

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 668 421**

51 Int. Cl.:

H04N 13/00 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.12.2011 E 11191987 (4)**

97 Fecha y número de publicación de la concesión europea: **17.01.2018 EP 2466903**

54 Título: **Un procedimiento y dispositivo para detección de rango de disparidad**

30 Prioridad:

14.12.2010 TR 201010436

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

18.05.2018

73 Titular/es:

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

72 Inventor/es:

CIGLA, CEVAHIR

74 Agente/Representante:

ARPE FERNÁNDEZ, Manuel

ES 2 668 421 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

Un procedimiento y dispositivo para detección de rango de disparidad

5 Campo técnico

[0001] La presente invención proporciona un procedimiento y un dispositivo para detectar un rango de disparidad de contenido estéreo, tal como como, en particular, vídeo estéreo digital, imágenes estéreo digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de emisión en vivo en tiempo real o a través de un medio de almacenamiento como un disco duro, escáner, disco óptico o similar.

Técnica anterior

[0002] Se sabe que el contenido estéreo se proporciona, por ejemplo, por al menos dos cámaras estéreo, que están situadas en el centro de un equipo de cámara y que están definidas como cámaras izquierda y derecha; sin embargo, el contenido estéreo también se puede proporcionar de diferentes maneras.

[0003] En el procesamiento de vídeo estéreo conocido, se requiere una gran cantidad de cálculos para la estimación del mapa de disparidad (aproximadamente 95%) que determina el desplazamiento horizontal de cada píxel entre las vistas izquierda y derecha para lograr una imagen 3D en una pantalla, en particular, por ejemplo, de un televisor o un monitor de ordenador.

[0004] Dependiendo de la geometría de las cámaras estéreo o de visión múltiple, cada desplazamiento horizontal corresponde a una profundidad percibida por el espectador, por ejemplo, en una pantalla de televisión 3D. La información 3D extraída del contenido estéreo o de vistas múltiples es esencial para pantallas 3D en TV 3D particulares para el control de profundidad adaptativo, ya que la mejora basada en la profundidad depende de la geometría 3D de las escenas.

[0005] Se conoce una pluralidad de procedimientos para la estimación de mapas de disparidad que realizan diversas técnicas de optimización que tienen ventajas y desventajas en términos de complejidad de cálculo, precisión y facilidad de implementación. Sin embargo, todos los procedimientos realizan una búsqueda de disparidad dentro de un rango predeterminado independiente del contenido estéreo o de vistas múltiples, de modo que el rango de búsqueda es constante en particular para cada tipo de escena, como una escena interior, exterior, un paisaje o una macro-vista.

[0006] Esto conduce a una variedad de desventajas que provocan que los algoritmos se apilan en mínimos locales, para perder valores reales de disparidad debido a un amplio rango de búsqueda de disparidad y para realizar cálculos adicionales para candidatos de disparidad fuera de rango.

[0007] El documento de patente US6226396B1, describe un procedimiento y sistema para la extracción de objetos. De acuerdo con este documento, para extraer objetos, se utilizan una primera imagen y una pluralidad de segundas imágenes. En dichas imágenes, se calculan las coordenadas tridimensionales de las características y la región característica se extrae como un objeto objetivo de dicha primera imagen.

[0008] Otro documento de patente US2007255480A1, divulga un sistema y procedimiento para detectar y rastrear objetos. De acuerdo con este documento, las imágenes estéreo tomadas por cámaras estéreo se utilizan para detectar la ubicación de los objetos.

Breve descripción de la invención

[0009] Este objetivo se resuelve mediante un procedimiento según la reivindicación 1 y un dispositivo de acuerdo con la reivindicación 6 de esta invención.

[0010] De acuerdo con la invención, se proporciona un dispositivo y un procedimiento para detectar un rango de disparidad de contenido estéreo, en particular tal como vídeo digital estéreo, imágenes digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de transmisión en directo en tiempo real o un medio de almacenamiento tal como un disco duro, escáner, disco óptico o similar, de la vista izquierda y la vista derecha de, al menos, dos cámaras, que comprende las siguientes etapas:

- calibrar el contenido estéreo con respecto a la vista izquierda y la vista derecha,
- detección de píxeles distintivos en el contenido estéreo, en particular a través de la detección de bordes o la extracción del mapa de prominencia,
- calcular valores de costo para los valores de disparidad candidatos de cada píxel distintivo detectado dentro de una ventana del contenido estéreo,
- determinar los valores de disparidad candidatos a la vista de una minimización de los valores de costes calculados para extraer datos fiables píxeles de los valores de disparidad candidatos,
- crear un histograma basado en los píxeles fiables en los que los candidatos de disparidad mínima y máxima son determinados a la vista de un umbral del histograma para detectar el rango de disparidad del contenido estéreo.

[0011] La detección de rango de disparidad según la invención se usa para proporcionar una reducción del cálculo para procedimientos de estimación de disparidad densa en los que se detecta el rango 3D percibido para mejorar la formación de imágenes 3D, por ejemplo en una pantalla de TV, un monitor de ordenador o un dispositivo similar. En particular, la detección del rango de disparidad según la invención se realiza antes de la estimación del mapa de disparidad para reducir la complejidad del cálculo para una estimación del mapa de disparidad y además, la

detección del rango de disparidad de acuerdo con la invención determina el rango 3D percibido del contenido estéreo o multivistas en televisores listos para 3D.

[0012] El contenido estéreo proporcionado se procesa de acuerdo con la invención y puede procesarse adicionalmente con una unidad de emparejamiento estéreo, una unidad de controlador de profundidad, una unidad de visualización en pantalla o una variedad de formas diferentes y, en particular, una analizador de video estereoscópico que realiza cualquier proceso adicional en datos estéreo tal como una estimación de mapa de disparidad.

[0013] Dado que el contenido estéreo puede ser proporcionado por dos cámaras estéreo que definen la vista izquierda y derecha, y por lo tanto fotogramas izquierdo y derecho, la detección del rango de disparidad según la invención depende de la geometría de las cámaras estéreo pueden corresponder al rango 3D existente del contenido estéreo.

[0014] Según una realización preferida, los valores de coste de cada píxel distintivo detectado dentro de una ventana de al menos, se calculan al menos un fotograma estéreo, que tiene un ancho de w y una altura de h mediante las siguientes ecuaciones:

$$\begin{aligned} \text{Costo Izquierda } (x, y, d) &= \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IL(x+i, y+j) - IR(x+i+d, y+j)| \\ \text{Costo Derecha } (x, y, d) &= \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IR(x+i, y+j) - IL(x+i-d, y+j)| \end{aligned}$$

donde d se refiere a la disparidad de cada píxel (x, y) que está rodeado por la ventana (w, h) .

