



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

(21) BR 112019016041-8 A2



(22) Data do Depósito: 31/01/2018

(43) Data da Publicação Nacional: 31/03/2020

(54) **Título:** APARELHO PARA CAPTURAR UM MAPA DE PROPRIEDADES DE IMAGEM, MÉTODO DE PROCESSAMENTO DE UM MAPA DE PROPRIEDADES DE IMAGEM E PRODUTO DE PROGRAMA DE COMPUTADOR

(51) **Int. Cl.:** H04N 13/00; G06T 5/00; H04N 13/128.

(30) **Prioridade Unionista:** 07/02/2017 EP 17155076.7.

(71) **Depositante(es):** KONINKLIJKE PHILIPS N.V..

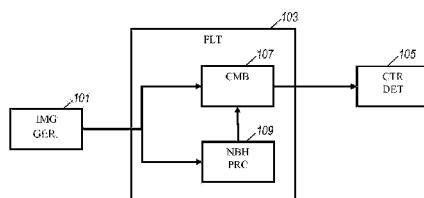
(72) **Inventor(es):** CHRISTIAAN VAREKAMP; PATRICK LUC ELS VANDEWALLE.

(86) **Pedido PCT:** PCT EP2018052318 de 31/01/2018

(87) **Publicação PCT:** WO 2018/145961 de 16/08/2018

(85) **Data da Fase Nacional:** 02/08/2019

(57) **Resumo:** A presente invenção se refere a um aparelho que compreende um receptor (101) que recebe uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem. Uma unidade de filtro (103) é disposta para filtrar o mapa de propriedades de imagem em resposta à imagem de intensidade de luz e ao mapa de confiança. Especificamente, para uma primeira posição, a unidade de filtro (103) determina um valor de propriedade de imagem combinado da vizinhança em resposta a uma combinação ponderada de valores de propriedade de imagem vizinhos em uma vizinhança em torno da primeira posição, o peso para um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição ser dependente de um valor de confiança para o primeiro valor de propriedade vizinho e de uma diferença entre valores de intensidade de luz para a primeira posição e a segunda posição; e determina um primeiro valor de propriedade de imagem filtrado para a primeira posição como uma combinação de um primeiro valor de propriedade de imagem na primeira posição no mapa de propriedades de imagem e o valor de propriedade de imagem combinado.



APARELHO PARA CAPTURAR UM MAPA DE PROPRIEDADES DE IMAGEM, MÉTODO DE PROCESSAMENTO DE UM MAPA DE PROPRIEDADES DE IMAGEM E PRODUTO DE PROGRAMA DE COMPUTADOR

Campo da invenção

[001] A invenção se refere a um método e a um aparelho para o processamento do mapa de propriedades de uma imagem e, em particular, mas não exclusivamente, para o processamento de um mapa de profundidade ou de disparidade.

Antecedentes da invenção

[002] Tradicionalmente, imagens digitais e quadros de vídeo simplesmente constituíam uma única imagem fornecendo informações de intensidade de luz (incluindo a cor) que refletiam as propriedades de luz da cena. No entanto, nos últimos anos, a representação de imagens se tornou cada vez mais complexa com informações adicionais que representam, muitas vezes, várias propriedades diferentes. Por exemplo, ao invés de uma imagem que consiste simplesmente de uma única imagem de intensidade de luz, muitas representações de uma cena compreendem atualmente imagens adicionais de intensidade de luz, mapas de profundidade, mapas de transparência, mapas de confiança etc.

[003] Tais informações adicionais podem ser fornecidas para possibilitar que a parte de recepção/renderização tenha mais flexibilidade e controle sobre a imagem sendo renderizada. Um exemplo é a disponibilização de mais informações de profundidade, por exemplo o fornecimento de um mapa de profundidade (ou um mapa de disparidade equivalentemente) para a imagem de intensidade de luz. Isso pode possibilitar que o receptor/renderizador

faça um deslocamento de ponto de vista ou que, por exemplo, gere imagens estéreo para visualização em 3D.

[004] No entanto, em muitas situações práticas, as propriedades de imagem podem ser imperfeitas e afetadas por ruído ou erros. Por exemplo, as informações de profundidade podem não ser tão precisas quanto o desejado. De fato, em muitos cenários, as informações de profundidade são geradas por estimativa e pela extração de valores de profundidade por meio da comparação de vistas de diferentes direções de visão. Entretanto, tais estimativas tendem a ser relativamente propensas a ruído e a erro.

[005] Especificamente, em muitas aplicações, são capturadas três cenas tridimensionais como imagens estéreo com o uso de duas câmeras em posições ligeiramente diferentes. Os valores de profundidade específicos podem, então, ser gerados mediante estimativa das disparidades entre objetos de imagem correspondentes nas duas imagens. Entretanto, tal extração e estimativa de profundidade é problemática e tende a resultar em valores de profundidade não ideais. Isto pode resultar novamente em artefatos e em uma qualidade de imagem tridimensional degradada.

[006] Os artefatos e a degradação da imagem tridimensional tendem a ser particularmente significativos para transições entre diferentes objetos de imagem. Adicionalmente, a determinação de informações de profundidade com base na estimativa de disparidade para imagens associadas também está geralmente relacionada à consideração de características de objetos de imagem. Geralmente, os algoritmos de estimativa de disparidade buscam correspondências entre uma imagem esquerda e uma imagem

direita, por meio de comparação local de diferenças de cor entre um ponto na imagem esquerda e seu ponto correspondente na imagem direita. No entanto, as propriedades de imagem e o processo podem ter um nível relativamente alto de ruído, fazendo com que o mapa de profundidade gerado também tenha um nível relativamente alto de ruído.

[007] Para melhorar tais mapas de profundidade, foi proposta a pós-filtragem do mapa de profundidade. Uma abordagem específica de pós-filtragem é um filtro adaptativo bilateral de cor e/ou de luminância no qual o núcleo de filtragem é adaptado para refletir as propriedades visuais da imagem. Esse tipo de filtro bilateral pode causar a adaptação do mapa de profundidade para que ele siga mais de perto as características da imagem, e pode resultar em melhor consistência e estabilidade temporal das disparidades estimadas, ou pode, por exemplo, fornecer uma transição de profundidade mais brusca entre objetos de imagem diferentes.

[008] A filtragem bilateral conjunta pode incluir um aumento artificial da taxa de amostragem - *upsampling* - e é uma técnica usada para filtrar ou aumentar artificialmente a taxa de amostragem de, por exemplo, um mapa de profundidade a partir de uma imagem de intensidade de luz associada como um sinal de "orientação". A idéia geral é que o sinal de "orientação" tem menos ruído ou tem resolução mais alta e que ele pode, dessa forma, fornecer informações relacionadas a quais amostras devem receber peso menor ou maior em um cálculo da média ponderada de uma vizinhança tipicamente espaço-temporal em torno do pixel a ser filtrado.

[009] Entretanto, embora tal filtragem, e especificamente a filtragem bilateral, possa melhorar o mapa

da propriedades de imagem que está sendo filtrado, ela tende a não fornecer um desempenho ideal e o mapa de propriedades de imagem resultante pode ainda incluir artefatos, erros, imprecisões e/ou ruído.

[010] Portanto, uma abordagem aprimorada do processamento de um mapa de propriedades de imagem seria vantajosa e, em particular, seria vantajosa uma abordagem que possibilitasse maior flexibilidade, implementação mais fácil, menor complexidade, redução de artefatos/erros/imprecisões e/ou ruído, e/ou qualidade e/ou desempenho aprimorados.

#### Sumário da invenção

[011] Consequentemente, a invenção busca, de preferência, mitigar, aliviar ou eliminar uma ou mais das desvantagens mencionadas acima, individualmente ou em qualquer combinação.

