia Hausdorff.

De acuerdo con un primer aspecto de la presente invención, se proporciona un método de generación de un método de generación de un mapa de profundidad binario desde dos cuadros en una secuencia de vídeo 2D, comprendiendo el método:

estimar un vector de movimiento global y los vectores de movimiento de bloque para el primer cuadro;
 proyectar los vectores de movimiento de bloque en el vector de movimiento global para obtener parámetros de correlación;
 calcular una métrica para el número de bloques estáticos en el primer cuadro;
 determinar si el fondo es estático o dinámico de acuerdo con los vectores de movimiento, el número de bloques estáticos y los parámetros de correlación; y,
 calcular un mapa de profundidad binario clasificando las partes de la imagen como el fondo y el primer plano de acuerdo con los vectores de movimiento, el número de bloques estáticos, los parámetros de correlación y la determinación de si el fondo es estático o dinámico.

Este método realiza un mapa de profundidad binario, por ejemplo clasificando las partes de la imagen como "fondo" o "primer plano". El método genera un vector de movimiento global y vectores de movimiento de bloque entre los cuadros consecutivos en la entrada de vídeo y encuentra la correlación entre estos utilizando una técnica de proyección del vector. Se calcula la diferencia de cuadros entre los cuadros consecutivos también para los bloques para dar una indicación del número de bloques estáticos. Estos parámetros son primeramente utilizados para determinar si el fondo de la imagen está estático o en movimiento. Luego, esta determinación es utilizada en conjunto con la información de

proyección del vector y la información de diferencia de cuadro para calcular un mapa de profundidad binario aplicando varias métricas a la información de cada bloque.

5 El método presente es bastante modesto en el sentido que sacrificando la precisión de robustez y simplicidad se trata de obtener un mapa de profundidad binario, enfocándose solamente en la clasificación de las regiones como regiones de primer plano o de fondo. El método utiliza una operación de proyección de vector simple en un solo vector representando el vector de cuadro de fondo estimado el cual es considerado como un producto secundario de cualquier bloque emparejado con base en el método de estimación de movimiento. Las operaciones pueden ser simples y económicas computacionalmente. Esto es en contraste con las técnicas de la técnica anterior la cual confía en las características de movimiento apropiado de parámetros de cámara múltiple con el fin de encontrar un mapa de profundidad, el cual es costoso computacionalmente y propenso a artefactos.

10 La realización preferida proporciona un sistema sistemático el cual diferencia las escenas estáticas de las escenas no estáticas y explica cómo esta información en conjunto con los vectores de movimiento, métricas de emparejamiento de bloque SAD y sus estadísticas de cuadro, son utilizadas para obtener y aplicar la morfología en los mapas de profundidad y decidir permitir/inhabilitar la operación de conversión 2D-3D deseada.

15 Preferiblemente el método comprende la determinación de si el fondo es estático o dinámico que comprende jerárquicamente la aplicación de reglas sí/no a los vectores de movimiento, el número de bloques estáticos y los parámetros de correlación. Esto proporciona una manera sencilla de utilizar los parámetros calculados del cuadro de entrada para decidir o no si el fondo es estático o dinámico.

20 Preferiblemente el fondo está determinado para ser dinámico si 1) el vector de movimiento global es distinto a cero y 2) hay una gran cantidad de buenos ajustes de los parámetros de correlación aplicados en todos los bloques del cuadro, los buenos ajustes son que al menos el 70% de los bloques tienen correlaciones de al menos el 75% con el vector de movimiento global.

Preferiblemente el fondo está determinado para ser estático si 1) hay un gran número de bloques estáticos o 2) los vectores de movimiento más dominantes en el cuadro no tienen la suficiente dominancia.

25 Preferiblemente el fondo está determinado para ser dinámico si 1) el vector de movimiento global es diferente de cero y 2) hay una gran cantidad de buenos ajustes de los parámetros de correlación aplicados en todos los bloques del cuadro, de lo contrario el fondo está determinado para ser estático si 1) hay un gran número de bloques estáticos o 2) los vectores de movimiento dominantes en el cuadro no tienen la suficiente dominancia.

30 Preferiblemente si el fondo está sin clasificar por cualquier otra regla, el fondo está determinado por defecto para ser dinámico.

Preferiblemente la determinación de un mapa de profundidad binario comprende jerárquicamente la aplicación de reglas sí/no a los vectores de movimiento, el número de bloques estáticos, los parámetros de correlación y la determinación de si el fondo es estático o dinámico en una base por bloque para llegar a una decisión binaria de si el bloque es el fondo o el primer plano.

35 Preferiblemente un bloque está determinado para estar en una región del primer plano si las métricas del error de emparejamiento del bloque en un vecindario de bloques son altas y las variaciones del vector son altas.

Preferiblemente, en el caso de que el fondo esté determinado para ser dinámico, el bloque es determinado a ser el fondo si 1) el parámetro de correlación o 2) la correlación del vector de movimiento de bloques con el vector de movimiento global es alta.

40 Preferiblemente en el caso de que el fondo esté determinado para ser estático, el bloque está determinado para ser el fondo si 1) el número de bloques estáticos da un muy buen ajuste o 2) el vector es el vector cero o 3) el vector es muy pequeño y tiene una pequeña diferencia métrica.

45 Preferiblemente el método comprende un bloque que está determinado para estar en una región del primer plano si las métricas del error de emparejamiento de bloque en un vecindario de bloques son altas y las variaciones del vector son altas; de lo contrario en el caso de que el fondo esté determinado para ser dinámico, el bloque está determinado para ser el fondo si 1) el parámetro de correlación o 2) la correlación del vector de movimiento de los bloques con el vector de movimiento global es alta; en el caso de que el fondo esté determinado para ser estático, el bloque está determinado

para ser el fondo si 1) el número de bloques estáticos da un muy buen ajuste o 2) el vector es un vector cero o 3) el vector es muy pequeño y tiene una pequeña diferencia métrica.

Preferiblemente el método comprende si el bloque está sin clasificarse como el fondo o el primer plano por cualquier regla anterior, el bloque está determinado para ser el primer plano.

5 Preferiblemente el método comprende la aplicación de morfología espacial y/o temporal al mapa de profundidad.

Preferiblemente la aplicación de morfología comprende: la determinación del bloque a ser el primer plano si la suma de las decisiones de profundidad binarias es muy baja, 0 representando el primer plano y 1 representando el fondo, en el bloque y los bloques en su vecindario, en el cuadro presente F_t o en el cuadro anterior F_{t-1} ; y/o determinando el bloque a ser el fondo si la suma de las decisiones de profundidad binarias es muy alta, 0 representando el primer plano y 1 representando el fondo, en el bloque y los bloques en su vecindario, en el cuadro presente F_t o en el cuadro anterior F_{t-1} . El mapa de profundidad binario hace el mapa de profundidad suavizado computacionalmente simple.

10

Preferiblemente el método comprende

la generación de vídeo 3D desde el vídeo 2D, que comprende para al menos un cuadro de una entrada de vídeo 2D, la generación de una imagen 3D de acuerdo con el mapa de profundidad que se proporciona por el método descrito anteriormente.

15

Preferiblemente el método comprende la incapacidad de la generación 3D para un cuadro de entrada si el número de bloques estáticos es alto el cuadro es alto.