[0015] Según una opción adicional, el procedimiento de la invención se caracteriza porque después de la calibración del contenido estéreo, se realiza un sub-muestreo de los fotogramas izquierdo y derecho antes de que se detecten los píxeles distintivos.

[0016] En caso de que el contenido estéreo haya sido sub-muestreado después de la calibración estéreo, se realiza un aumento de escala opcional realizado de acuerdo con la invención después de determinar los valores de disparidad candidatos.

[0017] Según un ejemplo, es ventajoso que después de crear un histograma basándose en los píxeles fiables, en el que los candidatos de disparidad mínima y máxima se determinan con vistas a un umbral del histograma para detectar el rango de disparidad del contenido estéreo, aplicándose un filtrado temporal al rango de disparidad con respecto al fotograma anterior basado en las siguientes ecuaciones:

$$\begin{aligned} \text{disp_min} &= 0.05 * \text{disp_min} + 0.95 * \text{previous_disp_min} \\ \text{disp_max} &= 0.05 * \text{disp_max} + 0.95 * \text{previous_disp_max} \end{aligned}$$

[0018] De acuerdo con un aspecto adicional de la invención, se proporciona un dispositivo para detectar un rango de disparidad del contenido estéreo, en particular tal como video estéreo digital, imágenes digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento tal como un disco duro, escáner, disco opcional o similar de la vista izquierda y la vista correcta de, al menos, dos cámaras que comprende:

- un calibrador estéreo con respecto a la vista izquierda y la vista derecha del contenido estéreo,
- un detector de píxel distintivo para detectar píxeles distintivos en el contenido estéreo, en particular a través de la detección de bordes o la extracción del mapa de prominencia,
- un calculador de costes para calcular los valores de costes para los valores de disparidad candidatos de cada píxel distintivo detectado dentro de una ventana del contenido estéreo,
- medios para determinar los valores de disparidad candidatos con vistas a minimizar los valores de coste calculados para extraer píxeles fiables de los valores de disparidad candidatos,
- medios para crear un histograma basándose en los píxeles fiables, en el que se determinan los candidatos de disparidad mínima y máxima a la vista de un umbral de un histograma para detectar el rango de disparidad del contenido estéreo.

[0019] Estos y otros aspectos, características y ventajas de la presente invención se describirán o resultarán evidentes a partir de la siguiente descripción detallada de las realizaciones preferidas, que debe leerse en relación con la combinación de dibujos.

Objeto de la invención

[0020] El objeto de la presente invención es proporcionar un procedimiento y un dispositivo para detectar el rango de disparidad del contenido estéreo para reducir la complejidad del cálculo para una estimación de mapa de disparidad.

Descripción de las figuras

5 **[0021]** Las realizaciones de la presente invención se describirán ahora en detalle con referencia a la dibujos, en los que:

- La figura 1 es una realización de un dispositivo para detectar un intervalo de disparidad de acuerdo con la invención;

10 - La figura 2 muestra un diagrama de flujo de la detección del rango de disparidad de acuerdo con la invención;

- La figura 3 muestra la distinción entre los bordes verticales y horizontales durante la búsqueda de disparidad según la invención;

- La figura 4 es una ilustración ejemplar del cálculo dentro de una ventana arbitraria alrededor de un píxel en color e imágenes integrales según la invención;

15 - La figura 5 es una ilustración ejemplar de un control de fiabilidad entre los mejores tres candidatos detectados de disparidad que minimizan la función de costo de acuerdo con la invención;

- La figura 6 es una ilustración ejemplar de la verificación de fiabilidad basada en las características de la función de coste según la invención; y

20 - La figura 7 muestra un umbral de histograma de disparidad para la detección de valores mínimos y máximos de la disparidad rango según la invención.

[0022] Los números de referencia como se usan en las figuras pueden poseer los siguientes significados;

Calibrador de estéreo (1)

Sub-muestreador (2)

Detector de píxel distintivo (3)

25 Calculador de coste (4)

Detector de píxel fiable (5)

Extractor de histograma de disparidad (6)

Dispositivo de umbral de histograma (7)

Dispositivo de escalado (8)

30 Filtro temporal IIR (9)

Calibración estéreo (10)

Sub-muestreo (20)

Detección de píxel distintivo (30)

Cálculo de coste (40)

35 Detección de píxel fiable (50)

Extracción histograma de disparidad (60)

Umbral de histograma (70)

Aumento escala (80)

Filtro temporal (90)

40 Proveedor de contenido estéreo (101)

Unidad de procesamiento estéreo (102)

Detector de rango de disparidad (103)

Unidad de emparejamiento estéreo (104)

Unidad de controlador de disparidad (105)

45 Unidad de presentación en pantalla (106)

Vista izquierda (107)

Vista derecha (108)

Imagen integral (109)

50 Descripción detallada de la invención

[0023] De aquí en adelante, las realizaciones de la presente invención se describirán en detalle con referencia a los dibujos adjunto. La presente invención no está restringida a las siguientes realizaciones, y son posibles muchas variaciones dentro del alcance de la presente invención. Las realizaciones de la presente invención se proporcionan con el fin de explicar más completamente ejemplos de la presente invención a un experto en la técnica.

55 **[0024]** Se debe mencionar que el contenido estéreo puede estar en varios formatos y se puede suministrar a través de varias fuentes, que pueden proporcionar video estéreo digital, imágenes estéreo digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento tal como un disco duro, escáner, disco óptico o similar.

60 **[0025]** La figura 1 muestra un detector de rango de disparidad de acuerdo con la invención, que es un dispositivo que determina el rango de la imagen 3D percibida en un contenido estéreo.

[0026] Una unidad de procesamiento estéreo (102) recibe el contenido estéreo suministrado a través del proveedor de contenido estéreo (101) y realiza el procesamiento del contenido estéreo para diversos fines. Para este propósito, la unidad de procesamiento estéreo (102) se puede construir en un ordenador o sistema de TV 3D o similar que explota contenido estéreo para visualización y creación de percepción tridimensional. Como alternativa, la unidad de procesamiento estéreo (102) puede incluso construir en una cámara que habilite ciertas herramientas para el video

estéreo o el procesamiento de imágenes. La unidad de procesamiento estéreo (102) se puede usar para diferentes propósitos tales como extracción de profundidad, control de profundidad percibido, mejora basada en profundidad o incrustación en pantalla. La invención se dirige y comprende en particular una unidad detectora de rango de disparidad (103) que está incorporada en la unidad de procesamiento estéreo (102) y que determina el rango de profundidad o disparidad del contenido estéreo definiendo la profundidad mínima y máxima. La unidad detectora de disparidad (103) según la invención proporciona un rango de trabajo específico de la disparidad para posibles unidades que pueden incorporarse en la unidad de procesamiento estéreo (102) tal como una unidad de ajuste estéreo (104) que estima el mapa de disparidad y la estructura 3D en el video estéreo, una unidad controladora de profundidad (105) que permite el ajuste de 3D percibido, o una unidad de visualización en pantalla (106) que crea parches, subtítulos u opciones de menú dependiendo de la estructura 3D en el contenido estéreo. La invención también puede explotarse para diversas aplicaciones adicionales que requieren una profundidad percibida mínima y máxima en el contenido estéreo.

