

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 633 812**

51 Int. Cl.:

**H04N 13/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.12.2011 E 11191981 (7)**

97 Fecha y número de publicación de la concesión europea: **24.05.2017 EP 2466902**

54 Título: **Cálculo de la disparidad basado en la permeabilidad de la información contenida en un vídeo estéreo**

30 Prioridad:

**14.12.2010 TR 201010438**

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

**25.09.2017**

73 Titular/es:

**VESTEL ELEKTRONIK SANAYI VE TICARET A.S.  
(100.0%)  
Organize Sanayi Bölgesi  
45030 Manisa, TR**

72 Inventor/es:

**CIGLA, CEVAHIR y  
ALATAN, ABDULLAH AYDIN**

74 Agente/Representante:

**ARPE FERNÁNDEZ, Manuel**

**ES 2 633 812 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**

Cálculo de la disparidad basado en la permeabilidad de la información contenida en un vídeo estéreo

5 **[0001]** La presente invención aporta un procedimiento para el cálculo de la disparidad de un contenido estéreo. Concretamente, la invención se refiere a la extracción de la información 3D incluida en el contenido estéreo, permitiendo la creación de diversas aplicaciones para TV en 3D, como el ajuste de la profundidad y la mejora basada en la profundidad.

10 Técnica anterior

**[0002]** El cálculo del mapa de disparidad constituye una de las herramientas fundamentales y más importantes para el procesamiento de imágenes en estéreo (multi-view [vistas múltiples]). A medida que ha ido aumentando la disponibilidad de las TVs 3D en el mercado, también ha ido aumentando el interés por el procesamiento estéreo. De este modo, resulta inevitable la estimación del mapa de disparidad, que facilita información acerca de la posición 3D de los objetos o píxeles incluidos en el contenido estéreo (multi-view [vistas múltiples]). En la literatura científica existen diversos procedimientos con los que se obtienen mapas de disparidad a partir de contenidos 3D con presencia de vídeo estéreo o multi-view. No obstante, la complejidad de las operaciones de cálculo constituye una de las limitaciones fundamentales que restringen la utilización de diversas técnicas de optimización que permiten obtener unas estimaciones de gran calidad. De este modo, en los últimos años se han popularizado diversos procedimientos locales que no requieren unos procedimientos globales de optimización que consuman una gran cantidad de recursos. Entre las muchas alternativas existentes, el procedimiento de ponderación de soporte adaptable ofrece el mapa de disparidad más preciso, y se calcula mediante el soporte ponderado de los píxeles adyacentes.

25 **[0003]** No obstante, los procedimientos conocidos para la estimación de una disparidad de contenido estéreo realizan diversas técnicas de optimización que conllevan una gran complejidad de cálculo, y que por lo tanto no son aplicables a los productos de electrónica de consumo. Asimismo, conforme a los procedimientos conocidos, los píxeles similares, atendiendo a sus colores, que se refieren a diferentes objetos pueden aportar unas ponderaciones de soporte muy altas, con lo que se pueden obtener unos mapas de disparidad incorrectos.

30 **[0004]** En la obra "Fast variable window for stereo correspondence using integral images" de Veksler O. se presenta un enfoque de ventanas variables rápido y preciso para la correspondencia estéreo. Las dos principales ideas para conseguir precisión consisten en seleccionar una gama útil de tamaños/formas de ventanas para su evaluación y el desarrollo de un nuevo coste de ventana que sea especialmente apropiado para la comparación de ventanas de diferentes tamaños.

35 **[0005]** En "Cross-Based Local Stereo Matching Using Orthogonal Integral Images" de Ke Zhang, et al. se propone un algoritmo de equiparación estéreo local basado en área para una estimación precisa de la disparidad en todas las regiones de imagen, en el que en primer lugar, para cada píxel de referencia se construye de forma adaptable un esqueleto de apoyo local en forma de cruz vertical, con cuatro brazos de longitud variable decidida en función de la similitud del color y de las limitaciones de conectividad, y en segundo lugar en función de los resultados de las decisiones cruzadas a nivel local, se construye dinámicamente sobre la marcha una región de soporte totalmente adaptable a la forma, fusionando los segmentos horizontales de las cruces en la vecindad vertical.

40 **[0006]** En "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms" de D. Scharstein et al. se presenta una taxonomía de procedimientos estéreo densos de dos tramas.

45 Breve descripción de la invención

**[0007]** La presente invención facilita un procedimiento que presenta las características de la reivindicación 1.

**[0008]** El procedimiento fundamental para la solución de la estimación de la disparidad a partir del contenido estéreo depende de la optimización local, mediante la agregación de coste sobre áreas de soporte específicas para cada píxel. En general, las dimensiones del área de soporte afectan directamente a la complejidad de cálculo del algoritmo. Para eliminar la dependencia de la agregación de costes de la región de soporte, se propone una integración ponderada, en dos etapas, de los valores de coste en las direcciones horizontal y vertical, que también facilita la permeabilidad de la información en regiones carentes de textura y bloques de información situados en los píxeles fronterizos. De ese modo, las constantes de permeabilidad de cada píxel en cuatro direcciones (izquierda-derecha-arriba-abajo) se calculan haciéndolas corresponder con números entre (0-1). Estas constantes indican la proporción de la información que va a transportarse a través del píxel para la correspondiente dirección. Los valores de coste para cada píxel se integran en la dirección horizontal considerando de forma independiente sus ponderaciones de permeabilidad en barrido derecho e izquierdo. Después, los resultados de ambas direcciones de barrido se suman, a lo que sigue una operación similar en la dirección vertical. De este modo se integran también verticalmente las medidas de costes integradas horizontalmente, lo que proporciona una agregación ponderada para cada

60 **[0009]** píxel de toda la imagen completa en tiempo constante, independientemente de las dimensiones del área de soporte. Se efectúa una agregación ponderada en dos pasadas para las vistas izquierda y derecha independientemente, y se asigna a los píxeles la disparidad con el mínimo coste de entre los candidatos alternativos. De este modo, se obtienen dos mapas de disparidad para las vistas izquierda y derecha. Después se lleva a cabo una comprobación cruzada para detectar las estimaciones fiables y las estimaciones no fiables causadas por

oclusiones. Las regiones ocluidas se rellenan promediando sobre asignaciones de disparidad fiables vecinas, con lo cual llega a su fin el procedimiento de estimación de la disparidad.

**[0010]** La estimación de disparidad constituye una importante herramienta para la extracción de la información 3D oculta en el contenido estéreo. Independientemente de la aplicación de la información 3D, los mapas de disparidad estimada deberían conservar las discontinuidades de profundidad en los límites de los objetos, y proporcionar estimaciones fiables en las regiones sin textura. Además de la precisión, la complejidad de cálculo del algoritmo de estimación debería ser baja para que resulte adecuado para aplicaciones de electrónica de consumo. De ese modo, el procedimiento propuesto implica la herramienta de optimización local más sencilla, el procedimiento de “el ganador se lleva todo”, que puede utilizarse en sistemas de aplicación en tiempo real. En los procedimientos de “el ganador se lleva todo”, las asignaciones de disparidad se consiguen minimizando la función de coste agregada en una región de soporte, que pueden ser regiones de dimensiones fijas, arbitrarias o adaptables. En el procedimiento propuesto, el procedimiento ponderado de dos pasadas permite que toda la imagen constituya el soporte de cada píxel, y las regiones conectadas implican una mayor ponderación para píxeles similares en términos de color, con un esfuerzo de cálculo constante.