[012] De acordo com um aspecto da invenção, é fornecido um aparelho para processar um mapa de propriedades de imagem, sendo que o aparelho compreende: um receptor para receber uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem, sendo que a imagem de intensidade de luz compreende valores de intensidade de luz para pixels da imagem de intensidade de luz, sendo que o mapa de propriedade da imagem compreende valores de propriedade da imagem para os pixels da imagem de intensidade de luz, e que o mapa de confiança compreende valores de confiança para os valores de propriedade de imagem; uma unidade de filtro disposta para filtrar o mapa de propriedades de imagem em resposta à imagem de intensidade de luz e ao mapa de confiança, para gerar um mapa de propriedades de imagem filtrado, sendo que a unidade de filtro é disposta de modo a, para uma primeira

posição no mapa de propriedade da imagem filtrado: determinar um valor de propriedade de imagem da vizinhança em resposta a uma combinação ponderada de valores de propriedade de imagem vizinhos ser valores de propriedades de imagem vizinhos em torno da primeira posição, sendo que o peso por um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição é dependente de um valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e uma diferença entre um valor de intensidade de luz para a primeira posição e um valor de intensidade de luz para a segunda posição; e determinar o primeiro valor de propriedade da imagem filtrado para a primeira posição como uma combinação de um primeiro valor de propriedade de imagem na primeira posição no mapa de propriedades de imagem e o valor de propriedade de imagem da vizinhança combinados; sendo que a unidade de filtro é disposta de modo a determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a uma medida de diferença para o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e a um valor de confiança para o primeiro valor de propriedade de imagem, sendo que a determinação do peso é assimétrica em relação ao valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o valor de propriedade da primeira imagem, com o peso dependendo do sinal da medida de diferença.

[013] A invenção pode possibilitar a geração de um mapa de propriedades de imagem melhorado. O mapa de propriedades de imagem filtrado pode ser determinado em resposta a diferentes propriedades e modalidades e pode fornecer um resultado mais exato dos valores corretos de

propriedades da imagem em muitos cenários e modalidades. Por exemplo, em muitas modalidades, pode ser obtida uma filtragem melhorada de um mapa de profundidade, resultando em um mapa de profundidade de saída filtrado que pode ser usado, por exemplo, para fazer deslocamentos de ponto de vista com maior qualidade da imagem resultante. Dessa forma, uma experiência de uso melhorada pode, muitas vezes, ser alcançada com base no mapa de propriedades de imagem filtrado resultante.

[014] A abordagem pode fornecer uma filtragem melhorada na qual tanto os valores de confiança como os valores de intensidade de luz podem ser usados simultaneamente para controlar e guiar uma filtragem de um mapa de propriedades de imagem, como um mapa de profundidade. O uso das diferentes propriedades pode ser integrado em uma combinação de operações simultâneas e integradas. Isso pode reduzir a qualidade e fornecer melhor desempenho. Por exemplo, a abordagem específica possibilita o uso conjunto de ambos os valores de confiança e de intensidade de luz sem a interferência indesejada de um com o outro e sem o surgimento de erros ou artefatos.

[015] A diferença entre o valor de intensidade de luz para a primeira posição e o valor de intensidade de luz para a segunda posição pode ser determinada de acordo com qualquer medida de diferença adequada. Por exemplo, ela pode ser determinada como uma norma para a diferença entre os valores de intensidade de luz.

[016] A combinação ponderada de valores de propriedade da imagem vizinhos pode especificamente ser uma soma ponderada de valores de propriedades de imagem vizinhos. Em algumas modalidades, a combinação ponderada de valores

vizinhos de propriedades de imagem pode incluir todos os valores de propriedades de imagem vizinhos dentro da vizinhança. Em outras modalidades, ela pode incluir apenas um subconjunto desses valores, por exemplo uma subamostragem pode ser aplicada aos valores vizinhos de propriedades de imagem da vizinhança.

[017] A combinação do primeiro valor de propriedade da imagem na primeira posição no mapa de propriedades de imagem e do valor de propriedade de imagem vizinho combinado pode ser uma soma ponderada.

[018] A vizinhança pode ser uma vizinhança/região que inclui a primeira posição. Tal vizinhança/região pode também ser chamada de núcleo.

[019] O peso do primeiro valor de propriedade de imagem da vizinhança pode ser uma função monotonicamente crescente do valor de confiança para o primeiro valor de propriedade de imagem da vizinhança (onde o valor de confiança tem um valor crescente para aumentar a confiança).

[020] O peso do primeiro valor de propriedade de imagem da vizinhança pode ser uma função monotonicamente decrescente da diferença entre o valor de intensidade de luz para a primeira posição e o valor de intensidade de luz para a segunda posição

[021] O valor de confiança para o primeiro valor de propriedade de imagem da vizinhança pode ser o valor de confiança na segunda posição no mapa de confiança.

[022] A medida de diferença para os valores de confiança pode ser qualquer medida de diferença.

[023] De acordo com um recurso opcional da invenção, a unidade de filtro é disposta para:



[024] determinar um peso para o primeiro valor de propriedade de imagem em relação a um peso para o valor de propriedade de imagem da vizinhança combinado em resposta a um valor de confiança para o primeiro valor de propriedade de imagem.

[025] Isso pode resultar na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[026] Os pesos para a combinação do primeiro valor de propriedade da imagem na primeira posição do mapa de propriedades de imagem e do valor combinado de propriedade da imagem vizinho podem depender do valor de confiança para o valor da primeira propriedade da imagem / na primeira posição no mapa de confiança.

[027] O peso para o valor da primeira propriedade da imagem em relação ao peso para o valor de propriedade de imagem vizinho combinado pode ser uma função monotonicamente crescente do valor de confiança para o primeiro valor de propriedade da imagem.

[028] De acordo com um recurso opcional da invenção, a unidade de filtro é disposta de modo a determinar o valor da propriedade de imagem vizinho combinado substancialmente como:

$$[029] \quad D_{\text{vizinho}} = \frac{\sum \alpha_j f_j g_j D_j}{\sum \alpha_j f_j g_j}$$

[030] onde  $f_j$  representa uma medida de diferença para um valor de intensidade de luz da segunda posição e um valor de intensidade de luz para um valor de propriedade de imagem da vizinhança na segunda posição;  $g_j$  representa uma medida de diferença para o valor de confiança  $C_i$  para o

primeiro valor de profundidade  $D_i$  e um valor de confiança para o valor de propriedade de imagem da vizinhança na segunda posição;  $D_j$  representa o valor de propriedade de imagem da vizinhança na segunda posição;  $\alpha_j$  representa um parâmetro de design, e as somas são sobre todas as posições  $j$  pertencentes à vizinhança.

[031] Isso pode fornecer uma operação particularmente vantajosa em muitas modalidades, e pode resultar especificamente na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[032] O parâmetro de design  $\alpha_j$  podem ser dependente de outros parâmetros, como, por exemplo, uma distância entre a primeira posição e a segunda posição.

[033] A determinação do peso é assimétrica em relação ao valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o valor de propriedade da primeira imagem.

[034] Isso pode fornecer uma operação particularmente vantajosa em muitas modalidades, e pode resultar especificamente na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[035] De acordo com um recurso opcional da invenção, a medida de diferença é limitada a um valor mínimo que é apenas ultrapassado se o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança exceder o limite do valor de confiança para o primeiro valor de propriedade de imagem.

[036] Isso pode fornecer uma operação particularmente vantajosa em muitas modalidades, e pode

resultar especificamente na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[037] O valor mínimo pode ser zero e/ou o limite pode ser zero.

[038] De acordo com um recurso opcional da invenção, a imagem de intensidade de luz, o mapa de confiança e o mapa da propriedade de imagem são parte de uma sequência temporal de imagens de intensidade de luz, mapas de confiança e mapas de propriedade da imagem, e a vizinhança tem uma extensão temporal.

[039] Isso pode proporcionar desempenho aprimorado em muitas modalidades e pode, em particular, fornecer muitas vezes melhor estabilidade e consistência temporais.