[0027] El detector de rango de disparidad (103) comprende una pluralidad de subunidades para extraer el rango de profundidad del estéreo contenido tal como se explica a continuación. Como primera etapa, el contenido estéreo se procesa dentro del detector de rango de disparidad (103) mediante el calibrador estéreo (1) que ecualiza las características de color de los fotogramas estéreo (izquierdo y derecho) si hay un desequilibrio entre las características del color. Entonces, una unidad opcional de sub-muestreo (2) puede disminuir la resolución del contenido estéreo para disminuir la complejidad de cálculo y lograr robustez contra el ruido. Como etapa siguiente, un detector de píxeles distintivo (3) determina los píxeles del borde o esquina en las vistas estéreo que son importantes para el emparejamiento estéreo. Luego, un calculador de coste (4) determina la similitud entre los píxeles distintivos en las vistas estéreo según las hipótesis de disparidad tomando las diferencias absolutas de los canales Rojo, Verde y Azul, y preferiblemente comprende medios para calcular los costes de cada píxel distintivo detectado dentro de una ventana de, al menos, un cuadro estéreo. Como etapa siguiente, un detector de píxel fiable (5) determina los píxeles que tienen asignaciones de disparidad consistentes de los píxeles fiables. Una unidad de extracción de histograma (6) crea un histograma de disparidad en función de las asignaciones de disparidad de los píxeles fiables. A continuación, un umbral de histograma (7) determina los valores de disparidad que tienen recuentos grandes sobre píxeles fiables, por lo tanto, un umbral existente. En caso de que los fotogramas estéreo se hayan muestreado de forma opcional después de la calibración estéreo, una unidad de escalado (8) compensa el rango de disparidad multiplicando el rango detectado con la razón de sub-muestreo para asignarlo a la resolución de contenido estéreo real en caso de usarse una unidad opcional de su-muestreo (2). Finalmente, el filtro temporal IIR opcional (Infinite Impulse Response [respuesta de impulso infinito]) (9) realiza una ponderación en el tiempo para eliminar las posibles fluctuaciones del rango de disparidad para caso de video estéreo.

[0028] Como ya se mencionó, la detección del rango de disparidad del contenido estéreo, en particular en forma de fotogramas estéreo para una vista izquierda y derecha es esencial para reducir la complejidad del cálculo para una estimación del mapa de disparidad que puede seguir la detección del rango de disparidad. Para la detección del rango de disparidad, una búsqueda inicial de píxeles distintivos tiene que ser realizada lo que se explicará con más detalle en la que se buscan los píxeles de los bordes o las regiones de mayor prominencia, que es una solución más robusta y reduce la complejidad del procesamiento posterior de los datos estéreo en comparación con la realización de una búsqueda completa de todos los píxeles, que es más compleja y puede implicar soluciones erróneas.

[0029] Sin embargo, antes de la detección de píxeles distintivos, los fotogramas estéreo se calibran con respecto al color e intensidad para garantizar que todas las vistas se refieran a la misma gama de colores de la escena. Por lo tanto, la calibración estéreo es requerida para normalizar inicialmente el color y la intensidad de los fotogramas estéreo como ya se mencionó para utilizar datos consistentes en todo el proceso de detección del rango de disparidad. Existen varios algoritmos de ecualización de histogramas relacionados con el procesamiento estéreo, por lo tanto, cualquiera de estos algoritmos puede utilizarse en esta etapa. La invención no depende estrictamente de la elección del procedimiento de normalización, sin embargo, el rendimiento de la etapa de procesamiento previo afecta al rendimiento de la detección del rango de disparidad. De acuerdo con un procedimiento conocido, la normalización se realiza de acuerdo con una imagen de referencia que se elige para ser el fotograma izquierdo para la presente invención. Por lo tanto, suponiendo que la vista de referencia tiene la mejor alineación de color con la escena observada, la vista derecha se corresponde con la vista de referencia basada en un histograma. La ecualización del histograma se logra para cada canal de datos RGB de forma independiente. Se pueden encontrar más detalles del procedimiento de normalización y ecualización en el documento "Histogram- Based Prefiltering for Luminance and Chrominance Compensation of Multiview Video" por Ulrich Fecker, Marcus Barkowsky y Andre Kaup, transacciones IEEE en circuitos y sistemas para tecnología de video, volumen 18, nº 9, septiembre de 2008, páginas 1258 a 1267.

[0030] La detección del rango de disparidad anterior dentro de las subunidades del detector de rango de disparidad (103) es la base para reducir la complejidad de cálculo para después el contenido estéreo procesado en particular para la estimación del mapa de disparidad. De esa manera, la invención proporciona un algoritmo de detección de rango de disparidad antes de realizar cualquier proceso adicional en datos estéreo.

[0031] De acuerdo con las figuras 2 a 7, el procedimiento para detectar un intervalo de disparidad en el contenido estéreo según la invención se describe con más detalle en donde la idea central del procedimiento de acuerdo con la invención es detectar lo visualmente atractivo que significa regiones prominentes en los fotogramas izquierdo y derecho, tales como píxeles de borde vertical, píxeles localmente distintivos a través de diferencias rodeando el centro y similares, como se explica con respecto a las figuras 2 a 7 de la invención.

[0032] El diagrama de flujo de la figura 2 muestra una descripción general del procedimiento para detectar un intervalo de disparidad según la invención, en el que el contenido estéreo se proporciona a una calibración estéreo (10) como primera etapa.

[0033] La equalización de color proporciona fotogramas estéreo para involucrar una distribución de color similar eliminando la posible diferencia de color resultante de diferentes calibraciones de la cámara. Una vez que se realiza la calibración, los fotogramas estéreo pueden ser sub-muestreados (20) en la siguiente etapa, para ahorrar cálculo. La complejidad de cálculo de la estimación de disparidad aumenta enormemente con el tamaño del fotograma, de esa manera trabajar con fotogramas más pequeños comprime el rango de búsqueda de disparidad y el número de píxeles. Una relación de sub-muestreo típica es de alrededor de 2 o 4. Sin embargo, es importante tener en cuenta que en la etapa final de detección del rango de disparidad, el rango se compensa aumentado a escala con la misma relación para que esté a la misma escala que los fotogramas izquierdo y derecho. En particular, la operación de sub-muestreo es opcional y no es fundamental para detectar el rango de disparidad con una precisión óptima. Solo proporciona una mayor reducción de la complejidad de cálculo.