**[0011]** Conforme a una realización de la invención, las constantes de permeabilidad de cada píxel en las cuatro direcciones se calculan en correspondencia con un valor  $\mu$  que oscila entre 0 y 1 a través de la siguiente ecuación:

$$\mu = \min(\exp(-\Delta R / \sigma), \exp(-\Delta G / \sigma), \exp(-\Delta B / \sigma))$$

Donde  $\Delta R$ ,  $\Delta G$  y  $\Delta B$  indican la distancia absoluta entre los valores Rojo, Verde y azul de los dos píxeles, y donde  $\sigma$  es el factor de alisado.

**[0012]** Opcionalmente, el procedimiento de la invención puede comprender un cálculo de los valores de coste para cada píxel de las vistas izquierda y derecha, lo que se efectúa mediante la siguiente ecuación:

$$Cost(x, y, d) = \sum_{i=1}^3 |I_{left}(x, y, i) - I_{Right}(x + d, y, i)|$$

Donde  $Cost(x, y, d)$  indica los valores de coste correspondientes al píxel  $(x, y)$  en la disparidad  $d$ , y donde  $I_{Left}$  e  $I_{Right}$  son las tres imágenes izquierda y derecha del canal RGB.

**[0013]** De acuerdo con un aspecto adicional de la invención, resulta ventajoso que a una agregación horizontal le siga una agregación vertical antes de minimizar los costes, donde los costes agregados (AC) se obtienen para cada dirección horizontal con arreglo a las ecuaciones:

$$AC^{Left}(x, y, d) = C(x, y, d) + \mu_{Right}(x - 1, y) * AC^{Left}(x - 1, y, d)$$

$$AC^{Right}(x, y, d) = C(x, y, d) + \mu_{Left}(x + 1, y) * AC^{Right}(x + 1, y, d)$$

$$AC^H(x, y, d) = AC^{Left}(x, y, d) + AC^{Right}(x, y, d)$$

y donde los costes agregados (AC) en la dirección vertical se calculan mediante las siguientes ecuaciones:

$$AC^{Down}(x, y, d) = AC^H(x, y, d) + \mu_{Down}(x, y - 1) * AC^{Down}(x, y - 1, d)$$

$$AC^{Up}(x, y, d) = AC^H(x, y, d) + \mu_{Up}(x, y + 1) * AC^{Up}(x, y + 1, d)$$

$$AC^{Final}(x, y, d) = AC^{Down}(x, y, d) + AC^{Up}(x, y, d)$$

**[0014]** Como opción de la invención, la verificación de la consistencia entre las vistas izquierda y derecha detecta las regiones fiables y ocluidas, y la posterior gestión de la oclusión se lleva a cabo mediante un filtrado ponderado en dos pases sobre unas estimaciones de disparidad fiables para la vista izquierda y derecha.

#### Objeto de la invención

**[0015]** Por lo tanto, el objeto de la presente invención consiste en proporcionar un procedimiento para la estimación de una disparidad de contenido estéreo que reduzca la complejidad del cálculo, logrando al mismo tiempo una estimación estéreo de gran calidad. Además, una estimación del mapa de disparidad de alta precisión a partir de vídeo estéreo con una baja complejidad de cálculo resulta una importante función de ampliación de las prestaciones de los televisores 3D, como el ajuste de la profundidad percibida y la mejora basada en profundidad.

De este modo, la extracción de la información 3D permite la inclusión de características adicionales que incrementen la percepción 3D.

Breve descripción de las figuras

- 5 **[0016]** A continuación se describirán en detalle una serie de ejemplos de la presente invención, haciendo referencia a las figuras adjuntas, en las que se representa lo siguiente:
- La figura 1 muestra un diagrama de flujo para el cálculo de la disparidad de contenido estéreo con arreglo a la invención;
  - 10 - La figura 2 muestra una ilustración de las constantes de permeabilidad en cuatro direcciones, para un píxel conforme a la invención;
  - La figura 3 muestra el barrido izquierdo y la ponderación de forma combinada, para calcular el soporte horizontal según la invención;
  - 15 - La figura 4 es una ilustración del barrido derecho y la ponderación de forma combinada, para calcular el soporte horizontal según la invención;
  - La figura 5 muestra un ejemplo de ponderaciones efectivas horizontales de un píxel situado en el centro de un objeto según la invención;
  - La figura 6 es una ilustración de las ponderaciones de soporte efectivas de un píxel en agregación vertical según la invención;
  - 20 - La figura 7 es una ilustración que indica que se han observado oclusiones en el trasfondo local, en torno a los límites del objeto según la invención; y
  - La figura 8 es un diagrama que muestra la función de fiabilidad basada en profundidad, que asigna unas mayores fiabilidades a unas mayores profundidades durante la gestión de la oclusión según la invención.

**[0017]** Los números de referencia utilizados en las figuras pueden poseer los siguientes significados;

25	Vista izquierda	(1)
	Vista derecha	(2)
	Profundidad máxima	(3)
	Cálculo de la permeabilidad	(10,20)
30	Cálculo del coste en función del píxel	(30)
	Agregación horizontal	(40, 50)
	Agregación vertical	(60, 70)
	Minimización	(80)
	Verificación cruzada	(90)
35	Gestión de la oclusión	(100)

Descripción detallada de la invención

40 **[0018]** A continuación se describirán en detalle diversos ejemplos de la presente invención, haciendo referencia a las figuras adjuntas. La presente invención no se limita a los siguientes ejemplos y es posible introducir muchos cambios dentro del alcance de la presente invención. Los ejemplos de la presente invención se facilitan para explicar de forma más completa los ejemplos de presente invención a un experto en la materia.

45 **[0019]** Ha de mencionarse que el contenido estéreo puede adoptar diversos formatos y que puede suministrarse a través de diversas fuentes, que pueden incluir vídeo estéreo digital, imágenes estéreo digitales, tramas estéreo de una copia de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento, como un disco duro, escáner, disco óptico o similar.