Preferiblemente el método comprende la incapacidad de la generación 3D para una entrada de un cuadro si hay una gran cantidad de buenos ajustes para la operación de proyección en un vector de movimiento global diferente de cero, y el emparejamiento del bloque con el vector cero da una gran cantidad de buenos ajustes, siendo los buenos ajustes donde al menos el 70% de los bloques tienen correlaciones de al menos el 75% con el vector de movimiento global.

20

De acuerdo con un segundo aspecto de la presente invención, se proporciona un aparato para la generación de un mapa de profundidad, el aparato está dispuesto para llevar a cabo el método descrito anteriormente.

Preferiblemente el aparato como es descrito anteriormente está dispuesto para llevar a cabo el método descrito anteriormente para la generación de un vídeo 3D desde una entrada de vídeo 2D.

25

Se apreciará que cualquiera de las características expresadas aquí como provistas "en un ejemplo" o como "preferibles" se pueden proporcionar en combinación con cualquiera de una o más de tales características en conjunto con cualquiera de uno o más de los aspectos de la presente invención.

Las realizaciones de la presente invención serán ahora descritas en detalle con referencia a los dibujos acompañantes, en los cuales:

30

La Figura 1 es una representación esquemática de un ejemplo de un método para la generación de un mapa de profundidad de acuerdo con una realización de la presente invención;

La Figura 2 es una representación esquemática de un método de generación de parámetros de estimación de movimiento en el método de la Figura 1.

35

Las Figuras 3A y 3B muestran un ejemplo de cálculos de proyección del vector como parte del método de la Figura 1;

La Figura 4 es una representación esquemática de un método para calcular proyecciones del vector;

La Figura 5 es una representación esquemática de un método para decidir si el fondo es estático o no;

La Figura 6 es una representación esquemática de un método de cálculo de un mapa de profundidad binario;

La Figura 7 es una representación esquemática de un método de morfología espacio-temporal; y,

40

La Figura 8 es una representación esquemática de un método de llegada en una decisión para permitir o inhabilitar la conversión.

- Un ejemplo del presente método será descrito ahora con referencia a la Figura 1 la cual ilustra esquemáticamente un sistema 1 para la generación de un mapa de profundidad para dos cuadros F_t y F_{t-1} de una secuencia de vídeo. Los dos cuadros consecutivos F_t y F_{t-1} de una secuencia de vídeo se ingresan a un bloque 2 de estimación de movimiento. Como se muestra en la Figura 2, el bloque 2 de estimación de movimiento determina los vectores de movimiento entre los dos cuadros F_t y F_{t-1} . Puede ser utilizado cualquier tipo de método de estimación de movimiento de emparejamiento de bloque para crear vectores de movimiento MV de cada bloque y una estimación del vector de movimiento global GMV del cuadro el cual denota el mejor movimiento de fondo representativo. En el presente ejemplo, los cuadros están divididos en bloques de píxeles de 4x4 para el emparejamiento de bloque. Sin embargo, se contemplan otros tamaños de bloque o esquemas de estimación del movimiento.
- Los vectores de movimiento MV y GMV son alimentados a un bloque 3 de proyección del vector el cual es utilizado para generar correlación en los parámetros los cuales indican cómo están correlacionados los vectores de bloque MV y el vector de movimiento global GMV, el cual como se describe posteriormente ayuda a la separación del cuadro en el fondo y el primer plano.
- Los dos cuadros de entrada son también alimentados al bloque 4 de diferencia de cuadro. Este bloque 4 calcula una diferencia métrica entre un bloque en los dos cuadros asumiendo un movimiento cero del bloque entre los cuadros, o puesto de otra manera, con un vector cero entre los bloques. Sin pérdida de generalidad, los bloques pueden ser bloques de píxeles de 16x16. La Suma de las Diferencias Absoluta (SAD) es utilizada en este ejemplo como una métrica de la diferencia entre los bloques correspondientes. Así, para cada bloque 16x16 en el primer cuadro, el SAD se calcula entre los píxeles en el primer cuadro y los píxeles en el segundo cuadro. El SAD es puesto a través de un bloque 5 de comparación de umbrales el cual prueba si el SAD excede algún valor del umbral predeterminado. Si la métrica del SAD es menor que el umbral, esto indica que hay un igualamiento cercano entre el bloque en el primer cuadro F_{t+1} y el segundo cuadro F_t asumiendo un movimiento cero, el cual es una indicación que el bloque puede ser estático. Inversamente, si la métrica del SAD es mayor que el umbral, esto indica que no hay un igualamiento cercano entre los bloques asumiendo cero movimiento, lo cual es una indicación de que el bloque puede estar en movimiento. El umbral puede ser determinado por experimento. Los resultados de las comparaciones son utilizados para generar un recuento del número de bloques estáticos en el cuadro. Se apreciará que pueden ser utilizadas otras métricas de error dentro del alcance de la invención.
- Los vectores de movimiento GMC y MVs, los parámetros de correlación y el número de bloques estáticos son pasados a un bloque 6 de decisión del fondo el cual decide con base en estas entradas si el fondo del cuadro F_t es estático o dinámico. Esta es una parte importante del proceso de estimación del mapa de profundidad y es discutido en más detalle posteriormente.
- Los parámetros de correlación, el número de bloques estáticos y la decisión del fondo son pasados en conjunto con los vectores de movimiento GMV y MVs al bloque 7 de Cálculo de Mapa de Profundidad, el cual calcula un mapa de profundidad binario para el cuadro F_t donde a cada bloque se le atribuye un valor de "1" o "0". Si se decide que el bloque es el fondo, este valor se convierte en "1"; de lo contrario, este valor se convierte en "0".
- En la práctica, es probable que las decisiones de movimiento y por lo tanto las decisiones del mapa de profundidades estén correlacionadas espacial y temporalmente. Por lo tanto la morfología espaciotemporal es aplicada por el bloque 8 en el mapa de profundidad binario inicial para crear un mapa de profundidad binario suave y final.
- El mapa de profundidad final es utilizado entonces por el bloque 9 en la conversión de la entrada de vídeo 2D en un vídeo 3D.
- Los diversos bloques del sistema 1 son ahora descritos con más detalles.
- Bloque de proyección del vector
- El bloque 3 de proyección del vector calcula las proyecciones en el vector de movimiento global y las métricas de emparejamiento del bloque en el vector cero.
- Después del paso de estimación de movimiento que se muestra en la Figura 2, el cuadro es dividido en bloques (teniendo elementos sin pérdida de generalidad de píxeles 4x4 en este ejemplo). Cada bloque constituye un cuadro que tiene su propio bloque de vector de movimiento (MV), y cada cuadro tiene un vector de movimiento global (GMV). Utilizando estos vectores, se calculan las proyecciones del vector en el vector de movimiento global y las métricas de emparejamiento del bloque en el vector cero como sigue.