[0034] La tercera etapa del procedimiento de acuerdo con la invención como se muestra en la figura 2, es la detección de píxeles distintivos (30) y juega un papel importante para aumentar la robustez. Los píxeles distintivos en esta invención se pueden determinar de dos maneras, a saber, detección en bordes y extracción de prominencia. En la detección de bordes, se buscan los bordes verticales ya que el rango de disparidad que se detectará implica solo un desplazamiento horizontal. Los bordes horizontales pueden no proporcionar coincidencia fiable en dirección horizontal, como se ilustra en la figura 3, por lo tanto, los dos candidatos de disparidad (d_1 , d_2) no proporcionan distinción para los píxeles en el borde horizontal, donde la diferencia es clara para el borde vertical. Las disparidades de los píxeles en un borde vertical son más distintivas y se pueden obtener coincidencias únicas. Sin embargo, para el borde horizontal, el número de posibles disparidades entre candidatos es casi igual a la longitud del borde; por lo tanto, no se pueden asignar asignaciones consistentes de disparidad. Para determinar los bordes verticales, se aplica un filtro de paso alto a las vistas izquierda y derecha como se detalla en las ecuaciones (A):

$$\begin{aligned} \text{Izquierda borde } (x, y) &= \left| \sum_{i=-M}^M h(i) I_L(x+i, y) \right| \\ \text{Derecha borde } (x, y) &= \left| \sum_{i=-M}^M h(i) I_R(x+i, y) \right| \end{aligned}$$

(A)

[0035] En las ecuaciones (1), h corresponde a un filtrado paso alto con longitud $(2M + 1)$ que se puede elegir arbitrariamente; I_L e I_R corresponden a los fotogramas izquierdo y derecho. Una vez que se obtienen los mapas de borde de la vista izquierda (107) y la vista derecha (108), los píxeles que tienen un valor de borde mayor que un umbral se seleccionan para ser distintivos y se explotan en el siguiente paso de cálculo de costo (40).

[0036] Además de los píxeles de borde vertical, los píxeles prominentes también se pueden explotar de la siguiente manera. Los píxeles prominentes se determinan en función de las diferencias de centro y contorno que corresponden a una medida de distinción de píxeles de sus contornos, lo que significa que los píxeles salientes se comparan con los píxeles vecinos. De esa manera, las intensidades de píxeles se restan del valor medio de sus alrededores. El valor medio envolvente de cada píxel se calcula tomando la intensidad promedio entre ciertas ventanas. La prominencia de un píxel para un tamaño de ventana dado se calcula como en la ecuación (B); donde h corresponde al índice del tamaño de ventana utilizado:

$$C_k(x, y) = I_{RGB}(x, y) - \frac{1}{w * h} \sum_{i=-w/2}^{w/2} \sum_{j=-h/2}^{h/2} I_{RGB}(x+i, y+j)$$

(B)

[0037] Al utilizar tamaños de ventana grandes durante la estimación del mapa de prominencia, por tanto calculando los valores medios dentro de la ventana para cada píxel de forma independiente requiere un gran procesamiento. Sin embargo, utilizando una imagen integral (109) disminuye drásticamente la complejidad de cálculo. Una imagen integral (109) se define mediante la ecuación (C) y el valor medio dentro de un tamaño de ventana arbitrario se puede hallarse mediante $(P4 + P1-P2-P3)$ como se ilustra en la figura 4.

$$\text{Integral } (x, y) = \sum_{i=1}^x \sum_{j=1}^y I_{RGB}(i, j)$$

(C)

[0038] Por lo tanto, la complejidad de la operación es independiente del tamaño de la ventana, ya que para cada tamaño de ventana solo se requieren cuatro operaciones. Después de calcular las diferencias centro-contorno para cada píxel a través de cinco tamaños de ventana predeterminados diferentes, el mapa de prominencia final se obtiene sumando directamente esos valores de acuerdo con la ecuación (D). De esta forma, se obtiene la medida de cómo se diferencian los píxeles de sus entornos.

$$\text{Mapa Prominencia } (x, y) = \sum_{i=1}^5 |C_i(x, y)|$$

(D)

[0039] La extracción del mapa de prominencia se realiza para las imágenes izquierda y derecha y los píxeles que tienen valores de prominencia más grande que un umbral se consideran píxeles distintivos en ambas imágenes.

[0040] De este modo, los puntos característicos que se utilizarán en las unidades de detección de rango de disparidad y cálculo de costes se determinan mediante detección de borde o prominencia. Estos píxeles proporcionan una medida de costo fiable, ya que implican características distintivas.

[0041] La siguiente etapa del algoritmo de detección de rango de disparidad es el cálculo del costo (40). Inicialmente, un rango de búsqueda de disparidad se resuelve considerando el ancho del fotograma, como (-ancho / 5, ancho / 5). Se supone que, en cualquier contenido estéreo, la disparidad fuera de este rango no puede proporcionar una sensación 3D confortable. La elección del rango de búsqueda inicial puede ser (ancho, ancho) también, sin embargo, tener tales disparidades no tiene sentido en términos de percepción 3D. Para cada uno de los píxeles distintivos extraídos en la etapa anterior, los valores de costo se calculan dentro del rango de disparidad especificado para los fotogramas izquierdo y derecho de forma independiente. Durante el cálculo del costo, se puede explotar la suma de la distancia absoluta o cualquier otra medida de similitud, como se muestra en las ecuaciones (E).

$$\begin{aligned} \text{Costo Izquierda } (x, y, d) &= \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IL(x+i, y+j) - IR(x+i+d, y+j)| \\ \text{Costo Derecha } (x, y, d) &= \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IR(x+i, y+j) - IL(x+i-d, y+j)| \end{aligned}$$

(E)

[0042] En las ecuaciones (E), los valores de costo de disparidad (d) se calculan dentro de una ventana (w, h) alrededor de la correspondiente píxel (x, y) para disminuir el efecto del ruido en los fotogramas. En una implementación típica, w y h se pueden elegir en particular para ser 5. Luego, las mejores asignaciones de disparidad (d1, d2 y d3) se determinan minimizando la función de costo, ya que el ruido y los patrones repetidos pueden inducir a error en la minimización. Si las tres mejores asignaciones de disparidad están escasamente separadas entre sí, entonces se concluye que el píxel correspondiente no proporciona una asignación de disparidad fiable. La razón subyacente en este enfoque se ilustra en la figura 5 donde se dan dos posibles distribuciones de funciones de costos. En la gráfica superior, las tres mejores asignaciones están mutuamente próximas, lo que indica que el píxel correspondiente proporciona una estimación de disparidad fiable. Sin embargo, para la gráfica de abajo, las mejores asignaciones están separadas posiblemente debido a un patrón repetido o a una superficie sin textura, lo que proporciona características de disparidad no fiables. Por lo tanto, para los píxeles distintivos se realiza una comprobación de fiabilidad (50) en la etapa entre las tres mejores asignaciones de disparidad para detectar píxeles fiables en términos de distribución de costes entre las disparidades candidatas como en las ecuaciones (F). En las ecuaciones (F), "DispThres" se utiliza para verificar si los dos valores de disparidad están más cerca el uno del otro que un umbral. En una ejecución típica, este valor se puede establecer dentro de un rango de (1, 5).