[040] De acordo com um recurso opcional da invenção, a unidade de filtro é disposta de modo a aumentar um valor de confiança para o valor de uma propriedade de imagem, se este for um valor de propriedade da imagem filtrado.

[041] Isso pode fornecer uma operação particularmente vantajosa em muitas modalidades, e pode resultar especificamente na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[042] O valor de propriedade de imagem em muitas modalidades pode ser considerado um valor de propriedade de uma imagem filtrado, se for parte de um mapa de propriedades de imagem filtrado gerado por uma filtragem anterior de um mapa de propriedades de imagem da sequência de mapas de propriedade de imagem.

[043] De acordo com um recurso opcional da invenção, a vizinhança é espacialmente assimétrica em relação à primeira posição.

[044] Isso pode proporcionar um desempenho aprimorado em muitas modalidades.

[045] De acordo com um recurso opcional da invenção, o mapa de propriedades de imagem é um mapa de indicação de profundidade que compreende valores determinados a partir de estimativa de disparidade em uma primeira imagem representativa de um primeiro ponto de vista e uma segunda imagem representativa de um segundo ponto de vista e que tem o primeiro ponto de vista como base para a estimativa de disparidade; e pela vizinhança se estender mais no sentido do segundo ponto de vista para o primeiro ponto de vista do que no sentido do primeiro ponto de vista para o segundo ponto de vista.

[046] Isso pode proporcionar um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades onde a estimativa de disparidade é usada. Isso pode refletir a assimetria na estimativa de disparidade subjacente entre as imagens correspondentes a diferentes pontos de vista.

[047] De acordo com um recurso opcional da invenção, a ponderação depende adicionalmente de uma diferença entre o primeiro ponto de vista e o segundo ponto de vista.

[048] Isso pode proporcionar um mapa de propriedade de uma imagem filtrado aprimorado em muitos cenários e modalidades.

[049] De acordo com um recurso opcional da invenção, o receptor é disposto adicionalmente para receber um

segundo mapa de propriedades de imagem e um segundo mapa de confiança, sendo que o segundo mapa de confiança compreende segundos valores de confiança para segundos valores de propriedade de imagem do segundo mapa de propriedades de imagem; e sendo que a unidade de filtro é adicionalmente disposta de modo a determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a um segundo valor de propriedade da imagem para a segunda posição e a um segundo valor de confiança para a segunda posição.

[050] Isso pode fornecer uma operação particularmente vantajosa em muitas modalidades, e pode resultar especificamente na geração de um mapa de propriedades de imagem filtrado aprimorado em muitos cenários e modalidades.

[051] De acordo com um recurso opcional da invenção, o mapa de propriedades de imagem é um dentre: um mapa de profundidade; um mapa de disparidade; um mapa de estimativa de movimento; e um mapa de transparência.

[052] A invenção pode, em muitas modalidades, possibilitar a geração de um mapa de profundidade; mapa de disparidade; mapa de estimativa de movimento; ou mapa de transparência aprimorado. Isso pode possibilitar um melhor processamento de imagens subsequente, por exemplo, da imagem de intensidade de luz e, assim, melhorar a qualidade da imagem e fornecer uma experiência de usuário aprimorada.

[053] De acordo com um aspecto da invenção, é fornecido um método de processamento de um mapa de propriedades de imagem, sendo que o método compreende: receber uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem, sendo que o mapa de propriedades de imagem compreende valores de propriedade de imagem para os

pixels da imagem de intensidade de luz, sendo que o mapa de propriedades de imagem compreende valores de propriedade da imagem para os pixels da imagem de intensidade de luz, e que o mapa de confiança compreende valores de confiança para os valores de propriedade de imagem; filtrar o mapa de propriedade da imagem em resposta à imagem de intensidade de luz e ao mapa de confiança, para gerar um mapa de propriedades de imagem filtrado, sendo que a filtragem compreende, para uma primeira posição no mapa de propriedade da imagem filtrado: determinar um valor de propriedade de imagem da vizinhança combinado em resposta a uma combinação ponderada de valores de propriedade de imagem vizinho que são valores de propriedades de imagem em uma vizinhança em torno da primeira posição, sendo que o peso para um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição é dependente de um valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e uma diferença entre um valor de intensidade de luz para a primeira posição e um valor de intensidade de luz para a segunda posição; e determinar um primeiro valor de propriedade de imagem filtrado para a primeira posição como uma combinação de um primeiro valor de propriedade de imagem no mapa de propriedades de imagem na primeira posição e o valor combinado de propriedade da imagem vizinho; sendo que a determinação do valor combinado de propriedade de imagem vizinho compreende determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a uma medida de diferença para o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e a um valor de confiança para o primeiro valor de propriedade de imagem, sendo que a determinação do peso é assimétrica em relação ao valor de

confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o valor de propriedade da primeira imagem, com o peso dependendo do sinal da medida de diferença.

[054] Esses e outros aspectos, recursos e vantagens da invenção ficarão evidentes e serão elucidados com referência à modalidade descrita (ou modalidades descritas) a seguir.

#### Breve descrição dos desenhos

[055] As modalidades da invenção serão descritas, apenas a título de exemplo, com referência aos desenhos, nos quais:

[056] a Figura 1 ilustra um exemplo de elementos de um aparelho para processar um mapa de propriedades de imagem, como um mapa de profundidade, de acordo com algumas modalidades da invenção;

[057] a Figura 2 ilustra um exemplo de um núcleo usado pelo aparelho da Figura 1.

#### Descrição detalhada das modalidades

[058] A descrição a seguir enfoca as modalidades da invenção aplicáveis ao processamento de um mapa de propriedades de imagem sob a forma de um mapa de profundidade. Entretanto, deve-se considerar que a invenção não se limita a essa aplicação, mas ela pode ser aplicada a muitos outros tipos de mapa de propriedades de imagem.

[059] A Figura 1 ilustra um exemplo de um aparelho para o processamento de um mapa de propriedades de imagem, o qual, no exemplo específico, é um mapa de profundidade.

[060] O aparelho compreende um receptor 101 que é disposto de modo a receber dados de imagem que compreendem uma imagem de intensidade de luz e um mapa de confiança associado e mapa de propriedades de imagem, o qual pode ser especificamente um mapa de profundidade.

[061] A imagem de intensidade de luz pode compreender valores que refletem um valor de intensidade de luz para os pixels individuais de uma imagem a ser renderizada/exibida por uma tela.

[062] Deve-se compreender que os valores de intensidade de luz podem ser qualquer valor indicativo de uma intensidade de luz, e podem ser especificamente uma intensidade de luz para, por exemplo, um canal de cor. Por exemplo, um valor de intensidade de luz pode ser um valor R, G ou B de uma representação RGB, ou pode ser, por exemplo, um valor Y de uma representação Yuv, ou pode ser, de fato, um valor u ou v de tal representação Yuv. É observado que os valores u e v podem ser valores de croma e que estes também fornecem informações relacionadas à intensidade de luz de canais de cor individuais durante a renderização. Dessa forma, um valor de intensidade de luz pode ser uma luminância, crominância ou, de fato, valor de croma de uma representação de cor. Tipicamente, a imagem de intensidade de luz compreende valores de pixel de intensidade de luz indicativos de uma propriedade de crominância para os pixels correspondentes. Uma propriedade de crominância para um pixel pode ser indicativa de uma propriedade de luminância, uma propriedade de saturação, ou ambas as propriedades de saturação e luminância para o pixel.

[063] O mapa de propriedades de imagem compreende valores de propriedade de imagem para os



pixels/posições da imagem de intensidade de luz. No exemplo específico, o mapa de propriedades de imagem é um mapa de profundidade que compreende informações sobre as profundidades dos pixels na imagem de intensidade de luz. O mapa de profundidade pode compreender especificamente valores de profundidade que indicam diretamente uma profundidade para os valores de intensidade de luz da imagem de intensidade de luz, como uma coordenada cartesiana.