**[0020]** La figura 1 muestra el procedimiento de estimación de mapas de disparidad con arreglo a la invención, en el que se lleva a cabo un cálculo de la permeabilidad por separado para una vista izquierda y una vista derecha del contenido estéreo, donde en particular la permeabilidad se calcula en la vista izquierda (10) y la permeabilidad de la vista derecha se calcula en una etapa independiente (20). Como indican las dos flechas, se lleva a cabo un cálculo de coste por píxel de los cálculos de permeabilidad de la vista izquierda y de la vista derecha (30). En la siguiente fase, se calcula una agregación horizontal para la vista izquierda (40) y se calcula una agregación horizontal para la vista derecha (50). Posteriormente se lleva a cabo una agregación vertical para la vista izquierda (60) y se efectúa una agregación vertical par la vista derecha (70). Se efectúa una minimización de la vista izquierda y de la vista derecha (80) para recibir los mapas de disparidad iniciales en función de las asignaciones de disparidad correspondientes a la vista izquierda y la vista derecha, en las que se ha realizado una verificación cruzada (90) basada en una comprobación de la consistencia de la vista izquierda y de la vista derecha, que se utiliza para detectar las regiones ocluidas (100) mediante una gestión de la oclusión realizada sobre unas estimaciones fiables de disparidad, donde un algoritmo de estimación de disparidad proporciona dos mapas para que las vistas izquierda y derecha reciban los mapas de disparidad finales. Dichos mapas de disparidad finales para las vistas izquierda y derecha a partir del contenido estéreo pueden utilizarse para extraer la información 3D que se precisa para realizar el ajuste de profundidad y la mejora basada en la profundidad en TV 3D.

60 **[0021]** De acuerdo con las figuras 2 a 8, el procedimiento para la estimación de la disparidad en contenidos estéreo según la invención se describe con mayor detalle utilizando un procedimiento local de agregación de costes rápido y eficiente.

[0022] Como se muestra en la figura 2, en primer lugar se calcula la permeabilidad de la información de cada píxel en cuatro direcciones, a saber, en la izquierda y la derecha, así como en la dirección superior e inferior. Este cálculo de la permeabilidad determina la proporción de la información que atraviesa el píxel en la dirección correspondiente.

[0023] La permeabilidad se mide comparando los valores de intensidad del píxel central con cuatro situados en su vecindad, como se muestra en la figura 2. La similitud del color (RGB) del píxel central con los píxeles situados en su vecindad determina la permeabilidad de la información que pasa a través del píxel en la dirección correspondiente. La permeabilidad es un valor de ponderación entre (0, 1) calculado mediante la siguiente función de correspondencia de la ecuación (A) que toma las distancias de color absolutas entre el píxel central y los situados en su vecindad.

$$\mu = \min(\exp(-\Delta R / \sigma), \exp(-\Delta G / \sigma), \exp(-\Delta B / \sigma)) \quad (A)$$

[0024] En la ecuación (A),  $\Delta R$ ,  $\Delta G$  y  $\Delta B$  indican la distancia absoluta entre los valores Rojo, Verde y azul de los dos píxeles, y  $\sigma$  es el factor de alisado. La permeabilidad ( $\mu$ ) se determina tomando la ponderación mínima entre diferencias de color RGB. Estos valores se calculan para cuatro direcciones izquierda-derecha-arriba-abajo y por lo tanto, para cada píxel se calcula una lista de cuatro ponderaciones de permeabilidad. Por ejemplo, el cálculo de la permeabilidad de un píxel arbitrario ( $x, y$ ) en dirección izquierda viene dado por la siguiente ecuación (B).

$$\begin{aligned} [\Delta R \ \Delta G \ \Delta B] &= |I_{RGB}(x, y) - I_{RGB}(x-1, y)| \\ \mu_{Left}(x, y) &= \min(\exp(-\Delta R / \sigma), \exp(-\Delta G / \sigma), \exp(-\Delta B / \sigma)) \end{aligned} \quad (B)$$

[0025] Una vez que se determinan las ponderaciones de la permeabilidad para cada píxel, la siguiente fase consiste en calcular los valores de coste en función de las posibles asignaciones de disparidad. De esa forma, las sumas de las diferencias absolutas de valores RGB se explotan como se indica en la siguiente ecuación (C):

$$Cost(x, y, d) = \sum_{i=1}^3 |I_{left}(x, y, i) - I_{right}(x+d, y, i)| \quad (C)$$

Donde  $Cost(x, y, d)$  indica los valores de coste correspondientes al píxel ( $x, y$ ) en la disparidad  $d$ ;  $I_{Left}$  e  $I_{Right}$  son las tres imágenes izquierda y derecha del canal RGB. El coste se calcula para cada píxel, lo que no es recomendable en caso de ruido, regiones sin textura e incluso oclusiones. De este modo, la agregación de la función de coste sobre una región de soporte aumenta la potencia del algoritmo. En esta invención, la agregación viene dada por la integración ponderada en dos pases basada en la permeabilidad de la información. La metodología de agregación propuesta implica dos fases principales, una primera integración horizontal seguida de una integración vertical.

[0026] La integración horizontal se consigue mediante una integración ponderada siguiendo el orden de barrido izquierdo y derecho consecutivamente para cada fila de forma independiente. De esa manera, los valores de coste por píxel se suman hacia las direcciones izquierda y derecha por los factores de ponderación de permeabilidad. En este caso, se obtienen dos costes agregados (AC) diferentes para cada píxel desde direcciones opuestas, como resulta de la siguiente ecuación (D). La información procedente de la dirección izquierda se transporta mediante la permeabilidad de los píxeles en dirección derecha, y viceversa para la dirección izquierda.

$$\begin{aligned} AC^{Left}(x, y, d) &= C(x, y, d) + \mu_{Right}(x-1, y) * AC^{Left}(x-1, y, d) \\ AC^{Right}(x, y, d) &= C(x, y, d) + \mu_{Left}(x+1, y) * AC^{Right}(x+1, y, d) \\ AC^H(x, y, d) &= AC^{Left}(x, y, d) + AC^{Right}(x, y, d) \end{aligned} \quad (D)$$

[0027] Los dos valores de coste agregados se suman a fin de obtener el valor de agregación horizontal definitivo. Es importante señalar que durante la integración ponderada, la permeabilidad permite que la información penetre a través de las superficies lisas, al tiempo que impide que pase la información procedente de la región de los bordes. Esta propiedad permite que los límites de los objetos se preserven durante el cálculo del soporte de los valores de coste. Como resultado de la integración horizontal, en función de las características de textura de la fila correspondiente, para cada píxel se obtiene una suma ponderada entre toda la fila (que corresponde a la ventana horizontal de ancho de imagen) con una complejidad de cálculo constante (2 multiplicaciones y 3 sumas para cada píxel). La información procedente de los píxeles adicionales penetra por las multiplicaciones consecutivas de los

valores de permeabilidad; en las figuras 3 y 4 se muestra una distribución típica de las ponderaciones, en la que la ponderación del píxel pasa de "0" a "1" en la dirección de la flecha. Dicho de otro modo, los píxeles oscuros indican unas ponderaciones cercanas a "0", y los píxeles brillantes corresponden a unas ponderaciones cercanas a "1". El procedimiento de la penetración facilita áreas de soporte conectadas para cada píxel cuando las ponderaciones efectivas se reducen al aumentar la distancia al píxel correspondiente.

[0028] La característica de preservación del borde también puede observarse en una imagen de ejemplo que se muestra en la figura 5, en la que la distribución de ponderaciones de los píxeles marcados con círculos viene dada en la dirección horizontal. Es evidente que la ponderación de la permeabilidad impide que pase información a través de los bordes.