Primero, como se muestra en la Figura 3, se calculan el producto interior (producto interior_{GMV&MV}) de cada vector de movimiento de bloque MV y el vector de movimiento global GMV:-

$$\text{producto interior}_{\text{GMV\&MV}} = \text{MV}_{\text{vertical}} \times \text{GMV}_{\text{vertical}} + \text{MV}_{\text{horizontal}} \times \text{GMV}_{\text{horizontal}}$$

Además del producto interior, se calcula la potencia del vector de movimiento global (potencia_{GMV}):-

5
$$\text{potencia}_{\text{GMV}} = \text{GMV}_{\text{vertical}} \times \text{GMV}_{\text{vertical}} + \text{GMV}_{\text{horizontal}} \times \text{GMV}_{\text{horizontal}}$$

Luego, como se muestra en la Figura 4, se calculan las proyecciones vertical y horizontal del vector de movimiento global GMV en cada vector de movimiento de bloque:-

$$\text{proyección}_{\text{vertical}} = \text{GMV}_{\text{vertical}} \times (\text{producto interior}_{\text{GMV\&MV}} / \text{potencia}_{\text{GMV}})$$

$$\text{proyección}_{\text{horizontal}} = \text{GMV}_{\text{horizontal}} \times (\text{producto interior}_{\text{GMV\&MV}} / \text{potencia}_{\text{GMV}})$$

10 y se calcula la potencia de la proyección:

$$\text{potencia}_{\text{proyección}} = \text{proyección}_{\text{vertical}} \times \text{proyección}_{\text{vertical}} + \text{proyección}_{\text{horizontal}} \times \text{proyección}_{\text{horizontal}}$$

Aunque la proyección de potencia dada por la anterior ecuación no siempre es negativa, para esta métrica se desea que tenga el mismo signo que el producto interior. De acuerdo con esto, si el producto interior para el bloque es negativo, entonces la potencia_{proyección} para ese bloque es multiplicada por -1 para cumplir con este requerimiento.

15 Por lo tanto, la potencia de proyección (también llamada aquí el parámetro de correlación) da un parámetro indicando la correlación del GMV en los MVs. Un valor altamente positivo da una indicación que el movimiento del bloque es similar al movimiento global del cuadro.

Con el fin de preservar la causalidad, en otras palabras para el bien de una implementación simple, las métricas calculadas desde el cuadro anterior F_{t-1} son utilizadas para calcular el campo de profundidad del cuadro actual F_t .

20 Bloque 6 Decisión del Fondo

La Figura 5 muestra un diagrama de flujo para el proceso de toma de decisión para decidir si cada bloque en el cuadro es el fondo o el primer plano.

Verificación 1 (verificación con la más alta prioridad)

25 Una primera verificación es llevada a cabo en el paso 51. Esta verificación prueba SI (1) el movimiento global es diferente de cero (por ejemplo GMV es diferente de cero) Y (2) hay una gran cantidad de buenos ajustes de la operación de proyección aplicados en todos los bloques del cuadro (por ejemplo una gran proporción de los parámetros de correlación son altos significando que una gran proporción de vectores de movimiento MV igualan el vector de movimiento global GMV) ENTONCES es muy poco probable que el fondo sea estático. Si la verificación 1 es satisfecha entonces se decide que el cuadro tenga un fondo dinámico (por ejemplo un movimiento horizontal o vertical de la cámara).

30

El buen ajuste está relacionado con la resolución del vídeo. Sin embargo, en una realización preferida, un buen ajuste se puede definir cuando al menos que el 70% de los bloques tengan correlaciones de al menos el 75% con el vector de movimiento global.

Si la primera verificación no es satisfecha, se realiza entonces una segunda verificación en el paso 52.

35 Verificación 2 (Verificación de menor prioridad)

Una segunda verificación prueba Si (1) el bloque emparejándose con un vector cero da una gran cantidad de buenos ajustes aplicados en todos los bloques del cuadro (por ejemplo, hay un gran número de bloques estáticos determinados en los bloques 4 y 5) O (2) los pocos más dominantes (por ejemplo dos) vectores de movimiento en el cuadro que no tengan suficiente dominancia entonces este es un indicador que el fondo sea estático, o haya un escenario de movimiento muy caótico dando como resultado que el cuadro tenga un gran número de objetos en movimiento independientes y no haya dominancia del fondo. Si la Verificación 2 es satisfecha entonces se decide que el cuadro tenga un fondo estático (por ejemplo no hay un movimiento horizontal o vertical de la cámara).

40

- En una realización preferida, se determinan una gran cantidad de buenos ajustes en el vector cero a ser al menos el 60% de bloques que se "igualen" en el vector cero. "Dominante" significa que el número del tipo de vector particular es más grande que los otros tipos de vector. (El vector de movimiento dominante en el cuadro es llamado el GMV.) Sin embargo, estos son parámetros. Como se apreciará, estos pueden ser ajustados a través de la experimentación de acuerdo con la aplicación particular, para obtener buenos resultados.
- 5 Por defecto:
- Si ninguna de las dos verificaciones es satisfecha entonces el cuadro es considerado por defecto con un fondo dinámico. En la práctica sin embargo una de las dos verificaciones es satisfecha la mayoría de las veces.
- 10 Por lo tanto, el cuadro está determinado para tener un fondo estático o dinámico por un conjunto sencillo y jerárquico de operaciones de comparación utilizando métricas calculadas previamente.
- Bloque 7 Cálculo del mapa de profundidad binario
- La Figura 6 muestra un diagrama de flujo para el proceso de toma de decisión para decidir si cada bloque en el cuadro es el fondo o el primer plano. La meta principal de este paso es aplicar la segmentación primer plano - fondo (FG-BG) en el cuadro, por ejemplo poniendo cada bloque en una de estas dos categorías.
- 15 Verificación 1 (verificación con la más alta prioridad)
- Una primera verificación es llevada a cabo en el paso 61. Esta verificación prueba SI igualando el bloque en el vector cero en un vecindario de bloques da un buen ajuste Y las variaciones del vector son altas ENTONCES se decide que la región sea una región FG.
- 20 El emparejamiento de bloques en el vecindario de bloques significa que los valores SAD que vienen del bloque 3 de estimación de movimiento deben ser pequeños. En una realización preferida, para cada píxel hay como máximo una diferencia de valor de píxel de 8 a 16 (relacionado al rango de valor de píxel de 0...255) entre dos píxeles en el bloque. "La variación del vector" se refiere al número de vectores de movimiento diferentes. Si el número de vectores diferentes, por ejemplo la diversidad de vectores, es alto, entonces la variación del vector es alta en esa región..
- 25 La segmentación es en sí misma una operación mal condicionada. Esta generalmente nunca es perfecta ni precisa. Cuando hay una ambigüedad en los límites FG-BG y en los límites de objetos diferentes, se prefiere poner al bloque en la categoría FG para tener menos artefactos visuales. Teniendo un buen ajuste entre los bloques emparejados en un vecindario y una gran variación entre los vectores en un vecindario también indica una ambigüedad la cual puede corresponder a la posible presencia de un límite. El tamaño del vecindario puede ser elegido por experimentación para dar los mejores resultados. Por ejemplo, el tamaño de los vecindarios puede ser seleccionado como 3x3 o 3x5 de entre otras posibilidades. Por lo tanto, los bloques ambiguos son puestos en la categoría FG.
- 30 Si la primera verificación no es satisfecha, entonces se lleva a cabo una segunda verificación.
- Verificación 2 (verificación de menor prioridad)
- Esta verificación depende de si el fondo ha sido determinado para ser estático o dinámico por el bloque 6 de decisión de fondo. En la práctica, la decisión de fondo del cuadro anterior es utilizada preferiblemente en esta verificación para determinar el mapa de profundidad de un cuadro actual. Por lo tanto, el diagrama de flujo se ramifica en el paso 62 dependiendo en si el fondo es estático o dinámico.
- 35 Verificación2_d (BG dinámica)
- Si el fondo es dinámico, la segunda verificación es llevada a cabo en el paso 63. Esta verificación prueba (1) si la prueba de proyección del vector de bloque sobre el vector de movimiento es positiva (por ejemplo el parámetro de correlación es positivo y alto) O (2) no importa que resultado de proyección particular sea, si casi todos los resultados de proyección de otros bloques en el cuadro son altos (por ejemplo la correlación de los MVs con el GMV es alta significando casi total uniformidad con el escenario del cuadro en movimiento) ENTONCES el bloque es puesto en la categoría de fondo.
- 40 En una realización preferida, en una primera parte de la prueba, el parámetro de correlación se toma como alta si es mayor que la mitad del valor de gmv-potencia calculado en el bloque 3 de estimación. También, en una segunda parte

de la prueba, la correlación de los MVs con el GMV se toma como alta si al menos el 70% de todos los bloques tienen un "buen igualamiento" con el GMV, por ejemplo tienen correlaciones de al menos el 75% con el GMV.