[064] Deve-se compreender que informação de disparidade é uma forma de informação de profundidade e que o mapa de profundidade é, de forma equivalente, um mapa de disparidade. Em algumas modalidades, o mapa de profundidade pode compreender um valor de disparidade para os valores de intensidade de luz da imagem de intensidade de luz. Deve-se compreender que as referências à profundidade na seguinte descrição são igualmente aplicáveis às modalidades que usam valores de profundidade/z diretos e para modalidades que usam valores de disparidade.

[065] Embora a descrição a seguir enfocará modalidades em que a propriedade de imagem é uma propriedade de profundidade, a abordagem pode ser aplicada a outras propriedades de imagem e, dessa forma, em outras modalidades o mapa de propriedades de imagem pode ser um mapa indicativo de uma propriedade de imagem diferente. Especificamente, o mapa de propriedades de imagem pode ser um mapa de transparência que compreende valores de dados indicativos de uma transparência dos pixels da imagem de intensidade de luz. Esse mapa de transparência é também conhecido tipicamente na técnica como um mapa  $\alpha$ .

[066] O mapa de confiança compreende valores de confiança para os valores de propriedade da imagem, ou seja, para os valores de profundidade ou transparência do mapa de propriedades de imagem. Deve-se compreender que o conceito de valores de confiança para refletir uma confiabilidade ou a precisão de um parâmetro é bem conhecido no campo de processamento de imagens. O valor de confiança para um valor de propriedade de imagem pode indicar uma probabilidade (estimada) que um valor de propriedade de imagem tem o valor correto. Os valores de confiança são gerados como o resultado de muitos processos de estimativa que procuram estimar ou determinar um valor de parâmetro. Por exemplo, uma estimativa de profundidade tipicamente não gera apenas uma estimativa de profundidade que representa a profundidade estimada, mas também gera um valor de confiança que indica a confiança na estimativa de profundidade refletindo a real profundidade do objeto de imagem sendo estimado.

[067] Em muitas modalidades, o mapa de propriedades de imagem pode compreender um valor individual de propriedade de imagem para cada valor de intensidade de luz da imagem de intensidade de luz. Entretanto, em muitas modalidades, um valor de propriedade de imagem pode, por exemplo, ser comum para uma pluralidade de valores de intensidade de luz. Por exemplo, em muitas modalidades, o mapa de propriedades de imagem pode ser de uma resolução mais baixa que a imagem de intensidade de luz e, de fato, o processo descrito a seguir pode ser usado para aumentar artificialmente a taxa de amostragem do mapa de propriedades de imagem.

[068] De modo similar, o mapa de confiança pode ter uma resolução diferente da imagem de intensidade de luz e,

de fato, pode também, em algumas modalidades, ter uma resolução diferente do mapa de propriedades de imagem. Em algumas modalidades, um valor do mapa de confiança pode ser comum a uma pluralidade de valores de propriedade de imagem do mapa de propriedades de imagem e, dessa forma, pode ter uma resolução mais baixa. Entretanto, na maioria das modalidades, o mapa de confiança terá a mesma resolução que o valor de propriedade de imagem e compreenderá um valor de confiança para cada valor de propriedade de imagem.

[069] Em algumas modalidades, o receptor 101 pode receber um único conjunto independente de (pelo menos) uma imagem de intensidade de luz e um mapa de propriedades de imagem e um mapa de confiança associados. Um conjunto de uma imagem de intensidade de luz e dados associados incluindo o mapa de propriedades de imagem e o mapa de confiança (bem como, possivelmente, outros mapas de propriedade de imagem ou imagens de intensidade de luz associados, por exemplo, para outros pontos de vista) será chamado de uma imagem composta. Um imagem composta pode, dessa forma, ser considerada ser o conjunto de todos os dados de imagem relacionado a um instante no tempo específico.

[070] Em algumas modalidades, a imagem composta pode ser uma parte de uma sequência temporal de imagens compostas. Dessa forma, a imagem de intensidade de luz pode ser parte de uma sequência temporal de imagens de intensidade de luz, e cada imagem de intensidade de luz pode ter um ou mais mapas de propriedade de imagem e mapas de confiança associados. Consequentemente, o mapa de propriedades de imagem e o mapa de confiança para a imagem de intensidade de luz atual

podem também fazer parte de uma sequência temporal de mapas de propriedade de imagem e mapas de confiança.

[071] Especificamente, a imagem composta pode fazer parte de uma sequência de imagens de vídeo, sendo que, tipicamente, cada imagem tem um mapa de propriedades de imagem e um mapa de confiança associados.

[072] O receptor 101 é acoplado a uma unidade de filtro 103 que é disposta de modo a filtrar o mapa de propriedades de imagem em dependência tanto da imagem de intensidade de luz e do mapa de confiança. Conforme será descrito com mais detalhes abaixo, a unidade de filtro 103, dessa forma, realiza uma filtragem do mapa de propriedades de imagem que leva em conta ambos os valores de intensidade de luz e valores de confiança para o mapa de propriedades de imagem. A filtragem, conseqüentemente, ao mesmo tempo, leva em conta duas diferentes modalidades em adição ao próprio valor de propriedade de imagem.

[073] A unidade de filtro 103 é, no exemplo específico da Figura 1, acoplado a um transmissor 105 que é disposto de modo a gerar um sinal de saída adequado a partir da imagem de intensidade de luz filtrado e transmiti-lo a uma fonte adequada. Por exemplo, o transmissor 105 pode ser disposto de modo a aplicar um processo de codificação de imagem ou vídeo à imagem de intensidade de luz para gerar um sinal adequado para distribuição. Por exemplo, o transmissor pode aplicar uma codificação MPEG à imagem (ou imagens) de intensidade de luz filtrado, possibilitando que o fluxo de bits resultante seja distribuído. Um valor de propriedade de imagem e possivelmente o mapa de confiança pode ser incluído no fluxo de bits de saída em muitas modalidades. Em outros

exemplos, o transmissor 105 pode ser disposto de modo a gerar um sinal de acionamento de exibição adequado que pode ser alimentado diretamente a uma tela para acionar a renderização e apresentar a imagem (ou imagens) de intensidade de luz com base no mapa de propriedades de imagem filtrado.

[074] Um aprimoramento de um mapa de propriedades de imagem com base em uma outra propriedade, como uma propriedade de intensidade de luz ou uma propriedade de confiança, pode melhorar a qualidade do mapa de propriedades de imagem. No entanto, as abordagens convencionais são tipicamente projetadas para considerar apenas um parâmetro, e a combinação de diferentes modalidades e propriedades é um problema complexo e desafiador. Pode não apenas exigir um aumento significativo da demanda de recursos e complexidade, mas também pode ter efeitos imprevisíveis e indesejáveis (por exemplo, os diferentes processos para as diferentes modalidades podem, tipicamente, interferir uns com os outros de maneiras possivelmente imprevisíveis).

[075] No sistema da Figura 1, a unidade de filtro 103 é disposta de modo a executar um processamento combinado, integrado e simultâneo do mapa de propriedades de imagem com base em ambos os valores de intensidade de luz e valores de confiança. Em particular, os inventores não só perceberam que não é apenas possível mas também desejável processar o mapa de propriedades de imagem com base nas diferentes propriedades, mas também que ao invés de realizar um processo sequencial como seria considerado necessário ou adequado, é possível realizar uma filtragem integrada e combinada que pode combinar, simultaneamente, informações fornecidas por diferentes propriedades e modalidades.

[076] Esta operação será descrita a seguir com consideração específica a um mapa de profundidade sendo filtrado.