[0029] Esta misma idea también se lleva a la práctica en el caso de la integración ponderada a lo largo de la dirección vertical. En este caso, la suma se lleva a cabo con los valores de coste agregados horizontalmente, en las direcciones hacia arriba y hacia abajo. Durante la integración hacia abajo, el valor integrado anterior se pondera mediante la permeabilidad hacia abajo y se suma al valor actual; cuando esta operación sigue a través de la columna, cada píxel transporta información de los píxeles inmediatamente superiores. Para la integración hacia arriba se invierte esta misma operación, facilitando la transferencia de información para cada píxel, procedente de los píxeles inferiores. Una típica distribución efectiva de ponderaciones para un píxel se facilita en la figura 6, en la que la ponderación del píxel pasa de "0" a "1" en la dirección de la flecha, y las ponderaciones disminuyen debido a la sucesiva multiplicación de permeabilidades. Las fórmulas de actualización de los valores de coste hacia abajo y hacia arriba se indican en la siguiente ecuación (E), en cuya etapa final se suman los dos valores de coste sumado ponderado para obtener el coste agregado definitivo.

$$AC^{Down}(x, y, d) = AC^H(x, y, d) + \mu_{Down}(x, y - 1) * AC^{Down}(x, y - 1, d)$$

$$AC^{Up}(x, y, d) = AC^H(x, y, d) + \mu_{Up}(x, y + 1) * AC^{Up}(x, y + 1, d)$$

$$AC^{Final}(x, y, d) = AC^{Down}(x, y, d) + AC^{Up}(x, y, d) \quad (E)$$

[0030] Las sumas verticales de los valores integrados horizontalmente aportan un área de soporte rectangular, cuando las ponderaciones horizontales efectivas se ponderan adicionalmente mediante las ponderaciones verticales efectivas, obteniéndose para cada píxel las características definitivas de la región de soporte. Las mayores ponderaciones se sitúan en torno al centro, obteniéndose una región de soporte conectada. El procedimiento de dos pases puede facilitar unas áreas de soporte más grandes, dependiendo de las características de textura, con una complejidad constante de la operación, de modo que para cada píxel se realizan 6 operaciones de suma y 4 de multiplicación, independientemente del tamaño de la ventana, lo que resulta mucho más eficaz que los procedimientos conocidos. Por lo tanto, se obtienen regiones de ponderación de soporte adaptable sin restricciones en cuanto al área de soporte, y las características de la región vienen determinadas automáticamente por las variaciones de la textura a nivel local.

[0031] El procedimiento de suma en dos pasadas se aplica para cada candidato a disparidad, y se determinan los correspondientes valores de coste agregados. La estimación inicial de la disparidad se efectúa asignando a la disparidad el valor mínimo para cada píxel, independientemente. Esta operación se lleva a cabo también de forma independiente para las imágenes izquierda y derecha. Las regiones ocluidas y sin textura pueden ser problemáticas durante la asignación inicial de la disparidad. Además, las estimaciones de la disparidad izquierda y derecha han de ser coherentes en las regiones visibles. De este modo, se lleva a cabo una verificación de la coherencia en el lado izquierdo y derecho para determinar las asignaciones fiables y no fiables. De este modo, para cada píxel de la vista izquierda (derecha), se visita el píxel correspondiente de la vista derecha (izquierda). Se espera que, para un píxel fiable en la vista izquierda (derecha), el píxel visitado también pueda establecer una correspondencia con el píxel en cuestión. Las asignaciones no fiables no satisfacen esta verificación uno a uno, por lo que pueden detectarse con precisión. Especialmente, se determinan las regiones ocluidas que implican falta de información en una de las vistas. La fase de verificación cruzada es similar en todos los algoritmos de estimación de la disparidad, por lo que esta etapa es la única etapa que ya se encuentra presente en los procedimientos conocidos.

[0032] La siguiente y última etapa de esta invención consiste en la gestión de las regiones ocluidas y de las asignaciones no fiables de disparidad detectadas. En general, los píxeles ocluidos se encuentran situados entre las discontinuidades de profundidad cuando existen límites de objetos. Los objetos situados en primer término presentan grandes disparidades; por ello, bloquean el fondo, donde la profundidad es mayor. Como resultado de esta característica, puede extraerse la conclusión de que las regiones ocluidas pertenecen al trasfondo local en torno a las discontinuidades de profundidad.

[0033] Puede verse un ejemplo de esto en la figura 7, en la que la silueta de una persona indica que su profundidad está más cerca de la cámara. En la figura 7, las "sombras" de la silueta de la persona indican las regiones ocluidas de la imagen izquierda (1) y de la imagen derecha (2). Es evidente que se observan oclusiones en el segundo plano, producidas por discontinuidades de profundidad. De este modo, las asignaciones de disparidades ausentes de las regiones ocluidas pueden copiarse a través del fondo. El procedimiento de gestión de la oclusión debería igualmente tener en cuenta las características de las regiones ocluidas. De esa forma, el procedimiento propuesto da preferencia a las asignaciones de gran profundidad y realiza una estrategia de soporte similar a la

etapa de agregación de costes. Las asignaciones de disparidad de las regiones ocluidas se proporcionan mediante promediado ponderado sobre píxeles fiables, basado en la ordenación de la profundidad y las similitudes de intensidad.

**[0034]** Para cada píxel, en función de la asignación de disparidad fiable se asigna un valor de fiabilidad. El valor de fiabilidad correspondiente al píxel ocluido se fija en "0", que corresponde a la falta de fiabilidad. Se aplica una función lineal con la gama (0, 1) a los píxeles no ocluidos, que asigna "1" a los píxeles con asignaciones de profundidad más elevadas, "0" a los de asignaciones de profundidad más bajas y los valores situados entre ambos a los restantes píxeles. Las características de la función de correspondencia de la fiabilidad se muestran en la Figura 8. Esta función de correspondencia favorece a aquellos valores con mayores profundidades, ya que en todos los casos, las regiones ocluidas pertenecen a los trasfondos locales. Al igual que en la etapa de agregación de costes, se lleva a cabo un filtrado en dos pases sobre las asignaciones de disparidad escaladas, en dirección vertical y horizontal, mediante las constantes de permeabilidad. Inicialmente, las asignaciones de disparidad se escalan por los valores de fiabilidad, a fin de impedir que la información de las regiones ocluidas (no fiables) disminuya el efecto de las regiones del primer plano para la gestión de la oclusión y de más importancia a las regiones fiables del fondo, que son las que más soportan las oclusiones. La agregación en dos pases se lleva a cabo en mapas de disparidad estimada escalada, con lo que se obtiene la suma ponderada de los valores de disparidad. A fin de determinar la asignación de disparidad más adecuada para el píxel ocluido, la suma ponderada debe normalizarse mediante la suma de las ponderaciones, que puede obtenerse mediante la agregación en dos pases de datos de fiabilidad 2D. La suma ponderada de los valores de fiabilidad compuestos por datos 2D tiene como resultado la ponderación total utilizada durante la suma ponderada de los datos de profundidad para cada píxel. En este caso, la asignación final de la disparidad puede conseguirse mediante la etapa de normalización final, en la que se favorecen las asignaciones de disparidad de fondo para llenar las oclusiones.