Verificación 2_s (BG estático)

5 Si el fondo es estático, es llevada a cabo la segunda verificación en el paso 64. Esta verificación prueba SI (1) el emparejamiento del bloque con el vector cero da un muy buen ajuste (por ejemplo el número de bloques estáticos determinados en los pasos 4 y 5 es alto) O (2) el vector MV de bloque es el vector cero O (3) el vector MV del bloque es muy pequeño y también el vector tiene un valor SAD pequeño ENTONCES el bloque es puesto en la categoría de fondo.

10 En una realización preferida, en la primera parte, el número de bloques estáticos puede ser considerado como alto cuando el valor SAD es menor que $1 * \text{los números del píxel en el bloque}$, por ejemplo para cada píxel la diferencia máxima debe ser menor que "1" (relacionada al rango de valor de 0..255). En la segunda parte, el vector MV del bloque puede ser considerado como el vector cero si es cero o sustancialmente cero. En la tercera parte, el vector MV del bloque puede ser considerado como muy pequeño si el tamaño total del vector (la suma de los componentes del vector x y y) es menor que un límite pequeño (por ejemplo menor que 5 a 10). De nuevo, pueden ser sintonizados por experimento diversos límites utilizados en esta prueba para dar mejores resultados.

15 Por defecto:

Si ninguna de las dos verificaciones es satisfecha entonces el bloque es por defecto puesto en la categoría de primer plano de nuevo para tener menos artefactos visuales.

Entonces, se crea un mapa de profundidad binario inicial para el cuadro.

Bloque 8 morfología

20 En este bloque las operaciones de morfología espaciotemporales son aplicadas en el mapa de profundidad inicial calculado por el bloque 7. Las decisiones tomadas en el bloque 7 son independientes. Sin embargo en la práctica las decisiones de movimiento y por lo tanto las decisiones de profundidad están correlacionadas espacial y temporalmente. Por lo tanto se aplica una morfología espaciotemporal en las decisiones.

25 La Figura 7 muestra un diagrama de flujo de un proceso de toma de decisiones para decidir para cada bloque si anular o no la clasificación previa del bloque como BG-FG.

En el paso 71, es llevada a cabo una primera verificación. Esta verificación prueba si la suma de las decisiones de profundidad binarias es baja (asignando digamos 0 representando FG y 1 representando BG) en el bloque y bloques en su vecindario (definidas por un límite bajo) (1) en el cuadro presente F_t O (2) en el cuadro previo F_{t-1} ENTONCES el bloque es puesto en la categoría FG.

30 Por ejemplo, en una realización preferida, la verificación prueba si la suma de las decisiones de profundidad binarias en el cuadro es menor del 80% del número de bloques. Tomando por ejemplo el caso donde el vecindario es elegido para ser una ventana de 5x3 (aunque pueda ser elegida una ventana de cualquier tamaño razonable) hay 15 bloques y por lo tanto 15 decisiones para bloque que puedan ser el fondo o el primer plano. Entonces, el valor máximo de la suma de las decisiones es 15 (por ejemplo todos los bloques son BG con un valor de "1"). La verificación prueba si la suma de las
35 decisiones binarias es menor que 80% de 15, por ejemplo menos que 12, o la decisión binaria anterior del bloque actual es FG, la decisión binaria del bloque actual se convierte en FG.

40 En el paso 72, es llevada a cabo una segunda verificación para el caso opuesto. Esta verificación prueba si la suma de las decisiones de profundidad binarias es alta (asignando digamos 0 representando FG y 1 representando BG) al bloque y bloques en su vecindario (definida por un límite alto) (1) en el cuadro presente F_t O (2) en el cuadro previo F_{t-1} ENTONCES el bloque es puesto en la categoría BG. Por ejemplo, en una realización preferida, la verificación prueba si la suma de las decisiones de profundidad binarias en el cuadro es mayor que el 20% del número de bloques.

Como se apreciará, los límites precisos utilizados para estas pruebas pueden ser seleccionados por experimentación para producir los mejores resultados de acuerdo con la aplicación.

45 Si ninguna condición aplica entonces el bloque preserva su valor de profundidad inicial antes de ingresar a la morfología 8 (por ejemplo determinado por el bloque 7 de mapa de profundidad). El tamaño de los vecindarios puede ser seleccionado como 3x3 o 3x5 de entre otras posibilidades.

Inhabilitación de conversión 2D-3D

Como se describe, el mapa de profundidad final generado por el método descrito anteriormente puede ser utilizado para convertir la secuencia de vídeo 2D a un vídeo 3D por cualquier técnica adecuada aplicada por el bloque 9, de las cuales pueden ser descritas varias en la técnica anterior. Sin embargo, hay situaciones donde la conversión de vídeo 2D en vídeo 3D no es deseable para un cuadro o cuadros en particular. La Figura 8 muestra un esquema para decidir si inhabilitar o no la conversión 2D/3D para un cuadro. El bloque 2 de estimación de movimiento, el bloque 3 de proyección del vector, el bloque 4 de diferencia del cuadro y el bloque 5 de comparación del límite pueden ser los mismos que se muestran en la Figura 1. Un bloque 10 de decisión recibe los vectores de movimiento desde el bloque 2 de estimación de movimiento, los parámetros de correlación del bloque 3 de proyección del vector, y la métrica del número de bloques estáticos del bloque 5 de comparación del límite. El bloque 10 de decisión utiliza dos criterios los cuales son ahora discutidos donde la operación de conversión 2D-3D es inhabilitada para un cuadro de vídeo. El número de criterios pueden ser aumentados o reducidos dependiendo en los métodos de estimación de profundidad aplicados y su eficiencia.

Criterio 1:

Ya que el método de conversión 2D-3D está con base en una estimación del movimiento, si no hay casi ningún movimiento diferente de cero en el cuadro entonces se deshabilita la conversión. Esto se calcula verificando el número de la métrica de bloques estáticos contra el total de número de bloques totales. Este criterio es algo diferente del caso donde hay un fondo estático. En el caso donde hay un fondo estático todavía podría haber objetos en movimiento independientes, mientras que el presente criterio está verificando que no haya movimientos distintos de cero en el cuadro.

Si hay rutinas de estimación de profundidad adicionales las cuales no utilizan la estimación de movimiento, tales como métodos los cuales utilizan señales monoculares, por ejemplo colores, líneas horizontales, variación en las características del desenfoque, etc, entonces puede mantenerse la conversión habilitada en las escenas estáticas.