[077] A unidade de filtro 103 compreende um combinador 107 que é especificamente disposto de modo a determinar valores de saída filtrados do mapa de propriedade da imagem. No exemplo específico, a resolução do mapa de propriedades de imagem de entrada e do mapa de propriedades de imagem de saída resultante filtrado será da mesma resolução. Consequentemente, a posição de um valor de propriedade da imagem filtrado tem uma correspondência direta com um valor de propriedade de imagem do mapa de propriedades de imagem de entrada, ou, em outras palavras, o mapa de propriedades de imagem de entrada e o mapa de propriedades de imagem de saída filtrado têm pixels diretamente correspondentes. A posição no mapa de propriedades de imagem filtrado de saída para os quais o valor atual é determinado irá, a seguir, ser chamada de uma primeira posição e usará o índice  $i$ . Deve-se compreender que essa referência e índice pode igualmente se relacionar com a posição correspondente no mapa de propriedades de imagem de entrada ou, de fato, às posições correspondentes no mapa de confiança ou à imagem de intensidade de luz. De modo similar, uma posição em uma vizinhança da primeira posição  $i$  será chamada de uma segunda posição  $j$ . Deve-se compreender que uma posição no mapa de propriedades de imagem de saída filtrado tem posições correspondentes no mapa de propriedades de imagem de entrada, no mapa de confiança, e na imagem de intensidade de luz. De fato, deve-se notar que é prática padrão na técnica que diferentes imagens e mapas de um compósito/imagem combinada

ter posições correspondentes em todos os mapas/imagens (mesmo que estas estejam em diferentes resoluções).

[078] Um primeiro valor de propriedades de imagem de entrada e um primeiro valor de propriedades de imagem de saída filtrado na primeira posição podem ser chamados de índice  $i$ . Para o mapa de propriedades de imagem ser um mapa de profundidade, o índice  $i$  denota, assim, um primeiro valor de profundidade de entrada  $D_i$  e um primeiro valor de profundidade de saída filtrado  $D_{i,saída}$ .

[079] Deve-se compreender ainda que as posições geralmente se correlacionam com pixels e, dessa forma, a posição dos termos e pixel será usada de forma intercambiável conforme for adequado. Especificamente, um pixel ou valor com índice  $i$  representa o valor correspondente na primeira posição, e um pixel ou valor com índice  $j$  representa o valor correspondente na segunda posição.

[080] Para este primeiro valor de profundidade filtrado  $D_{i,saída}$ , (e, portanto, para a primeira posição) o combinador 107 gera um valor de saída  $D_{i,saída}$ , que depende do valor de entrada  $D_i$  e um valor de propriedade de profundidade/imagem que é determinado com base nos valores de propriedade de profundidade/imagem em uma vizinhança em torno da posição/do pixel  $i$ , ou seja, em torno da posição do valor de saída de  $D_{i,saída}$ , e do primeiro valor de profundidade  $D_i$ . Este valor determinado em uma região em torno da posição  $i$  é chamado de um valor de profundidade/propriedade de imagem vizinho  $D_{i,vizinhos}$ .

[081] A unidade de filtro 103 é, consequentemente, disposta de modo a determinar um primeiro valor de propriedade de imagem filtrado para o valor/pixel/

posição/índice  $i$  como uma combinação do primeiro valor de profundidade de entrada  $D_i$  e o valor de profundidade vizinho combinado  $D_{i,vizinhos}$ . O combinador 107 pode, conseqüentemente, determinar o valor de saída filtrado como:

$$[082] \quad D_{i,saída} = f(D_{i,entrada}, D_{i,vizinhos})$$

[083] onde  $f()$  é uma combinação ponderada.

[084] Em algumas modalidades, uma simples combinação predeterminada pode, por exemplo, ser usada, como:

$$[085] \quad D_{i,saída} = \alpha D_i + (1 - \alpha) D_{i,vizinhos}$$

[086] onde  $\alpha$  é uma constante predeterminada entre 0 e 1.

[087] O valor de profundidade vizinho combinado  $D_{i,vizinhos}$  é determinado por um processador 109 vizinho que faz parte da unidade de filtro 103 e é acoplado ao combinador 107.

[088] O processador 109 vizinho é, para cada valor de profundidade  $i$  no mapa de profundidade de saída, disposto de modo a determinar um valor de profundidade vizinho combinado  $D_{i,vizinhos}$  como uma combinação ponderada dos valores de profundidade em uma dada vizinhança. Tal combinação ponderada pode ser considerada corresponder a uma filtragem, a qual pode ser especificamente uma filtragem espacial, quando apenas uma vizinhança espacial é considerada, uma filtragem temporal, quando apenas uma vizinhança temporal é considerada, ou tanto uma filtragem espacial como temporal, quando ambas as vizinhanças espacial e temporal são consideradas. Deve-se compreender também que o termo "núcleo" é muitas vezes usado no campo para dita uma vizinhança que forma a base de filtragem, e os termos "vizinhança" e "núcleo" serão ambos usados, conseqüentemente, para se referir à vizinhança conforme adequado ao contexto.



[089] Dessa forma, no sistema, um núcleo ou vizinhança é definido e o processador 109 vizinho pode gerar um valor de profundidade vizinho combinado  $D_{i,vizinhos}$  mediante a combinação dos valores de profundidade que pertencem ao núcleo/vizinhança em torno da posição do valor de profundidade para a qual o valor de profundidade filtrado é determinado. O valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode, conseqüentemente, ser visto como uma média ponderada dos valores de profundidade no núcleo/vizinhança.

[090] O valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode, conseqüentemente, ser determinado como:

$$[091] \quad D_{i,vizinhos} = C_{j \in K_i}(w_j, D_j)$$

[092] onde  $K_i$  representa o núcleo/a vizinhança para a posição no mapa de confiança e imagem de intensidade de luz correspondente a pixel/posição  $i$  (ou seja, a primeira posição) nos mapas de profundidade, e  $j$  é o índice usado para a posição de pixel vizinho (ou seja, a segunda posição).  $C_{j \in K_i}$  denota a combinação ponderada sobre (tipicamente todos) os índices  $j$  que estão no núcleo/vizinhança  $K$  para pixel/posição  $i$  com os pesos para um dado valor de profundidade vizinho  $D_j$  sendo dados por  $w_j$ .

[093] Especificamente, em muitas modalidades, o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser dado por uma soma ponderada dos valores de profundidade vizinhos  $D_j$ :

$$[094] \quad D_{i,vizinhos} = \sum_{j \in K_i} w_j D_j.$$

[095] Dessa forma, o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  é determinado por um cálculo da média ponderada dos valores de profundidade vizinho  $D_j$  na

vizinhança em torno da posição do primeiro valor de profundidade  $D_i$ .

[096] Na abordagem, os pesos para essa combinação são determinados com base tanto na imagem de intensidade de luz como no mapa de confiança. Especificamente, para um primeiro valor de profundidade vizinho  $D_j$ , o peso  $w_j$  na combinação é dependente tanto do valor de confiança  $C_j$  para a posição  $j$  como de uma diferença entre o valor de intensidade de luz  $I_j$  para a posição  $j$  e o valor de intensidade de luz  $I_i$  para a posição  $i$ , ou seja, o peso  $w_j$  é dependente do valor de confiança para o primeiro valor de profundidade vizinho  $D_j$  e da diferença entre o valor de intensidade de luz  $I_j$  para o primeiro valor de profundidade vizinho  $D_j$  e o valor de intensidade de luz  $I_i$  para o primeiro valor de profundidade  $D_i$ . Em particular, o peso  $w_j$  na combinação é dependente do valor de confiança  $C_j$  para a posição  $j$  em relação ao valor de confiança para o pixel/a posição  $i$  para o qual o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  é determinado, ou seja, é determinado com base no valor do valor de confiança  $C_j$  em relação ao valor de confiança  $C_i$ . O peso  $w_j$  é determinado com base em uma medida da diferença entre o valor de confiança para o valor de propriedade de imagem da vizinhança  $C_j$  e o valor de confiança para o valor ou posição "âncora"  $C_i$ . O peso pode ser determinado especificamente como

$$[097] \quad w_j = f(d_c(C_j, C_i); d_I(I_j, I_i))$$

[098] onde  $f$  é uma função adequada dependente da modalidade específica,  $d_c(C_j, C_i)$  representa uma medida de diferença adequada para os valores de confiança na posição  $i$  e  $j$  respectivamente, e  $d_I(I_j, I_i)$  representa uma medida de diferença adequada para os valores de intensidade da luz na

posição  $i$  e  $j$ , respectivamente. As medidas de diferença adequadas para os valores de intensidade pode, por exemplo, ser uma norma adequada  $||\mathbf{I}_i; \mathbf{I}_j||$  baseada nas intensidades de luz e, em particular, no vetor de diferença de intensidade de luz  $\mathbf{I}_i - \mathbf{I}_j$ , ou seja,  $||\mathbf{I}_i - \mathbf{I}_j||$ . De modo similar, as medidas de diferença adequadas para os valores de intensidade podem, por exemplo, também ser uma norma adequada. Em muitas modalidades, um valor de diferença direta obtido por subtração simples pode ser usado, por exemplo,  $d_c(C_j; C_i) = C_j - C_i$ .