**[0035]** Como resultado de ello, las regiones ocluidas se rellenan filtrando las asignaciones de disparidad de las regiones fiables en las que se favorecen durante el filtrado los píxeles del fondo (cuya profundidad es superior). Esta operación se lleva a cabo para las imágenes izquierda y derecha de forma independiente, y se obtienen los mapas de disparidad definitivos.

**[0036]** Concretamente, la invención describe una estimación de la disparidad basada en la permeabilidad de la información en vídeo estéreo, cuyos mapas de disparidad ofrecen información acerca de la estructura 3D de la escena y permiten funciones extra, como el ajuste de la profundidad y la mejora de la imagen basada en la profundidad para televisiones 3D. Este procedimiento conforme a la invención para el cálculo de los mapas de disparidad a partir de vídeo estéreo resulta un procedimiento de agregación local de coste novedoso, rápido y eficiente. La estrategia de cálculo de soporte depende de las características de permeabilidad de la información de los píxeles en cuatro direcciones (izquierda-derecha-arriba-abajo), basada en las similitudes de color con los píxeles vecinos. Las constantes de permeabilidad permiten que la información circule a través de las regiones lisas, al tiempo que impiden el paso de información procedente de los píxeles de los bordes. De este modo, para cada píxel, se asignan cuatro constantes de permeabilidad, a lo cual sigue el cálculo del coste por píxeles en función de las diferencias RGB entre los píxeles de las vistas izquierda y derecha. Los valores de coste en función de los píxeles se agregan mediante una estrategia de integración ponderada de dos pases. De esa manera, utilizando las constantes de permeabilidad como ponderaciones para la correspondiente dirección, se consigue una suma ponderada horizontal por orden de barrido derecho e izquierdo. A continuación se aplica el mismo procedimiento entre los ejes verticales en las direcciones abajo y arriba sobre valores de coste integrados horizontalmente. El procedimiento en dos pases proporciona un área de soporte 2D para cada píxel, dependiendo de las características de textura, con una baja complejidad de cálculo (6 sumas y 4 multiplicaciones para cada píxel). Esta operación se lleva a cabo para cada candidato a la disparidad, y la disparidad que ofrece el mínimo coste soportado se asigna al píxel correspondiente. La asignación inicial de la disparidad se efectúa para ambas vistas, a lo que sigue una comprobación de la coherencia izquierda-derecha que detecta las regiones fiables y las ocluidas. La gestión de la oclusión se lleva a cabo mediante un filtrado en dos pases de un tipo similar, sobre unas estimaciones fiables de la disparidad. Durante esta fase, se da preferencia a los píxeles que tienen mayores asignaciones de profundidad, ya que la mayoría de las regiones ocluidas se encuentran en los píxeles del fondo, en torno a las discontinuidades de profundidad. La gestión de la oclusión finaliza con el algoritmo de estimación de la disparidad, que proporciona dos mapas para las vistas izquierda y derecha, modelando la estructura 3D de la escena y permitiendo el ajuste de la profundidad y la mejora de la imagen basada en la profundidad para Televisiones 3D a partir de contenidos en estéreo.

**[0037]** El procedimiento según la invención también puede extenderse a la extracción de modelos 3D a partir de vistas capturadas mediante múltiples cámaras.

**REIVINDICACIONES**

1. Procedimiento de cálculo de la disparidad de un contenido estéreo, concretamente un vídeo estéreo digital, imágenes digitales, tramas estéreo de una copia de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento, como un disco duro, escáner, disco óptico o similar, de la vista izquierda y de la vista derecha de al menos dos cámaras, que comprende las siguientes etapas sucesivas:

- para cada píxel de dichas vista izquierda y vista derecha, realización del cálculo de las constantes de permeabilidad en las direcciones arriba, abajo, izquierda y derecha, comparando los valores de intensidad de un píxel relevante y de un píxel vecino en las direcciones izquierda, derecha, arriba y abajo (10, 20),

- el cálculo de los valores de coste para cada píxel, en función de diferencias de color entre los píxeles de las vistas izquierda y derecha (30),

- para cada píxel de dichas vista izquierda y vista derecha, el cálculo de los valores de coste agregados, utilizando unas constantes de permeabilidad y unos valores de coste calculados (40, 50, 60, 70),

- la minimización de la función de coste para cada píxel, reduciendo al mínimo los valores de coste agregados en el interior de una región de soporte para atribuir un mapa inicial de disparidad a las vistas izquierda y derecha (80), caracterizado porque

se calculan las constantes de permeabilidad para cada píxel en cuatro direcciones, correspondiendo a un valor  $\mu$  que oscila entre 0 y 1, mediante la siguiente ecuación:

$$\mu = \min(\exp(-\Delta R / \sigma), \exp(-\Delta G / \sigma), \exp(-\Delta B / \sigma))$$

donde  $\Delta R$ ,  $\Delta G$  y  $\Delta B$  indican la distancia absoluta entre los valores Rojo, Verde y azul de los dos píxeles, y donde  $\sigma$  es el factor de alisado.

2. Procedimiento según la reivindicación 1, caracterizado porque el cálculo de los valores de coste para cada píxel de las vistas izquierda y derecha se lleva a cabo mediante la ecuación:

$$C(x, y, d) = \sum_{i=1}^3 |I_{left}(x, y, i) - I_{Right}(x + d, y, i)|$$

donde  $C(x, y, d)$  indica los valores de coste para que el píxel  $(x, y)$  tenga una disparidad  $d$ , donde  $I_{left}$  e  $I_{Right}$  representan las imágenes izquierda y derecha de los tres canales RGB.

3. Procedimiento según una de las anteriores reivindicaciones, **caracterizado porque** la agregación horizontal va seguida de una agregación vertical con anterioridad a la minimización de los costes, en la que los costes agregados (AC) se obtienen para cada píxel en una dirección horizontal, conforme a las ecuaciones

$$AC^{Left}(x, y, d) = C(x, y, d) + \mu_{Right}(x - 1, y) * AC^{Left}(x - 1, y, d)$$

$$AC^{Right}(x, y, d) = C(x, y, d) + \mu_{Left}(x + 1, y) * AC^{Right}(x + 1, y, d)$$

$$AC^H(x, y, d) = AC^{Left}(x, y, d) + AC^{Right}(x, y, d)$$

y en la que los costes agregados (AC) en una dirección vertical se calculan para las siguientes ecuaciones

$$AC^{Down}(x, y, d) = AC^H(x, y, d) + \mu_{Down}(x, y - 1) * AC^{Down}(x, y - 1, d)$$

$$AC^{Up}(x, y, d) = AC^H(x, y, d) + \mu_{Up}(x, y + 1) * AC^{Up}(x, y + 1, d)$$

$$AC^{Final}(x, y, d) = AC^{Down}(x, y, d) + AC^{Up}(x, y, d)$$

y donde  $\mu_{Left}$  se calcula para la siguiente ecuación:



$$[\Delta R \ \Delta G \ \Delta B] = |I_{RGB}(x, y) - I_{RGB}(x-1, y)|$$

$$\mu_{Left}(x, y) = \min(\exp(-\Delta R / \sigma), \exp(-\Delta G / \sigma), \exp(-\Delta B / \sigma))$$

5 Y donde  $\mu_{Right}$ ,  $\mu_{Up}$ , y  $\mu_{Down}$  se calculan de la misma forma pero basándose en los valores de intensidad en las otras direcciones.