Criterio 2:

La estimación de movimiento es un proceso mal condicionado. Por lo tanto este fallará en ciertas situaciones. El impacto de la falla será diferente dependiendo de para qué aplicación se utiliza la estimación de movimiento. Una aplicación de conversión 2D-3D es en general muy robusta en contra de los errores de estimación de movimiento. Sin embargo, hay un caso particular para el cual esta es muy sensible. Esto es cuando hay una gran cantidad de buenos ajustes desde la operación de proyección en un vector de movimiento global (fondo) diferente de cero Y el emparejamiento del bloque con el vector cero da una gran cantidad de buenos ajustes, cuando ambos son aplicados en todos los bloques del cuadro. Esta situación es como la situación cuando ambas verificaciones son satisfechas en el paso de proyección del vector, esto es, hay una correlación alta entre el GMV y los MVs y hay un número alto de bloques estáticos (los límites utilizados sin embargo pueden variar). En este caso, dos indicadores que miden métricas opuestas dan ambos valores altos, los cuales son difíciles y conflictivos, por ejemplo, mal condicionados, así como el caso para un esquema de estimación de movimiento.

Como se explicó anteriormente, en una realización preferida una cantidad alta de "buenos ajustes" puede ser considerada que es de al menos el 70% de bloques teniendo un buen ajuste, y siendo un buen ajuste la proyección que tiene un valor de al menos el 75%.

Las realizaciones de la presente invención han sido descritas con referencia particular a los ejemplos ilustrados. Sin embargo, se apreciará que pueden ser hechas variaciones y modificaciones a los ejemplos descritos dentro del alcance la presente invención.

Reivindicaciones

1. Un método para la generación de un mapa de profundidad binario desde dos cuadros en una secuencia de vídeo 2D, comprendiendo el método:
- 5 la estimación (2) de un vector de movimiento global y de los vectores de movimiento de bloque para el primer cuadro; la proyección (3) de los vectores de movimiento de bloque en el vector de movimiento global para obtener los parámetros de correlación;
- 10 el cálculo (4,5) de una métrica para el número de bloques estáticos en el primer cuadro;
- la determinación (6) de si el fondo es estático o dinámico de acuerdo con los vectores de movimiento, el número de bloques estáticos y los parámetros de correlación; y,
- 15 el cálculo de un mapa (7) de profundidad binario clasificando las partes de la imagen como fondo o primer plano de acuerdo con los vectores de movimiento, el número de bloques estáticos, los parámetros de correlación y la determinación de si el fondo es estático o dinámico.
2. Un método de acuerdo con la reivindicación 1, en donde la determinación de si el fondo es estático o dinámico comprende jerárquicamente la aplicación de reglas sí/no a los vectores de movimiento, el número de bloques estáticos y los parámetros de correlación
- 20 y/o
- la determinación de un mapa de profundidad binario que comprende jerárquicamente la aplicación de reglas sí/no a los vectores de movimiento, el número de bloques estáticos, los parámetros de correlación y la determinación de si el fondo es estático o dinámico en una base por bloque para llegar a una decisión binaria de si el bloque es el fondo o el primer plano.
- 25
3. Un método de acuerdo con la reivindicación 1 o la reivindicación 2, en donde el fondo está determinado para ser dinámico si
- 30
- 1) el vector de movimiento global no es cero
y
2) hay una gran cantidad de buenos ajustes de los parámetros de correlación aplicados en todos los bloques del cuadro, siendo los buenos ajustes cuando al menos el 70% de los bloques tienen correlaciones de al menos el 75% con el vector de movimiento global.
- 35
4. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 3, en donde el fondo está determinado para ser estático si
- 40
- 1) hay un gran número de bloques estáticos
o
2) los vectores de movimiento más dominantes en el cuadro no tienen suficiente dominancia.
- 45
5. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde si el fondo no está clasificado por cualquier otra regla, el fondo está determinado para ser dinámico por defecto.
6. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en donde un bloque está determinado para estar en la región del primer plano si las métricas de error en el emparejamiento del bloque en un vecindario de bloques son altas Y las variaciones del vector son altas.
- 50
7. Un método de acuerdo con la reivindicación 6, en donde, en el caso de que el fondo esté determinado para ser dinámico, el bloque está determinado para ser el fondo si

- 1) el parámetro de correlación
- 5 o
- 2) la correlación de los vectores de movimiento de bloque con el vector de movimiento global es alta.
8. Un método de acuerdo con la reivindicación 6 o la reivindicación 7, en donde, en el caso de que el fondo esté determinado para ser estático, el bloque está determinado para ser el fondo si
- 10 1) el número de bloques estáticos da un muy buen ajuste
- o
- 15 2) el vector es el vector cero
- o
- 3) el vector es muy pequeño y tiene una pequeña diferencia métrica.
- 20 9. Un método de acuerdo con la reivindicación 6, en donde se determina un bloque para estar en una región del primer plano si las métricas de error de emparejamiento del bloque en un vecindario de bloques son altas Y las variaciones del vector son altas; además en el caso de que se determina el fondo para ser dinámico, se determina el bloque para ser el fondo si
- 25 1) el parámetro de correlación
- o
- 30 2) la correlación de los vectores de movimiento de los bloques con el vector de movimiento global es alta; y en el caso de que se determine el fondo para ser estático, se determina el bloque para ser el fondo si
- 35 1) el número de bloques estáticos da un muy buen ajuste
- o
- 2) el vector es el vector cero
- 40 o
- 3) el vector es muy pequeño y tiene una pequeña diferencia métrica.
10. Un método de acuerdo con la reivindicación 6 o a la reivindicación 7, en donde si el bloque no está clasificado como el fondo o el primer plano por cualquier regla precedente, se determina el bloque para ser el primer plano.
- 45 11. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 10, que comprende aplicar morfología espacial y/o temporal al mapa de profundidad, lo cual comprende preferiblemente:
- 50 determinar el bloque para ser el primer plano si la suma de las decisiones de profundidad binarias es muy baja, representando 0 el primer plano y representando 1 el fondo, en el bloque y bloques en su vecindario, en el cuadro presente F_t o en el cuadro anterior F_{t-1} ; y/o
- 55 determinar el bloque para ser el fondo si la suma de las decisiones de profundidad binarias es muy alta, representando 0 el primer plano y representando 1 el fondo, en el bloque y bloques en su vecindario, en el cuadro presente F_t o en el cuadro anterior F_{t-1} .

12. Un método para la generación de vídeo 3D desde vídeo 2D, que comprende para al menos un cuadro de una entrada de vídeo 2D, la generación (9) de una imagen 3D de acuerdo con el mapa de profundidad que se proporciona por el método de cualquiera de una o más de las reivindicaciones 1 a 11.
- 5 13. Un método de acuerdo con la reivindicación 12, que comprende impedir la generación de 3D para un cuadro de entrada si
- a) el número de bloques estáticos es alto el cuadro es alto y/o
- 10 b) hay ambos
- una gran cantidad de buenos ajustes de la operación de proyección en un vector de movimiento global diferente de cero
- y
- 15 el emparejamiento del bloque con el vector cero da una gran cantidad de buenos ajustes, un buen ajuste es cuando al menos el 70% de los bloques tiene correlaciones de al menos el 75% con el vector de movimiento global.
- 20 14. Un aparato para la generación de un mapa de profundidad, estando dispuesto el aparato para llevar a cabo el método de cualquiera de las reivindicaciones 1 a 11.
- 15 Un aparato de acuerdo con la reivindicación 14, estando dispuesto el aparato para llevar a cabo el método de cualquiera de las reivindicaciones 12 a 13 para la generación de un vídeo 3D desde una entrada de vídeo 2D.