[099] O peso para um dado valor de profundidade vizinho  $D_j$  aumentará tipicamente para um valor de confiança crescente  $C_j$  em relação ao valor de confiança de referência  $C_i$ , ou seja, o valor de profundidade vizinho mais confiável  $D_j$  é considerado ser relativo à confiabilidade do valor de profundidade original para o pixel sendo filtrado, sendo que o maior valor de profundidade  $D_j$  é ponderado na combinação. Especificamente, o peso para o valor de profundidade vizinho  $D_j$  é tipicamente uma função monotonicamente crescente do valor de confiança  $C_j$  para o valor de profundidade vizinho  $D_j$  e é tipicamente uma função monotonicamente decrescente do valor de confiança valor  $C_i$  para o valor de profundidade da posição  $i$ . Em muitas modalidades, o peso é uma função monotonicamente crescente da medida de diferença para os valores de intensidade de confiança na posição  $i$  e  $j$ , respectivamente.

[100] O peso para um dado valor de profundidade vizinho  $D_j$  tipicamente diminuirá para um valor de diferença crescente, ou seja, quanto maior a diferença entre o valor de intensidade de luz  $I_j$  para a posição  $j$  e o valor de intensidade de luz  $I_i$  para a posição  $i$ , menor será a ponderada na combinação. Especificamente, o peso  $w_j$  para um valor de

profundidade vizinho  $D_j$  é tipicamente uma função monotonicamente decrescente do valor/medida de diferença indicativa da diferença entre o valor de intensidade de luz  $I_j$  para a posição  $j$  e o valor de intensidade de luz  $I_i$  para a posição  $i$ . Especificamente, a função  $f$  pode ser uma função monotonicamente decrescente de uma norma  $||I_i; I_j||$ , e especificamente  $||I_i - I_j||$ .

[101] A abordagem, dessa forma, determina um valor de profundidade vizinho combinado  $D_{i,vizinhos}$  que tem maior contribuição dos valores de profundidade vizinhos  $D_j$  para o qual o valor de confiança é alto em relação ao valor de confiança para o valor de profundidade original para a posição atual, e a diferença na intensidade de luz daquela do pixel atual é baixa. A abordagem, dessa forma, possibilita um ajuste eficiente, simultâneo e conjunto, e a variação das contribuições feitas por valores individuais com esta alteração é dependente de diferentes propriedades e modalidades.

[102] A unidade de filtro 103 pode ser disposta para prosseguir com a realização desta operação de filtragem para todas as posições/pixels do mapa de profundidade para gerar um mapa de profundidade filtrado.

[103] A abordagem pode proporcionar uma experiência de usuário aprimorada em muitas situações. Em particular, ela fornece uma filtragem transversal com base em múltiplas propriedades e modalidades. Adicionalmente, descobriu-se na prática que essa abordagem fornece excelentes resultados e possibilita não apenas um processo computacionalmente eficiente, mas também mitiga a interferência e reduz a interdependência entre as operações

de filtragem baseadas em diferentes modalidades. Descobriu-se que essa abordagem fornece substancialmente melhor desempenho em comparação com as abordagens mais simples e mais intuitivas de execução de processamento sequencial e em cascata, e filtragem bilateral.

[104] Para muitos cenários práticos, descobriu-se que uma imagem mais precisa e confiável do mapa de propriedades, e especificamente do mapa de profundidade, pode ser gerada, fornecendo um melhor suporte para, por exemplo, subsequente processamento de imagens da imagem de intensidade de luz. Por exemplo, a comutação do ponto de vista da imagem melhorada pode ser encontrada em muitos cenários práticos, levando a uma melhor qualidade de imagem e uma experiência de usuário aprimorada.

[105] Conforme mencionado, o combinador 107 é disposto de modo a gerar o valor de profundidade de saída filtrado  $D_{i,saída}$  mediante a combinação do primeiro valor de profundidade  $D_i$  e o valor de profundidade vizinho combinado  $D_{i,vizinhos}$ . Em algumas modalidades, esta combinação pode ser uma combinação ponderada com pesos predeterminados, como, por exemplo, a determinação do valor de profundidade de saída filtrado  $D_{i,saída}$  como a média do primeiro valor de profundidade  $D_i$  e o valor de profundidade vizinho combinado  $D_{i,vizinhos}$ .

[106] Entretanto, em muitas modalidades, o peso para o primeiro valor de profundidade  $D_i$  em relação ao valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser determinado em resposta ao valor de confiança para o primeiro valor de profundidade  $D_i$ , chamado de primeiro valor de confiança  $C_i$ . O peso para o primeiro valor de profundidade  $D_i$  em relação ao valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode

tipicamente ser uma função monotonicamente crescente do primeiro valor de confiança  $C_i$ . Dessa forma, quanto mais confiança houver no valor de profundidade existente, menor será a possível modificação com base nos valores de profundidade na vizinhança.

[107] Deve-se compreender que a ponderação relativa entre o primeiro valor de profundidade  $D_i$  e o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser alterada mediante a alteração de um peso do primeiro valor de profundidade  $D_i$ , um peso do valor de profundidade vizinho  $D_{i,vizinhos}$  ou, de fato, tanto os pesos do primeiro valor de profundidade  $D_i$  como do valor de profundidade vizinho combinado  $D_{i,vizinhos}$ .

[108] Em muitas modalidades, o valor de profundidade de saída filtrado  $D_{i,saída}$  pode ser especificamente determinado como:

$$[109] \quad D_{i,saída} = C_i D_i + (1 - C_i) D_{i,vizinhos}$$

[110] Dessa forma, em tal exemplo, no caso em que, por exemplo, uma estimativa de disparidade usada para gerar o primeiro valor de profundidade  $D_i$  é considerada confiante ( $C_i=1$ ), então os pixels vizinhos não contribuem para valor de profundidade de saída filtrado  $D_{i,saída}$ . No outro extremo, onde não há confiança no valor atual ( $C_i=0$ ), o valor de profundidade de saída filtrado  $D_i$ , tem por base apenas os pixels vizinhos.

[111] Em algumas modalidades, o processador 109 de vizinhança pode ser disposto ainda de modo a determinar um valor de confiança vizinho combinado para o valor de profundidade vizinho combinado  $D_{i,vizinhos}$ . Este valor de confiança vizinho pode ser determinado com base em valores de

confiança para os valores de profundidade vizinhos  $D_j$  (ou seja, os valores de confiança  $C_j$  dos pixels vizinhos). Por exemplo, a média dos valores de confiança vizinhos pode ser determinada ou, por exemplo, a média ponderada onde os pesos  $w_j$  são levados em consideração pode ser determinada.

[112] Em tal caso, a ponderação do primeiro valor de profundidade  $D_i$  em relação ao valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser determinado com base no nível de confiança para o primeiro valor de profundidade  $D_i$  em relação ao valor de confiança vizinho.