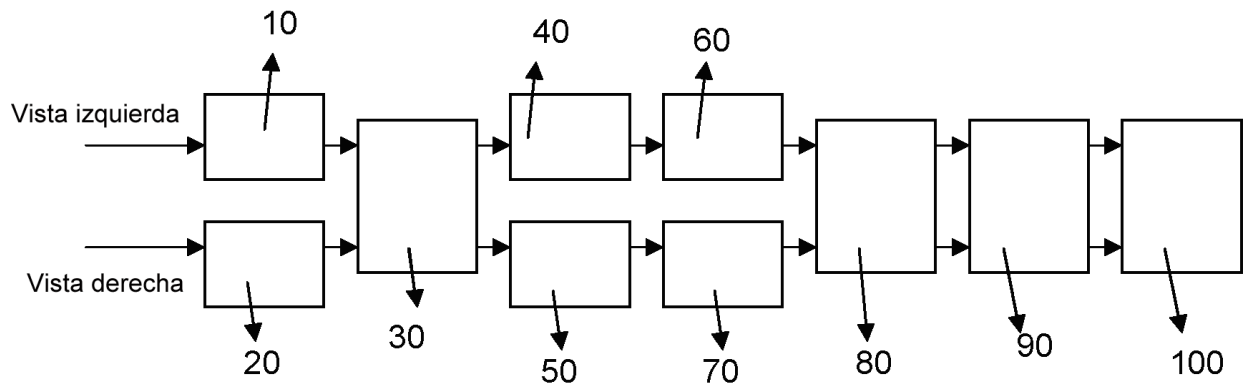


Figura 1

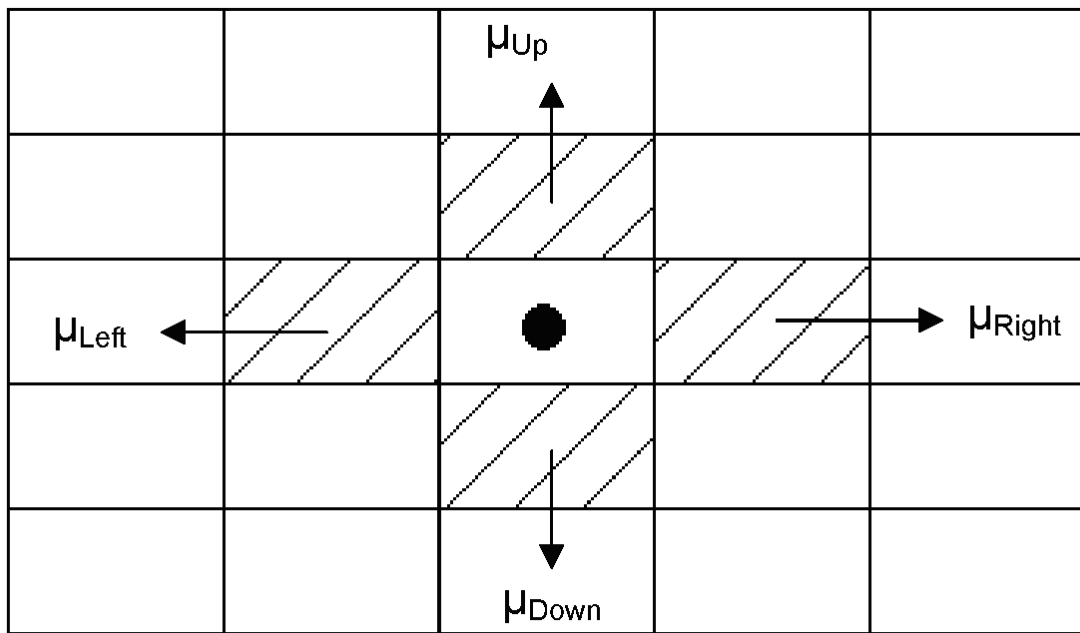


Figura 2

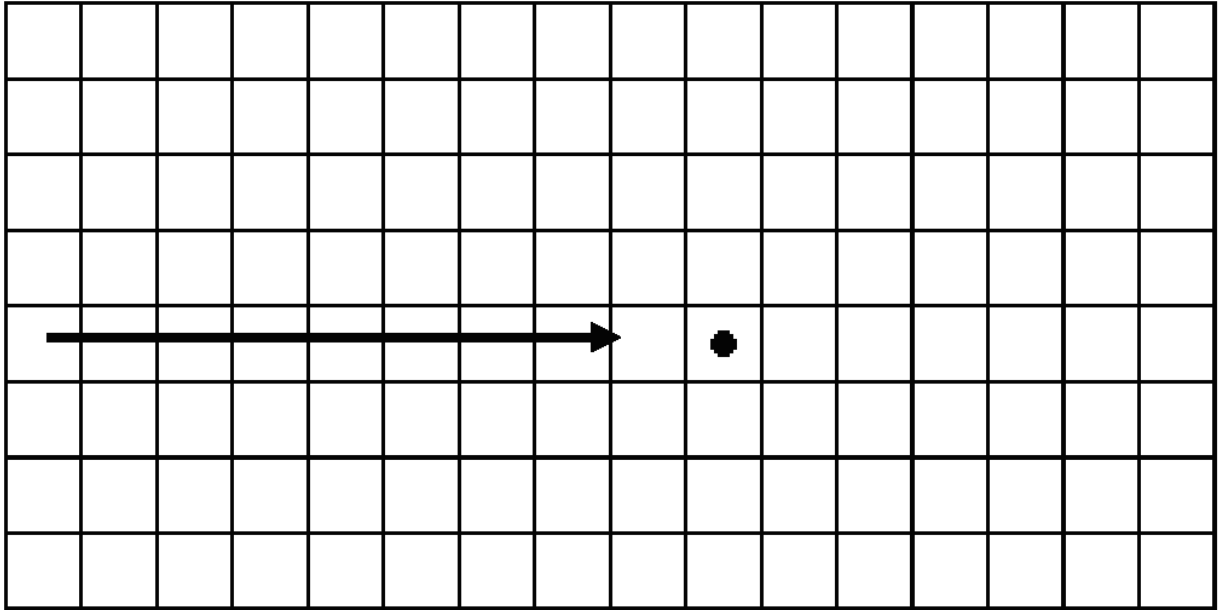