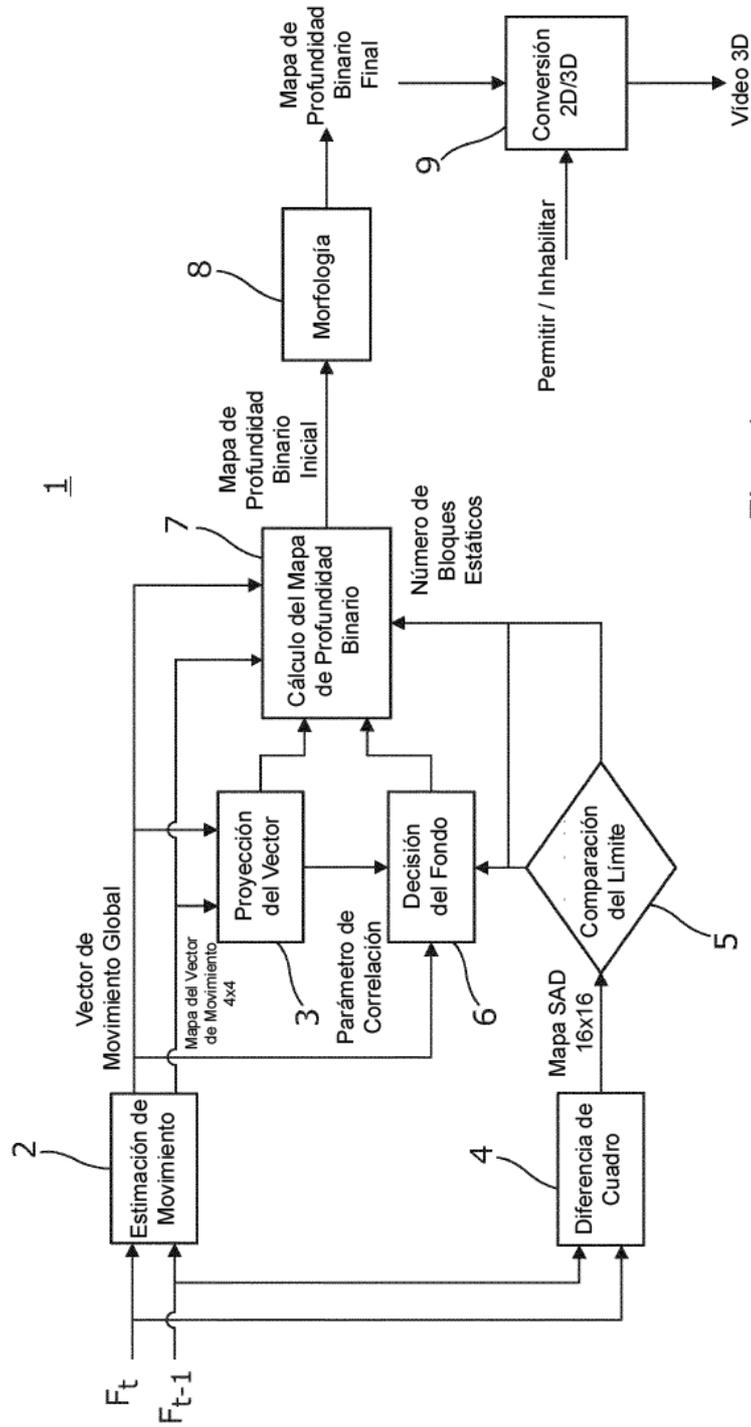


Fig. 1

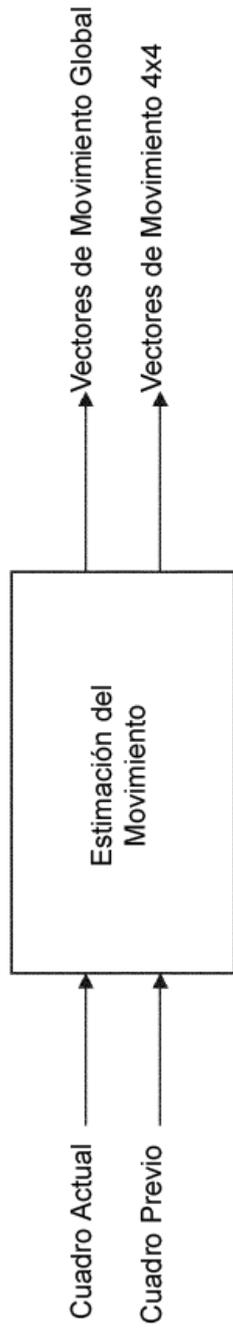


Fig. 2

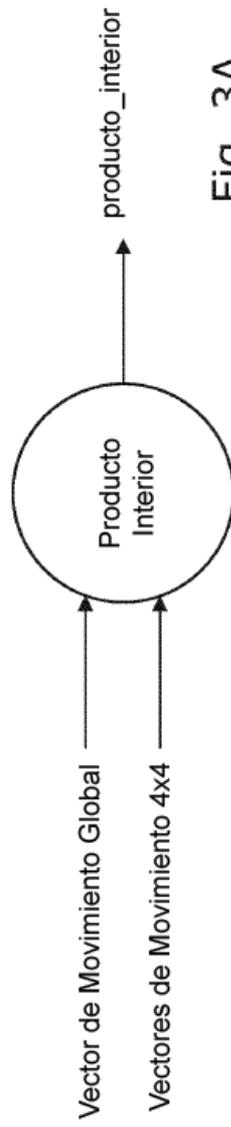


Fig. 3A