$$R(x, y) = \begin{cases} 1 & \text{si } |d1 - d2| < \text{DispThres y } |d2 - d3| < \text{DispThres} \\ 0 & \text{en otro caso} \end{cases}$$

(F)

[0043] Además de la mejor comprobación de asignación de tres disparidades, la distribución de los valores de costo también se considera determinar la fiabilidad de los píxeles. De esa manera, para cada píxel, se calcula la cantidad de disparidades candidato cuyo costo es más próximo a un umbral al costo mínimo. Si el número es alto, entonces

se puede decir que la distribución del costo no es suficiente para ser fiable. Como se ilustra en la figura 6, las funciones de costo pueden satisfacer la primera condición (cerca a tres mejores disparidades); sin embargo, no proporcionan buenas coincidencias. Este caso ocurre cuando la superficie no está texturizada de alguna manera. En esa etapa, también se puede considerar la varianza de la función de costo, de modo que si la varianza de la función de costo de un píxel es alta, es decir, una distribución uniforme de valores de costo, ese píxel se marca como no fiable. Estas dos comprobaciones de fiabilidad permiten manejar errores que pueden resultar de patrones repetidos o superficies sin textura. Una vez que se determinan los píxeles que dan dos condiciones del control de fiabilidad, se obtiene el histograma de las disparidades (60), considerando las tres mejores asignaciones de disparidad de estos píxeles. En ese punto, el recuento en el histograma se aumenta para cada tres asignaciones de disparidad; de ahí que se calcule el histograma de las mejores tres estimaciones en lugar de las mejores estimaciones, que disminuyen el efecto del ruido durante el cálculo del costo. En esta etapa, el rango de disparidad se detecta (70) al encontrar las disparidades mínima y máxima en el histograma cuyas frecuencias son superiores a 1%, como se ilustra en la figura 7. La última etapa es aplicar un filtrado temporal (90), en el que la ecuación (G) puede usarse preferiblemente para eliminar posibles cambios abruptos en el rango de disparidad debido al ruido temporal. En la naturaleza del video estereoscópico, el rango de disparidad cambia lentamente siempre que no haya cambios de escena o movimientos muy rápidos (comúnmente no se observados).

$$\begin{aligned} \text{disp_min} &= 0.05 * \text{disp_min} + 0.95 * \text{previous_disp_min} \\ \text{disp_max} &= 0.05 * \text{disp_max} + 0.95 * \text{previous_disp_max} \end{aligned} \quad (G)$$

- 20 **[0044]** Como se mencionó anteriormente, es una opción realizar un aumento de escala (80) en la etapa 80 en caso que el contenido estéreo se haya sub-muestreado (20) previamente. Sin embargo, esto no es necesario para llevar a cabo la invención.
- [0045]** A continuación, se describe brevemente una realización preferida de la invención, en la que en la primera etapa los fotogramas estéreo se calibran (10) en términos de color por equalización de histograma.
- 25 **[0046]** Como segunda etapa, los fotogramas estéreo son sub-muestreados (20) en un factor de 2 o 4 para ahorrar cálculo, lo cual es solo una opción, como se explicó anteriormente.
- [0047]** En la tercera etapa, los píxeles distintivos se detectan (30) en ambos fotogramas izquierdo y derecho a través de detección de borde o extracción de mapas de salientes, en la que se eligen los píxeles con valores de borde alto o medidas de elevada prominencia.
- 30 **[0048]** En la cuarta etapa el cuarto paso, los valores de costo se calculan (50) para valores de disparidad candidatos que se establecen para estar dentro de $(-\text{fotograma_ancho} / 5, \text{fotograma_ancho} / 5)$ mediante las ecuaciones (E).
- [0049]** En la etapa 5, los mejores tres candidatos de disparidad se determinan para cada píxel distintivo en función del menor costo.
- 35 **[0050]** En la etapa 6, se verifica si los tres mejores candidatos de disparidad están mutuamente próximos, si no, se ignora el píxel correspondiente.
- [0051]** En la etapa 7, la fiabilidad de las funciones de coste se determina contando los candidatos de disparidad cuyo valor de coste está más cerca que un umbral del costo mínimo, en el que el píxel correspondiente se ignora en caso de que el número sea superior al 10% del total de candidatos.
- 40 **[0052]** En la etapa 8, se construye un histograma de disparidad (60) considerando los mejores tres candidatos de disparidad de los píxeles con fiabilidad detectados en las etapas 6 y 7.
- [0053]** En la etapa 9, se detectan las disparidades mínima y máxima (70), en donde se compara si la frecuencia del histograma total es al menos el 1%.
- [0054]** En la etapa 10, los límites de disparidad son aumentados de escala (80) por la relación de sub-muestreo para compensar una opción sub-muestreo que podría haberse realizado en la segunda etapa de esta realización.
- 45 **[0055]** En la etapa 11, se aplica un filtrado temporal (90) para el rango de disparidad del fotograma anterior para eliminar parpadeo, donde se usan las ecuaciones (G): de acuerdo con la invención, se describen un procedimiento y un dispositivo utilizando píxeles visualmente distintivos en videos estéreo (multivistas) para detectar el rango de disparidad, indicando valores de disparidad máximo y mínimo disponibles.
- 50 **[0056]** De acuerdo con la invención, inicialmente se realiza una calibración del video estéreo con respecto al color para que los fotogramas izquierdo y derecho se originen en la misma gama de colores. En el caso de video de vistas múltiples, dos cámaras centrales son elegidas para ser las vistas izquierda y derecha y el rango de disparidad detectado se aplica a todo el equipo de cámara. Una vez que llevada a cabo la equalización de color, los fotogramas estéreo se pueden sub-muestrear para reducir aún más la complejidad del cálculo en lo siguiente. La detección de píxeles distintivos se puede lograr de dos maneras diferentes, a saber, mediante detección de borde vertical y extracción de mapa de prominencia. Los píxeles que tienen un borde alto o medidas de prominencia son asignados como distintivos. Se realiza una búsqueda de disparidad para los píxeles distintivos dentro de un rango fijo predeterminado, en particular tal como $(-\text{ancho} / 5, \text{ancho} / 5)$ explotando cualquier medida de similitud, suma de distancia absoluta, distancia absoluta media, error cuadrático medio. Los tres mejores candidatos de disparidad que minimizan la función de costo se almacenan para todos los píxeles distintivos y se realiza una verificación de fiabilidad según la distribución de la función de costo y los mejores candidatos de disparidad. Un histograma de disparidad está formado por los píxeles distintivos fiables y las tres mejores disparidades correspondientes. La detección de fiabilidad maneja posibles errores de asignación debido a regiones sin textura y patrones repetidos.
- 55 **[0056]** De acuerdo con la invención, inicialmente se realiza una calibración del video estéreo con respecto al color para que los fotogramas izquierdo y derecho se originen en la misma gama de colores. En el caso de video de vistas múltiples, dos cámaras centrales son elegidas para ser las vistas izquierda y derecha y el rango de disparidad detectado se aplica a todo el equipo de cámara. Una vez que llevada a cabo la equalización de color, los fotogramas estéreo se pueden sub-muestrear para reducir aún más la complejidad del cálculo en lo siguiente. La detección de píxeles distintivos se puede lograr de dos maneras diferentes, a saber, mediante detección de borde vertical y extracción de mapa de prominencia. Los píxeles que tienen un borde alto o medidas de prominencia son asignados como distintivos. Se realiza una búsqueda de disparidad para los píxeles distintivos dentro de un rango fijo predeterminado, en particular tal como $(-\text{ancho} / 5, \text{ancho} / 5)$ explotando cualquier medida de similitud, suma de distancia absoluta, distancia absoluta media, error cuadrático medio. Los tres mejores candidatos de disparidad que minimizan la función de costo se almacenan para todos los píxeles distintivos y se realiza una verificación de fiabilidad según la distribución de la función de costo y los mejores candidatos de disparidad. Un histograma de disparidad está formado por los píxeles distintivos fiables y las tres mejores disparidades correspondientes. La detección de fiabilidad maneja posibles errores de asignación debido a regiones sin textura y patrones repetidos.
- 60 La detección del rango de disparidad finalmente se logra al determinar los candidatos de disparidad mínima y máxima que tienen frecuencias de ocurrencia superiores al 1% en el histograma de disparidad. Para evitar cambios bruscos