Figura 3

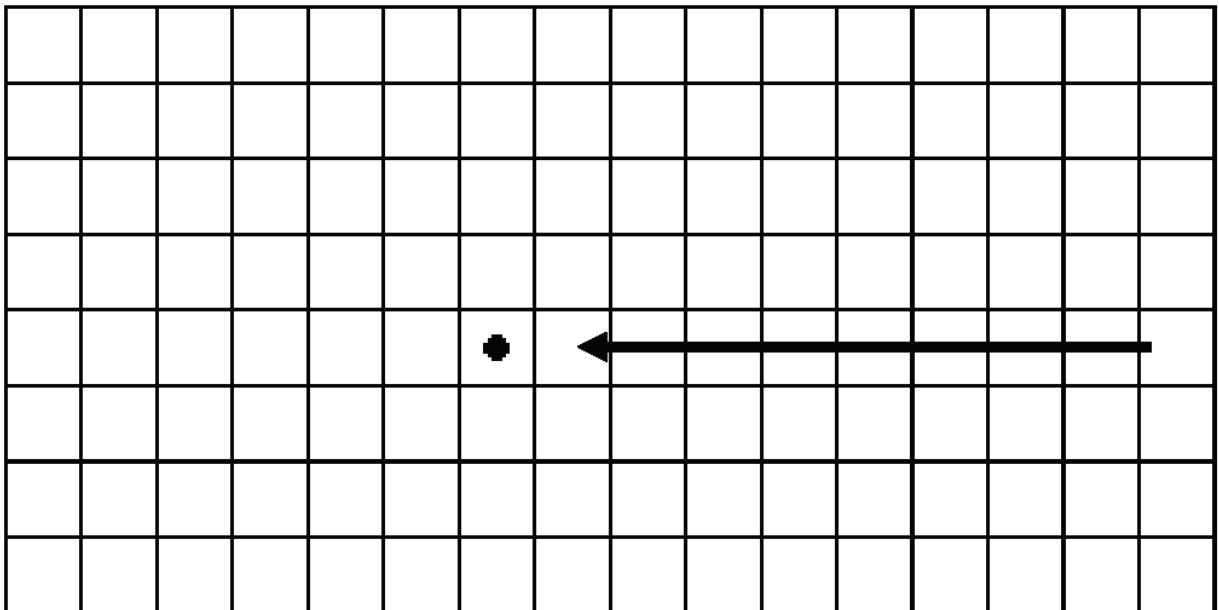


Figura 4

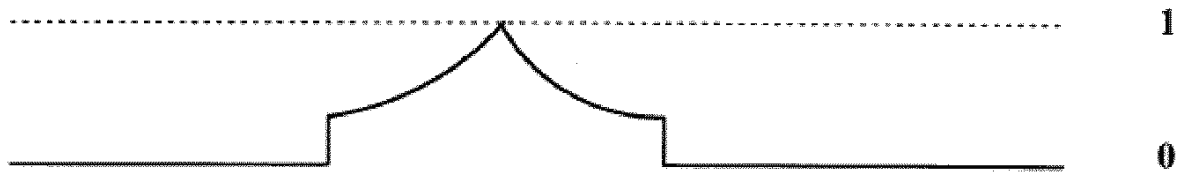
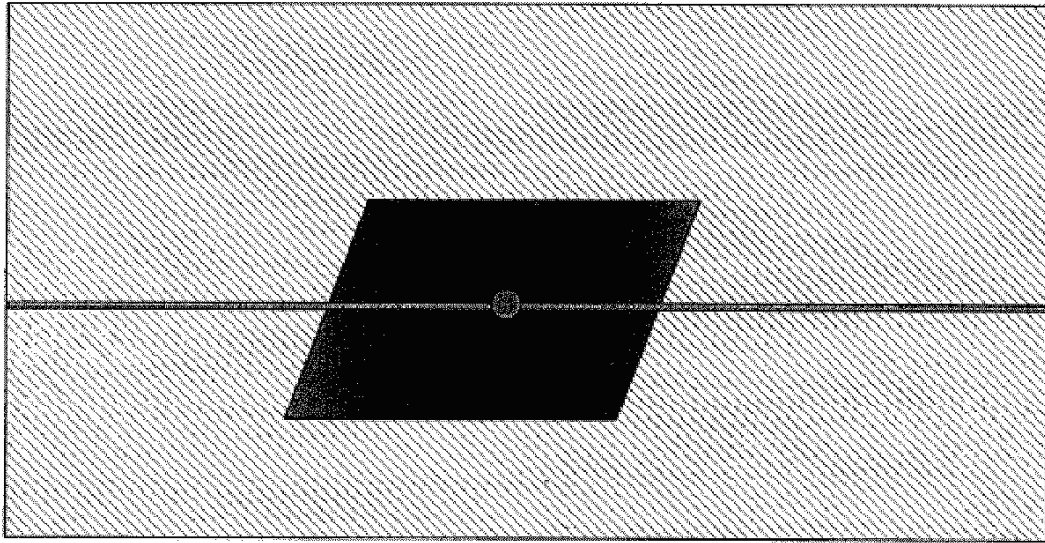