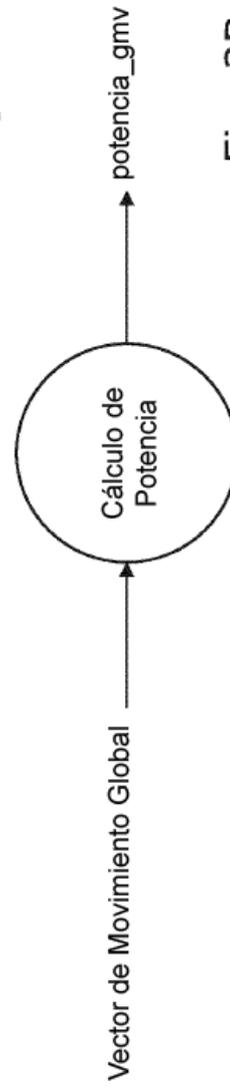


Fig. 3B

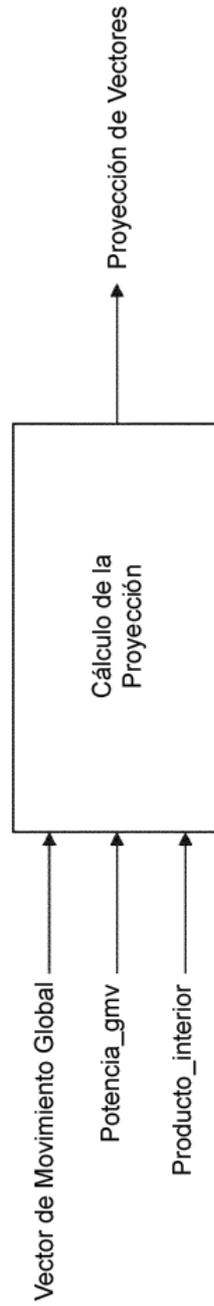


Fig. 4

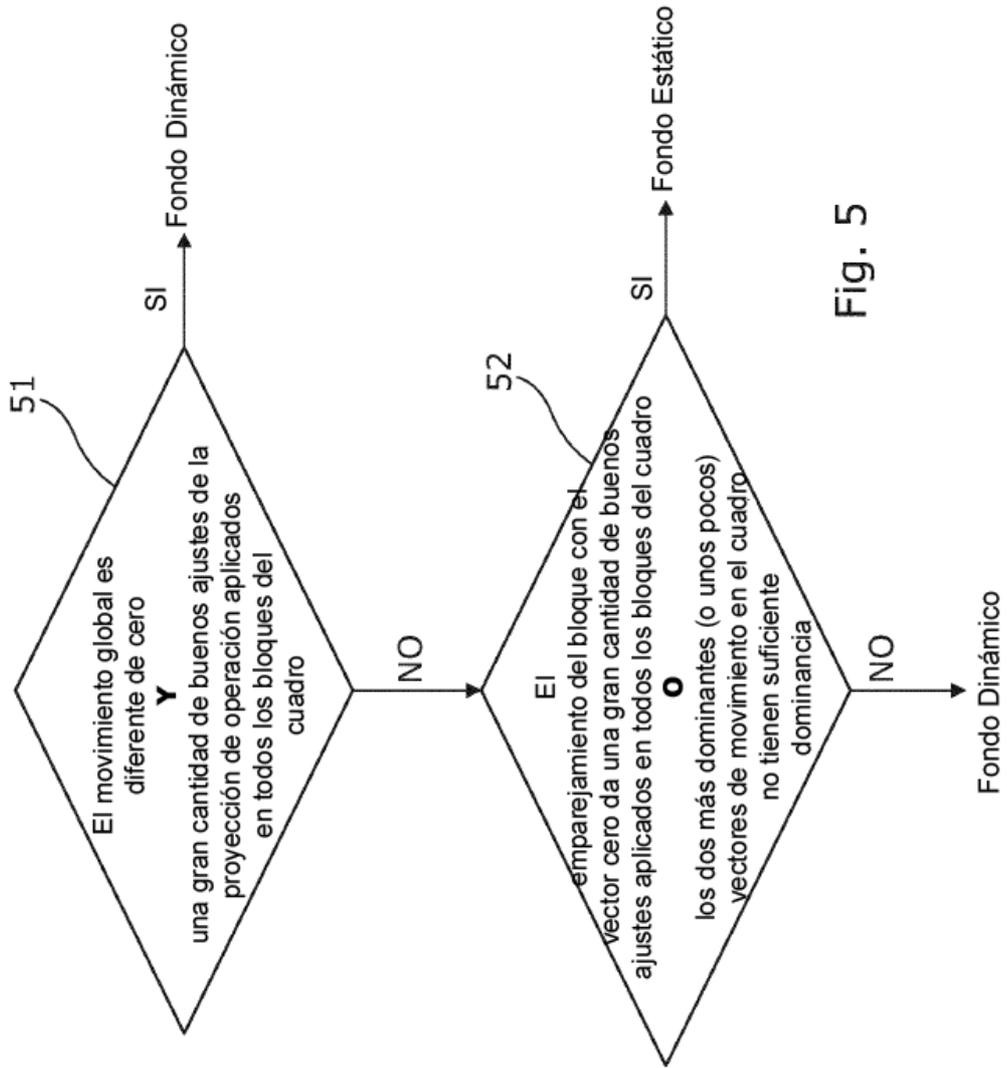


Fig. 5

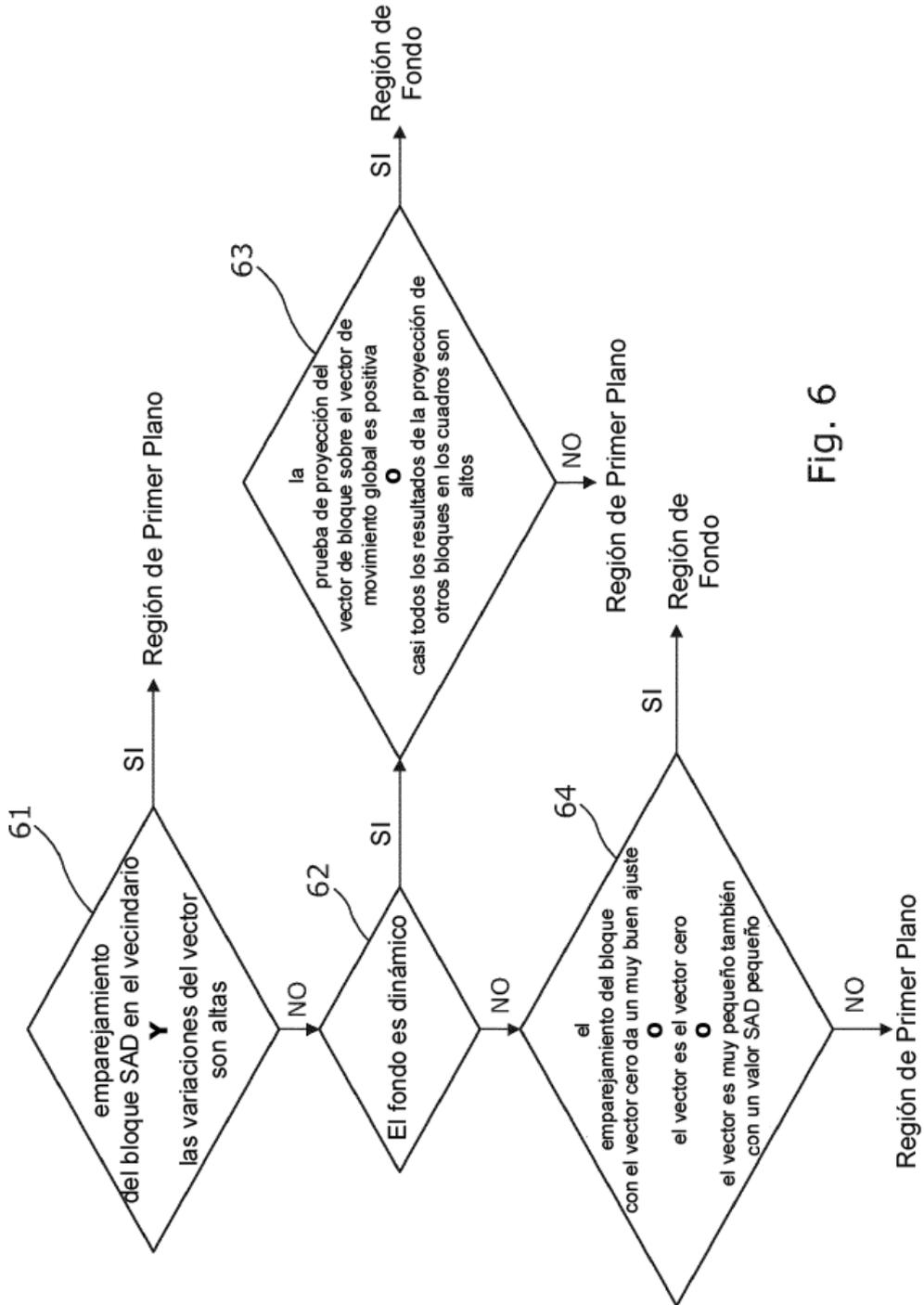