[113] Por exemplo, o valor de profundidade de saída filtrado  $D_{i,saida}$  pode ser determinado como:

$$[114] \quad D_{i,saida} = \frac{C_i}{C_i + C_{i,vizinho}} D_i + \frac{C_{i,vizinho}}{C_i + C_{i,vizinho}} D_{i,vizinhos}$$

[115] Onde  $C_{i,vizinho}$  representa o valor de confiança vizinho.

[116] Deve-se compreender que diferentes abordagens e funções podem ser usadas em diferentes modalidades para executar a combinação para determinar o valor de profundidade vizinho combinado  $D_{i,vizinhos}$ .

[117] Em algumas modalidades, o peso  $w_j$  para um valor de profundidade vizinho  $D_j$  é determinado ser proporcional a um valor  $g_j$  que é dependente do valor de confiança para o valor de profundidade vizinho  $D_j$  mas não aos valores de intensidade de luz. Adicionalmente, o peso  $w_j$  pode ser proporcional a um valor  $f_j$  que é dependente da medida de diferença para os valores de intensidade da luz mas não é dependente do valor de confiança. Dessa forma, em algumas modalidades, as medidas separadas podem ser determinadas com base no mapa de confiança e na imagem de intensidade de luz,

e estas medidas podem ser combinadas mediante a multiplicação ao se determinar o peso  $w_j$ .

[118] Em algumas modalidades, o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  para um dado valor de profundidade  $D_i$  pode ser determinado como

$$[119] \quad D_{vizinhos} = \frac{\sum \alpha_j f_j g_j D_j}{\sum \alpha_j f_j g_j}$$

[120] onde  $f_j$  representa uma medida de diferença para o valor de intensidade de luz  $I_i$  para o primeiro valor de profundidade  $D_i$  e o valor de intensidade de luz  $I_j$  para um valor de profundidade vizinho  $D_j$ ;  $g_j$  representa uma medida de diferença de confiança para o valor de confiança  $C_i$  para o primeiro valor de profundidade  $D_i$  e um valor de confiança  $C_j$  para um valor de profundidade vizinho  $D_j$ ; e  $\alpha_j$  representa um parâmetro de design. As somas são sobre os pixels/valores da vizinhança, ou seja,  $j \in K_i$  onde  $K_i$  é o núcleo/vizinhança para a posição  $i$ .

[121] Dessa forma, em algumas modalidades, o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser determinado como:

$$[122] \quad D_{i,vizinhos} = \sum w_j \cdot D_j$$

[123] Onde

$$[124] \quad w_j = \frac{\alpha_j f_j g_j}{\sum \alpha_j f_j g_j}.$$

[125] Em algumas modalidades,  $\alpha_j$  pode ser simplesmente definida como 1, ou seja:

$$[126] \quad w_j = \frac{f_j g_j}{\sum f_j g_j}.$$



[127] onde  $g_j$  depende de uma diferença entre  $C_j$  e  $C_i$ , e  $f_j$  tipicamente depende de uma norma para o vetor de diferença de cor:  $\mathbf{I}_i - \mathbf{I}_j$ .

[128] Deve-se considerar que abordagens diferentes para detectar as medidas  $f_j$  e  $g_j$  podem ser usadas em modalidades diferentes.

[129] Por exemplo, uma abordagem particularmente vantajosa para determinar o valor de  $f_j$  pode, em muitas modalidades, ser dada por:

$$[130] \quad f_j \equiv e^{-\gamma \|\mathbf{I}_i - \mathbf{I}_j\|}$$

[131] onde  $\|\mathbf{I}_i - \mathbf{I}_j\|$  é uma norma adequada para o vetor de diferença de cor:  $\mathbf{I}_i - \mathbf{I}_j$  e  $\gamma$  é um parâmetro de design que pode ser usado para controlar a importância relativa da diferença de intensidade de luz em relação ao valor de confiança.

[132] Uma abordagem particularmente vantajosa para determinar o valor de  $g_j$  pode, em muitas modalidades, ser:

$$[133] \quad g_j \equiv e^{\beta(C_j - C_i)}$$

[134] onde  $\beta$  pode ser um parâmetro de design que pode ser usado para controlar a importância relativa do valor de confiança em relação à diferença de intensidade de luz.

[135] O peso  $w_j$  é, consequentemente, não apenas dependente do valor de confiança  $C_j$  mas também do valor de confiança  $C_i$  para o primeiro valor de profundidade  $D_i$ . Especificamente, o peso  $w_j$  é, em muitas modalidades, vantajosamente dependente de uma medida de diferença entre o valor de confiança  $C_j$  e o valor de confiança  $C_i$ . Isso pode, por exemplo, ser alcançado, tornando-se o valor  $g_j$  dependente tanto do valor de confiança  $C_j$  como do valor de confiança  $C_i$ , e especificamente da diferença entre eles, por exemplo,

$$[136] \quad g_j = f(C_j - C_i).$$

[137] A função  $f(C_j - C_i)$  pode especificamente ser uma função linear que fornece um peso proporcionalmente maior aos valores de profundidade vizinhos  $D_j$  quanto mais o valor de confiança  $C_j$  estiver acima do valor de confiança  $C_i$ . Dessa forma, a abordagem aumentará o enfoque nos valores de profundidade vizinhos  $D_j$  que têm maior confiança que o primeiro valor de profundidade  $D_i$ . Dessa forma, em muitas modalidades, o peso  $w_j$  pode ser determinado como um peso assimétrico que é assimétrico em relação ao valor de confiança  $C_j$  e o valor de confiança  $C_i$ . Não apenas especificamente a diferença absoluta entre estes, mas também o sinal é importante e afeta o peso  $w_j$  gerado.

[138] Em muitas modalidades, o peso  $w_j$  pode ser dependente de uma medida de diferença indicativa de uma diferença entre o valor de confiança  $C_j$  e o valor de confiança  $C_i$  sendo que a medida de diferença é limitada a um valor mínimo que é apenas ultrapassado se o valor de confiança  $C_j$  exceder o limite do valor de confiança  $C_i$ . O valor mínimo e o limite podem ser especificamente zero, e a medida de diferença pode, por exemplo, ser determinada como:

$$[139] \quad \max(C_j - C_i, 0).$$

[140] Em tais modalidades, o peso  $w_j$  pode ser baseado em uma contribuição do valor de confiança  $C_j$  para o valor de profundidade vizinho  $D_j$  dado por:

$$[141] \quad g_j = f(\max(C_j - C_i, 0))$$

[142] onde  $f()$  é uma função adequada, como, por exemplo,

$$[143] \quad g_j = e^{-\beta(1 - \max(C_j - C_i, 0))}.$$

[144] Em algumas modalidades, o peso  $w_j$  pode também ser dependente de uma distância entre a posição  $i$  do primeiro valor de profundidade filtrado  $D_{i,saída}$  e a posição do valor de profundidade vizinho  $D_j$ .

[145] Por exemplo, o valor  $\alpha_j$  pode se tornar dependente da distância entre as posições correspondentes para os índices  $i$  e  $j$ . Deve-se compreender que qualquer medida de distância adequada pode ser usada, como, especificamente, uma medida de distância com base em uma norma adequada. Como um exemplo específico,  $\alpha_j$  pode refletir uma função de distância, como:

$$[146] \quad \alpha_j \equiv e^{-\psi((x_j-x_i)^2+(y_j-y_i)^2)}.$$

[147] onde  $\psi$  é um parâmetro de design e  $x, y$  refletem as posições Euclidianas que correspondem aos índices  $i$  e  $j$ .