Figura 5

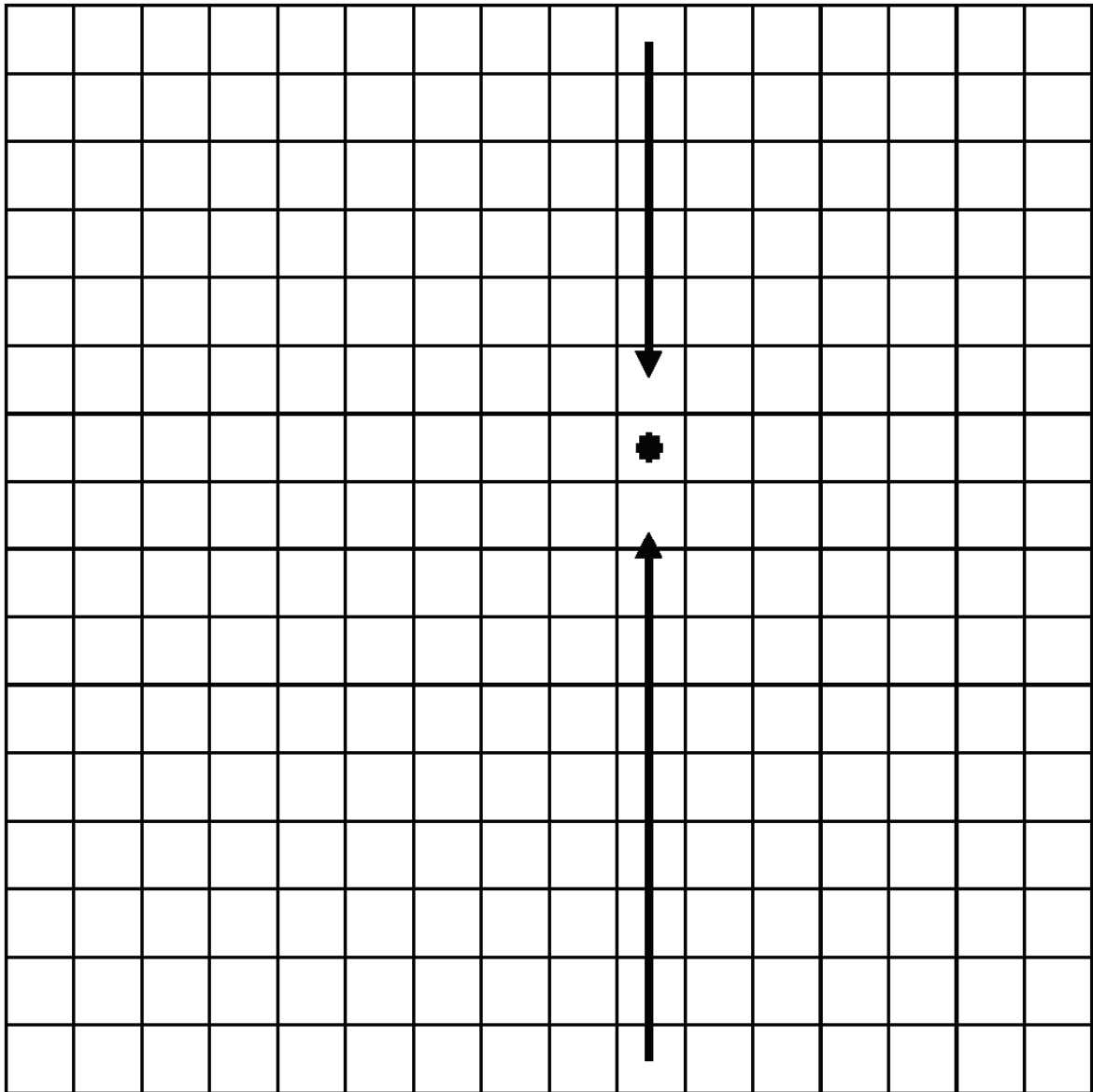


Figura 6

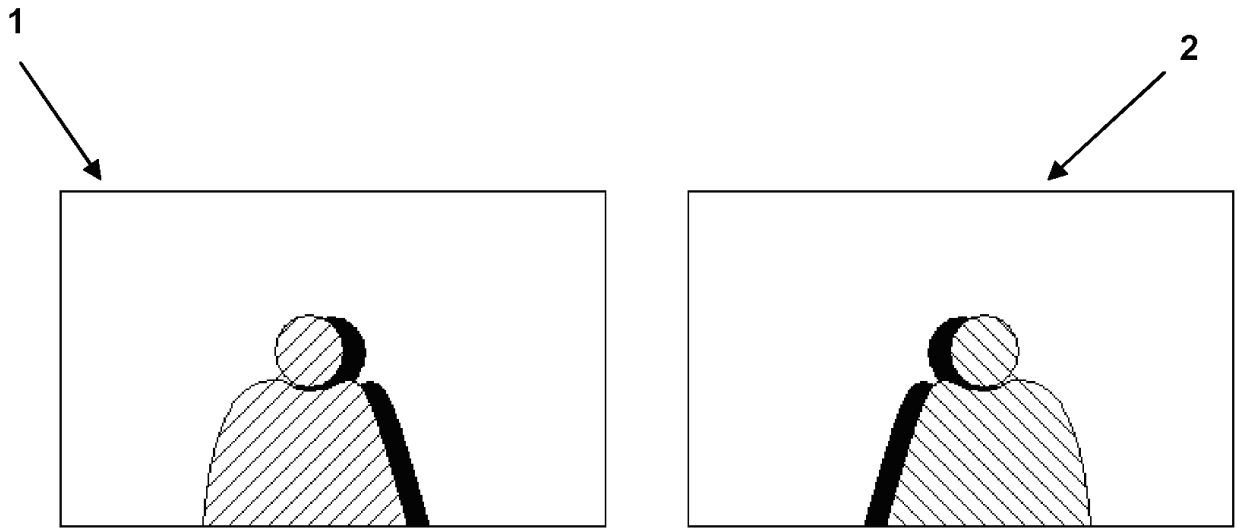


Figura 7

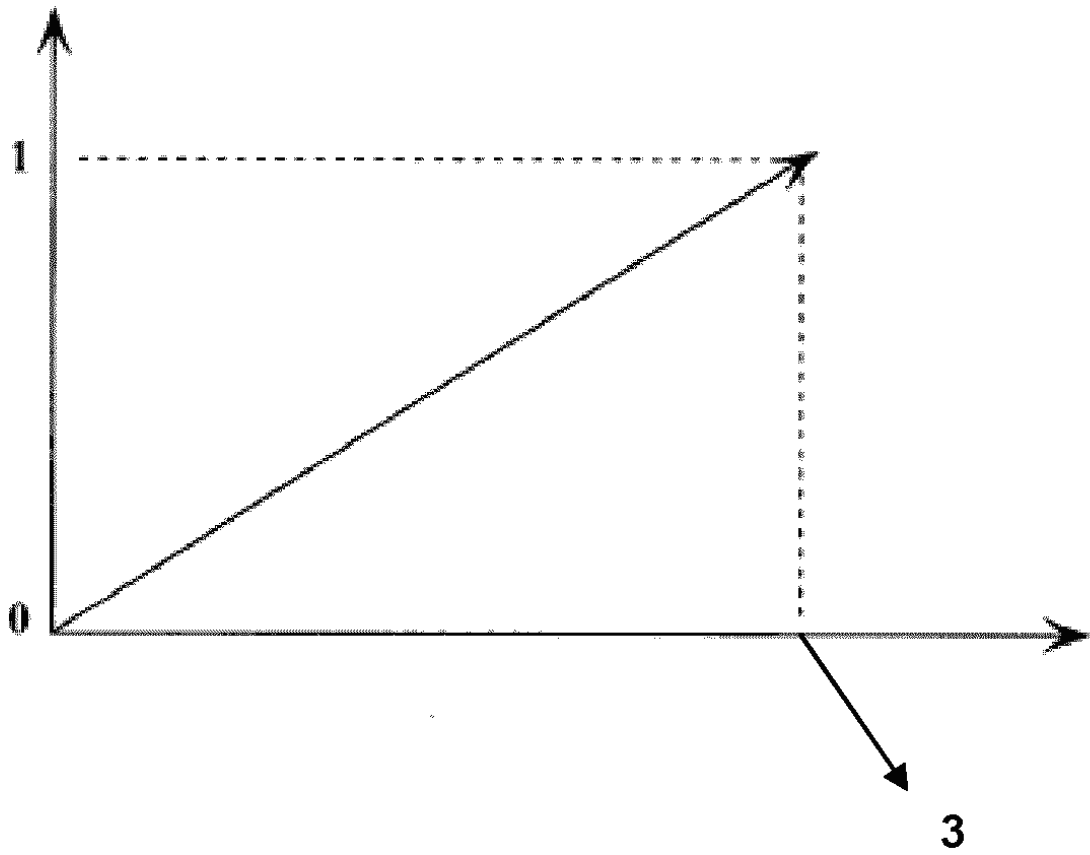


Figura 8

**REFERENCIAS CITADAS EN LA DESCRIPCIÓN**

La lista de referencias citada por el solicitante lo es solamente para utilidad del lector, no formando parte de los documentos de patente europeos. Aún cuando las referencias han sido cuidadosamente recopiladas, no pueden excluirse errores u omisiones y la OEP rechaza toda responsabilidad a este respecto.

**Bibliografía de patentes citada en la descripción**

- **D. SCHARSTEIN.** *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms* [0006]