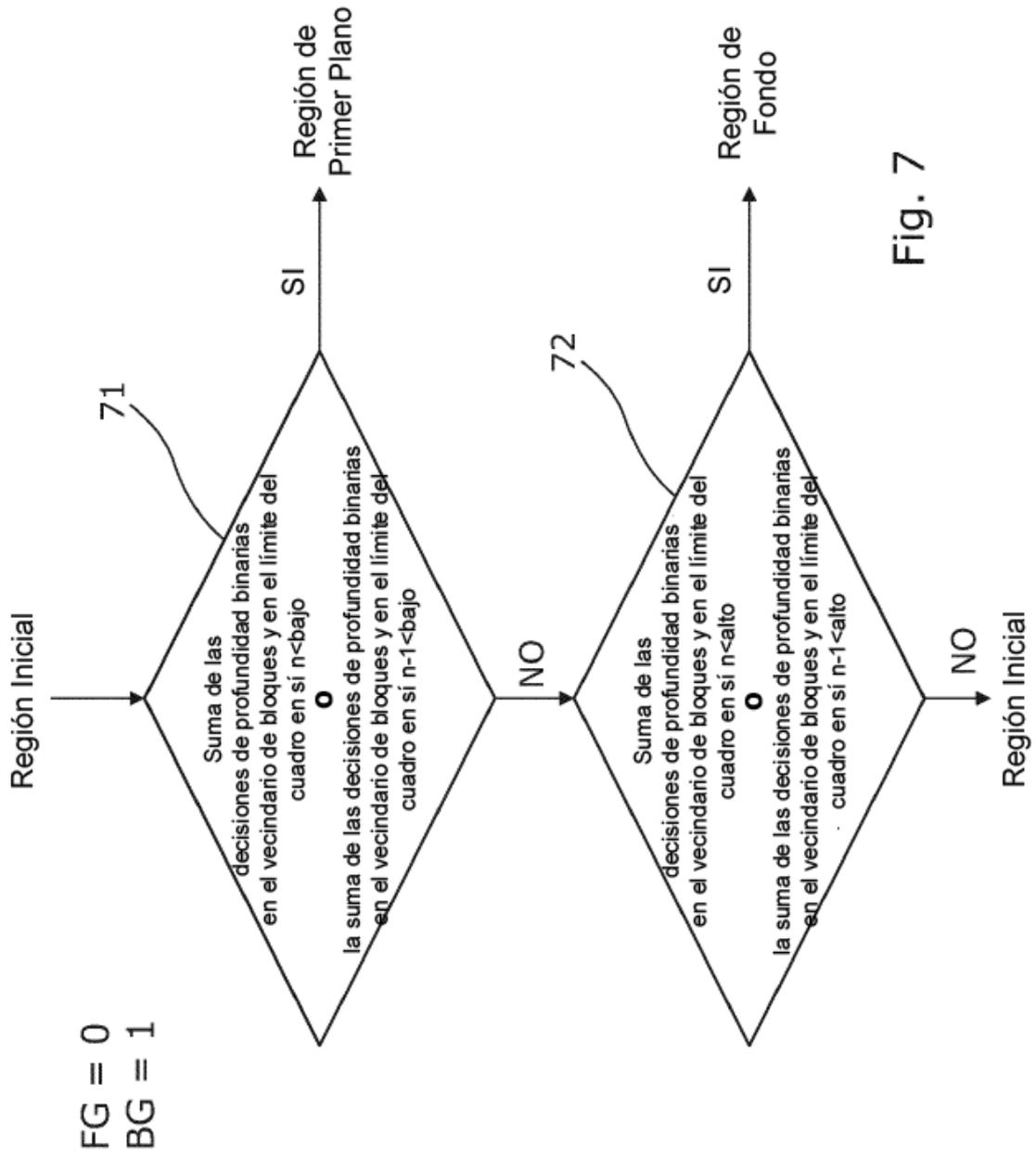
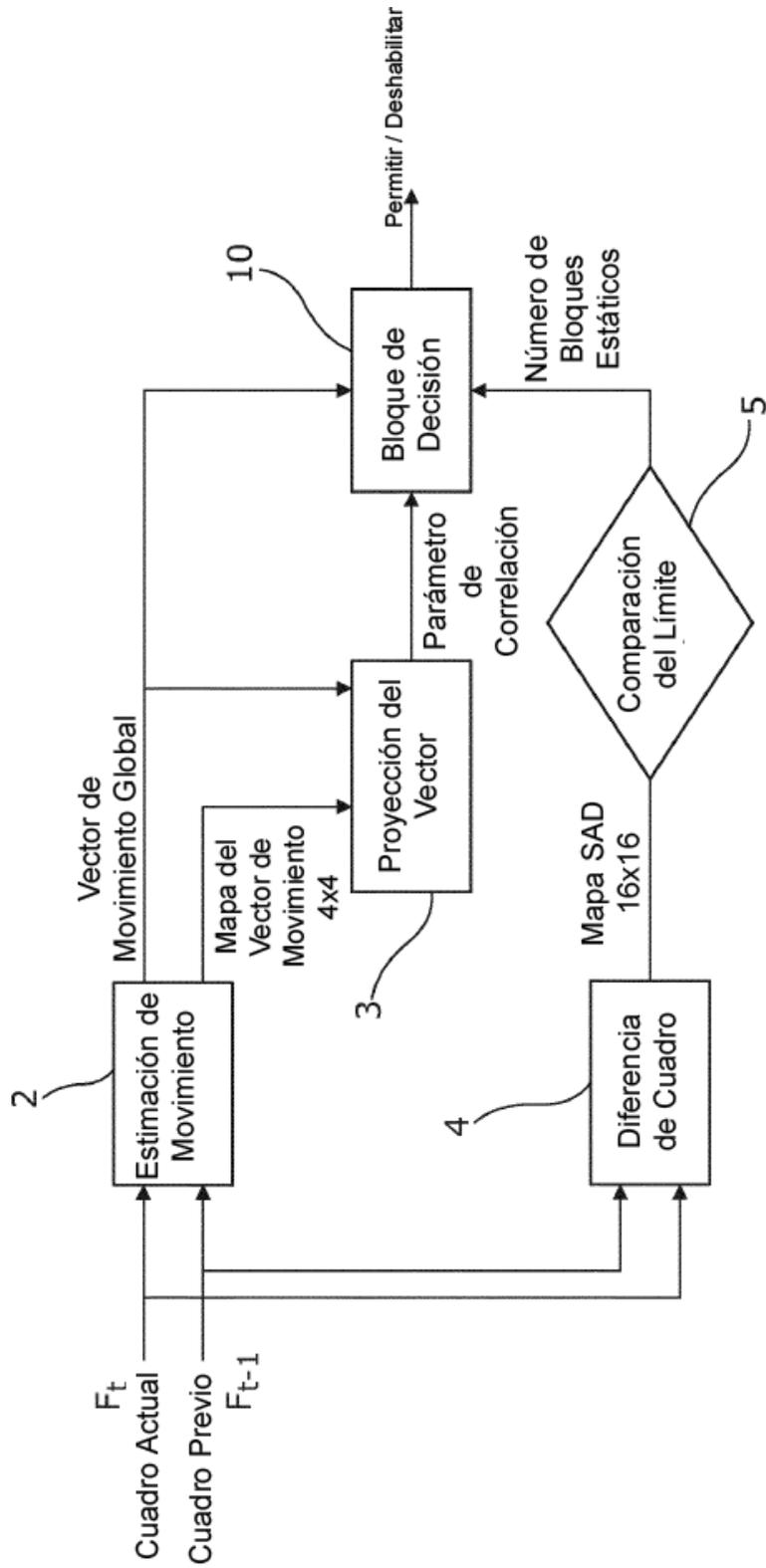


Fig. 6





Decisión de Permitir / Deshabilitar la Conversión

Fig. 8