del rango de disparidad en el tiempo, el filtrado temporal se realiza ponderando los rangos de disparidad de cuadro estéreo actual y anterior.

REIVINDICACIONES

1. Procedimiento para detectar un rango de disparidad de contenido estéreo, en particular, tal como video digital estéreo, imágenes digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento tal como un disco duro, escáner, disco óptico o similar de la vista izquierda y la vista derecha de, al menos, dos cámaras, que comprende las siguientes etapas:

- calibrar el contenido estéreo (10) con respecto a la vista izquierda y la vista derecha,
- detectar píxeles distintivos en el contenido estéreo (30), en particular a través de la detección de bordes o la extracción del mapa de prominencia,
- calcular los valores de coste (40) para cada píxel distintivo detectado dentro de una ventana de, al menos, un fotograma estéreo, teniendo un ancho y una altura,
- determinar los valores de disparidad candidatos (50) con vistas a una minimización de los valores de costes calculados, seleccionando candidatos de disparidad con valores de costo mínimo, para extraer píxeles fiables de los valores de disparidad candidatos, en el que se dice que un píxel es fiable cuando dicho píxel tiene asignaciones de disparidad consistentes,
- crear un histograma (60) a partir de los píxeles fiables en el que se determinan los candidatos de disparidad mínimo y máximo teniendo en cuenta un umbral del histograma para detectar el rango de disparidad del contenido estéreo (70)

caracterizado porque la verificación de fiabilidad de los píxeles se realiza mediante la siguiente ecuación;

$$R(x, y) = \begin{pmatrix} 1 \text{ si } |d1 - d2| < DispThres \text{ y } |d2 - d3| < DispThres \\ 0 \text{ en otro caso} \end{pmatrix} \quad (F)$$

en la que d_1 , d_2 y d_3 son las mejores tres asignaciones de disparidad del píxel, siendo $DispThres$ un valor umbral y siendo $R(x, y)$ el resultado del control de fiabilidad.

2. Procedimiento según la reivindicación 1, caracterizado porque el cálculo de los valores de coste (40) de cada píxel distintivo detectado dentro de una ventana de, al menos, un fotograma estéreo, que tenga un ancho de w y una altura de h , mediante las siguientes ecuaciones:

$$Costo Izquierda(x, y, d) = \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IL(x+i, y+j) - IR(x+i+d, y+j)|$$

$$Costo Derecha(x, y, d) = \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IR(x+i, y+j) - IL(x+i-d, y+j)|$$

donde d se refiere a la disparidad de cada píxel (x, y) que está rodeado por la ventana (w, h) , correspondiendo I_L e I_R a intensidades de píxel de fotogramas derecho e izquierdo.

3. Procedimiento de acuerdo con la reivindicación 1 o 2, caracterizado porque después de la calibración del contenido estéreo (10), se realiza un sub-muestreo de los fotogramas izquierdo y derecho (20) antes de detectarse los píxeles distintivos.

4. Procedimiento de acuerdo con la reivindicación 3, caracterizado porque después de determinar los valores de disparidad candidatos (50), se lleva a cabo un aumento de escala (80) opcional en caso que el contenido estéreo se sub-muestree después de la calibración estéreo.

5. Procedimiento según al menos una de las reivindicaciones precedentes 1 a 4, caracterizado porque el control de fiabilidad de los píxeles se realiza mediante

- para cada píxel, calculando el número de disparidades de candidato cuyo costo está más próximo a un umbral al costo mínimo
- si el número calculado es alto, determinar que la distribución de costos no es suficiente para ser fiable.

6. Dispositivo para detectar un rango de disparidad de contenido estéreo, en particular tal como video estéreo digital, imágenes digitales, fotogramas estéreo de impresión de película escaneada a través de un sistema de transmisión en directo en tiempo real o a través de un medio de almacenamiento tal como un disco duro, escáner, disco opcional o similar de la vista izquierda y de la vista derecha de, al menos, dos cámaras que comprenden:

- un calibrador estéreo (1) con respecto a la vista izquierda y la vista derecha del contenido estéreo,

- un detector de píxeles distintivo (3) para detectar píxeles distintivos en el contenido estéreo, en particular a través de la detección de bordes o la extracción del mapa de prominencia,
 - un calculador de costes (4) para calcular los valores de coste para cada píxel distintivo detectado dentro de una ventana de, al menos, un fotograma estéreo, que tiene un ancho y una altura,
 - medios para determinar los valores de disparidad candidatos (5) con vistas a una minimización de los valores de costo calculados, seleccionando candidatos de disparidad con valores de costo mínimo, para extraer píxeles fiables de los valores de disparidad candidatos, en los que se dice que un píxel es fiable cuando dicho píxel tiene asignaciones de disparidad consistentes,
 - medios para crear un histograma (6), basándose en los píxeles fiables, donde el máximo y el mínimo de candidatos de disparidad se determinan a la vista de un umbral de un histograma para detectar el rango de disparidad del contenido estéreo
- caracterizado porque dichos medios para determinar los valores de disparidad de candidato (5) extraen los píxeles fiables de acuerdo con la siguiente ecuación;

$$R(x, y) = \begin{cases} 1 & \text{si } |d1 - d2| < DispThres \text{ y } |d2 - d3| < DispThres \\ 0 & \text{en otro caso} \end{cases} \quad (F)$$

donde d_1 , d_2 y d_3 son las mejores tres asignaciones de disparidad del píxel, $DispThres$ es un valor de umbral y siendo $R(x, y)$ el resultado del control de confiabilidad.

7. Dispositivo según la reivindicación 6, caracterizado porque el calculador de costes (4) calcula los costes (40) de cada píxel distintivo detectado dentro de una ventana de, al menos, un fotograma estéreo, que tiene un ancho de w y una altura de h mediante las siguientes ecuaciones:

$$Costo Izquierda(x, y, d) = \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IL(x+i, y+j) - IR(x+i+d, y+j)|$$

$$Costo Derecha(x, y, d) = \sum_{j=-h/2}^{h/2} \sum_{i=-w/2}^{w/2} |IR(x+i, y+j) - IL(x+i-d, y+j)|$$

donde d se refiere a la disparidad de cada píxel (x, y) que está rodeado por la ventana (w, h) , correspondiendo I_L e I_R corresponden a intensidades de píxel de los fotogramas izquierdo y derecho.

8. Dispositivo según la reivindicación 6 o 7, caracterizado porque un elemento de sub-muestreo (2) se incorpora dentro del dispositivo que realiza el sub-muestreo de los fotogramas derecho e izquierdo (20) antes de que se detecten los píxeles distintivos después de la calibración del contenido estéreo (10).

9. Dispositivo según la reivindicación 8, caracterizado porque se proporciona un elemento de aumento de escala (8) dentro del dispositivo de aumento de escala (80) de los datos después de determinar los valores de disparidad candidato (50).

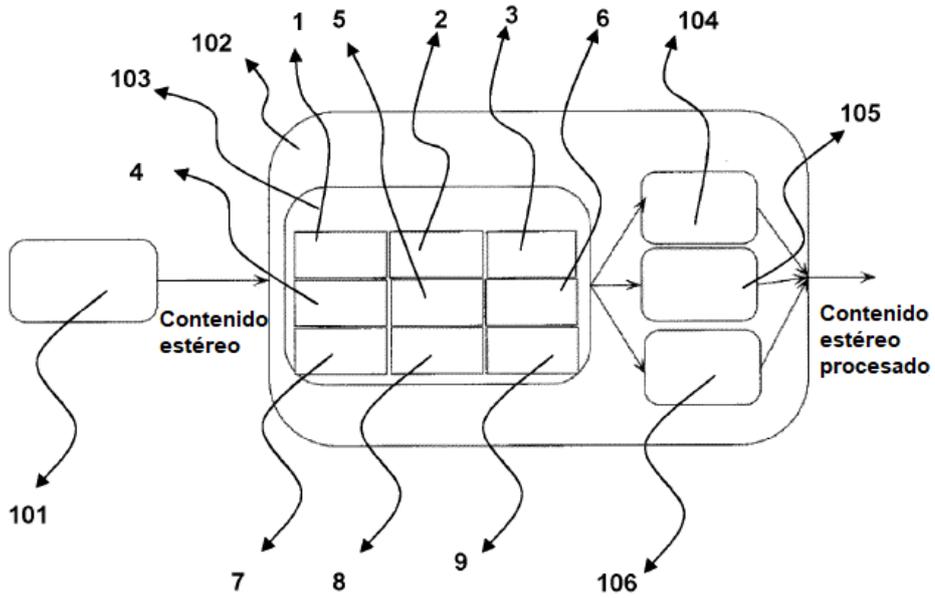


Figura 1

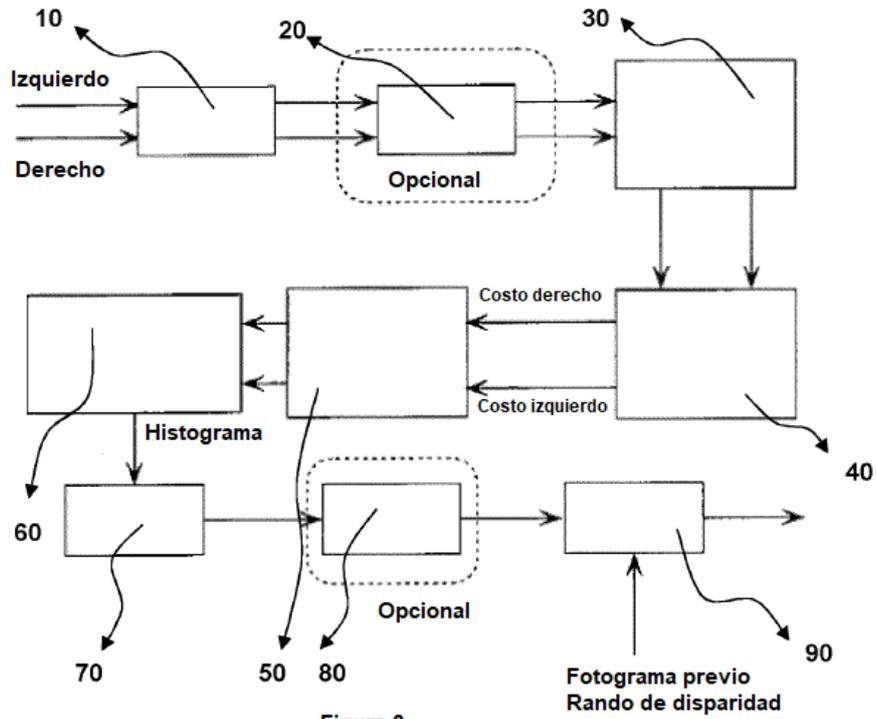


Figura 2

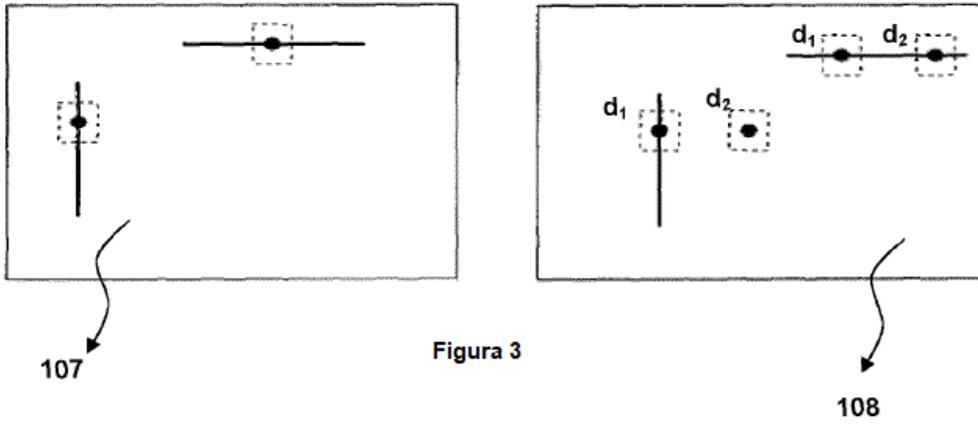


Figura 3

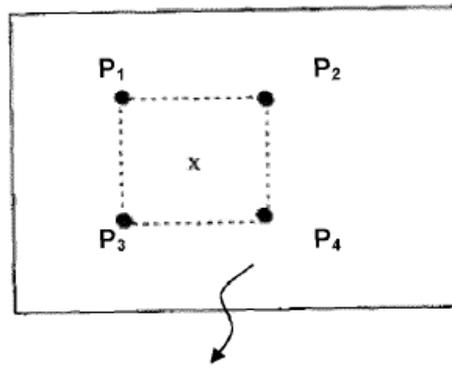


Figura 4

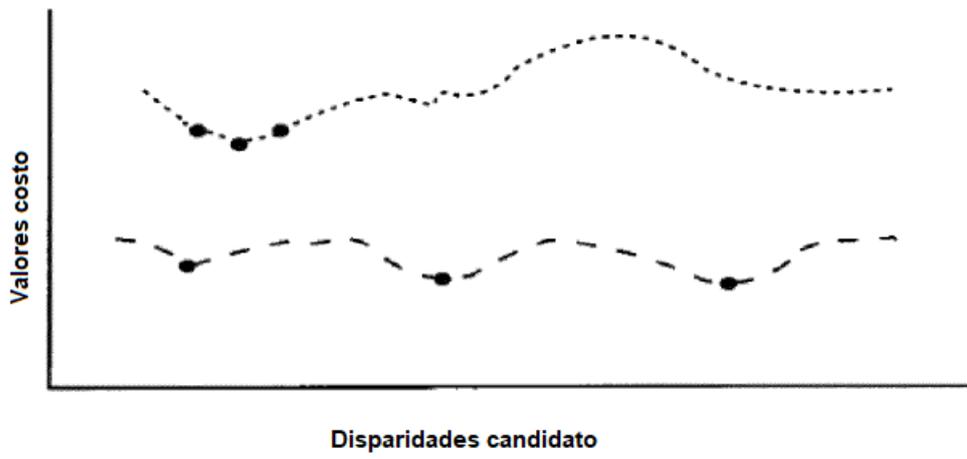


Figura 5

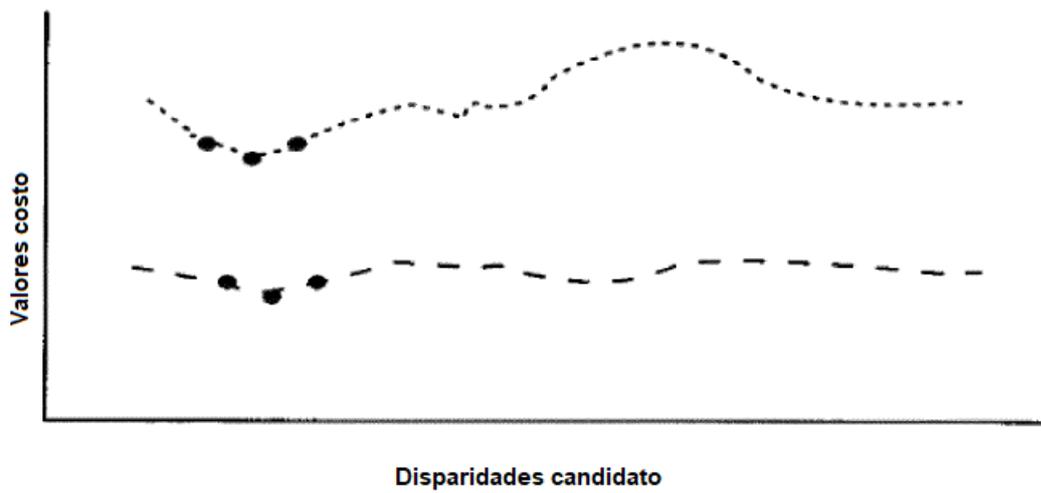


Figura 6

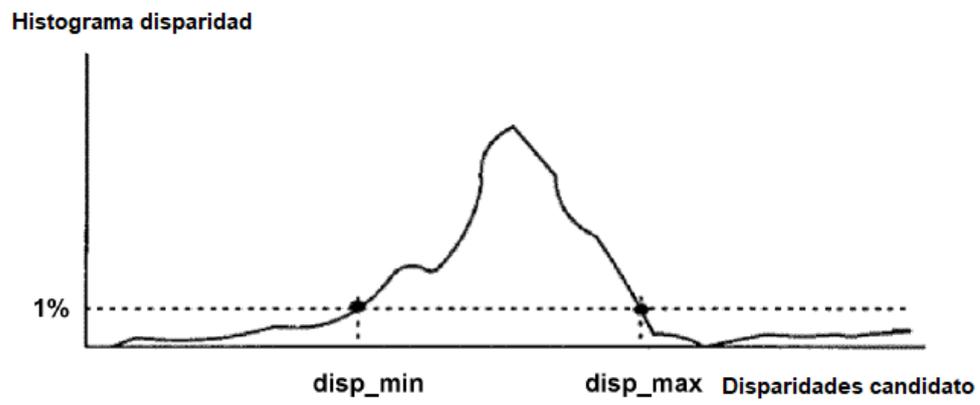


Figura 7

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

5 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.

Documentos de patente citado en la descripción

- US 6226396 B1 [0007]
- US 2007255480 A1 [0008]

10 **Bibliografía no de patentes citada en la descripción**

- **ULRICH FECKER.** Histogram- Based Prefiltering for Luminance and Chrominance Compensation of Multiview Video. *Marcus Barkowsky and Andre Kaup, IEEE transactions on Circuits and Systems for Video Technology*, September 2008, vol. 18 (9), 1258-1267 [0029]