[148] Conforme anteriormente mencionado, a imagem de intensidade de luz, o mapa de confiança e o mapa de propriedades de imagem podem fazer parte de um mapa de sequência temporal de imagens de intensidade de luz, mapas de confiança e mapas de propriedades de imagem, ou seja, o receptor 101 pode receber uma sequência de imagens combinadas. Especificamente, o receptor 101 pode receber um sinal de vídeo que compreende a sequência de imagens combinadas, e, especificamente, cada imagem de intensidade de luz pode ser um quadro de uma sequência de vídeo.

[149] Em tais exemplos, a vizinhança, ou seja, o núcleo, pode ter uma extensão temporal, ou seja, a vizinhança pode incluir valores de profundidade a partir de mapas de profundidade tipicamente para quadros/instantes no tempo anteriores ao quadro/instante de tempo atual (embora também

seja possível que a vizinhança inclua valores de profundidade que pertencem a quadros/instantes no tempo subsequentes).

[150] Em tais modalidades, a extensão espacial da vizinhança pode ser diferente para os diferentes quadros/instantes de tempo. Tipicamente, a extensão espacial da vizinhança é maior para a imagem de intensidade de luz e para o mapa de confiança fornecidos para o mesmo instante no tempo como o mapa de propriedades de imagem que está sendo filtrado do que para outros instantes no tempo.

[151] A extensão temporal pode proporcionar desempenho aprimorado e pode, em particular, fornecer consistência espacial aprimorada e ruído temporal percebido reduzido.

[152] Em algumas modalidades, o mapa de propriedades de imagem para um outro instante de tempo pode ser um que já tenha sido filtrado. Por exemplo, o aparelho pode ser disposto de modo a filtrar o mapa de propriedades de imagem atual com base em uma vizinhança/um núcleo que inclui também o mapa de propriedades de imagem para o quadro/instante no tempo anterior. No entanto, como isto já foi processado pela unidade de filtro 103, pode ser, consequentemente, um mapa de propriedades de imagem filtrado. Essa abordagem pode fornecer melhor desempenho na medida em que os valores de propriedade de imagem do mapa de propriedades de imagem filtrado podem ser mais confiáveis devido à filtragem.

[153] A unidade de filtro 103 pode ser disposta ainda de modo modificar os valores de confiança para refletir tal confiabilidade aprimorada. Especificamente, a unidade de filtro 103 pode ser disposta de modo a aumentar um valor de confiança para um valor de propriedade de imagem se este for

um valor de propriedade de imagem filtrado. O aumento pode, em algumas modalidades, ser um aumento relativo, como adicionar um deslocamento específico (por exemplo, aumentar o valor de confiança em 0,5 - com o resultado sendo limitado a um máximo de 1,0), ou pode ser uma configuração absoluta, como definir o valor em 1,0.

[154] Esse tipo de abordagem resulta em uma maior ponderação dos valores de propriedade de imagem considerados mais confiáveis devido à uma filtragem anterior e, conseqüentemente, pode proporcionar melhor filtragem do mapa de propriedades de imagem atual.

[155] Um exemplo de uma vizinhança/núcleo 201 que pode ser adequado para muitas modalidades é ilustrado na Figura 2.

[156] No exemplo específico, o núcleo 201 para um determinado valor de profundidade  $D_i$  compreende um grupo 205 de 20 posições de pixel mais próximas da posição 203 do primeiro valor de profundidade  $D_i$ . Este grupo 305 de posições de pixel reflete as posições de pixel espaciais que estão compreendidas no núcleo 201 para o mapa de profundidade atual e o mapa de profundidade anterior, ou seja, ele reflete um grupo espacial de pixels para o qual o núcleo 201 também tem uma extensão temporal.

[157] A parte restante do núcleo define 49 posições de pixel que só são incluídas no núcleo 201 para o instante no tempo/quadro atual. Dessa forma, para estas 49 posições de pixel, o núcleo 201 não tem extensão temporal.

[158] Deve-se também notar que, no exemplo, os valores de profundidade são subamostrados, ou seja, o núcleo 201 não inclui todos os valores de profundidade dentro da

região do núcleo, mas define apenas um subconjunto desses. No exemplo específico, a subamostragem podem ser por um fator de 3, e dessa forma apenas cada terceiro valor de profundidade é considerado para a filtragem/o núcleo 201.

[159] Consequentemente, embora o núcleo 201 consista de 69 posições espaciais, ele abrange uma região ou uma área correspondente a 40 por 13 pixels no mapa de profundidade. Essa subamostragem possibilita que uma área maior seja considerada pela filtragem, ao mesmo tempo em que possibilita uma redução de complexidade e requisitos de recurso de processamento, em outras palavras a subamostragem pode reduzir os cálculos enquanto mantém um grande núcleo de filtro.

[160] No exemplo, a vizinhança/o núcleo 201 é, além disso, assimétrica/o em relação à posição do primeiro valor de profundidade filtrado  $D_{i,vizinhos}$  e, dessa forma, em relação ao primeiro valor de profundidade  $D_i$  (ou seja, é assimétrica em relação à posição 203 do pixel atual, ou seja, no que diz respeito à posição  $i$ ).

[161] Esta assimetria pode ser particularmente adequada para cenários em que, por exemplo, o mapa de profundidade é gerado com base na estimativa de disparidade em ao menos duas imagens. Especificamente, um mapa de profundidade (incluindo um mapa de disparidade onde a profundidade é representada por uma disparidade) pode compreender valores de profundidade que tenham sido gerados por estimativa de disparidade entre duas imagens com diferentes pontos de vista. Por exemplo, a estimativa de disparidade pode ser realizada entre uma imagem de olho esquerdo e uma imagem de olho direito.

[162] Em tal um exemplo, a imagem de intensidade de luz pode ser fornecida para o ponto de vista de olho

esquerdo, e o mapa de profundidade filtrado pode ser gerado para este ponto de vista de olho esquerdo com o uso da imagem de intensidade de luz do olho esquerdo como base. Entretanto, em tal caso, é mais provável que a imagem de intensidade de luz e, de fato, o mapa de profundidade de entrada tenha regiões mais ocultas à direita de um objeto em primeiro plano do que à esquerda do objeto em primeiro plano, pois o ponto de vista do olho esquerdo possibilita que o plano de fundo seja visto à esquerda do objeto no primeiro plano, o que ficará oculto a um ponto de vista do olho direito.

[163] O oposto será para o caso do ponto de vista de olho direito, ou seja, neste caso as áreas ocultas tenderão a se estender mais para o lado esquerdo de um objeto no primeiro plano do que para o lado direito.

[164] Assim, em tal cenário, uma estimativa de disparidade pode não ser capaz de encontrar na imagem do olho direito a área da imagem correspondente a um plano de fundo que não está oculto na imagem do olho esquerdo, uma vez que pode estar oculto neste. Como consequência, a estimativa de disparidade tenderá a gerar áreas de alta incerteza à esquerda das imagens em primeiro plano. Adicionalmente, é provável que os valores de profundidade corretos para estas regiões caiam sobre à esquerda da área mais do que à direita, à medida que eles serão mais propensos a refletir o plano de fundo ao invés do objeto em primeiro plano.

[165] Isso pode ser levado em conta ao se ter um núcleo espacialmente assimétrico.

[166] Especificamente, quando se usa a imagem do olho esquerdo como referência, o núcleo pode em muitas modalidades vantajosamente se estender mais na direção

esquerda do que na direção direita, como de fato é refletido no exemplo da Figura 2. Claramente, a situação simétrica ocorre se a imagem do olho direito é usada como base para a estimativa de disparidade e, portanto, o núcleo em tal exemplo se estende ainda mais à direita do que à esquerda.

[167] Dessa forma, em cenários em que o mapa de profundidade compreende valores de profundidade determinados a partir da estimativa de disparidade em uma primeira imagem representando um primeiro ponto de vista e uma segunda imagem representando um segundo ponto de vista, e onde o primeiro ponto de vista é a base para a estimativa de disparidade, a vizinhança/núcleo pode ser projetada/o para se estender mais no sentido do segundo ponto de vista para o primeiro ponto de vista do que no sentido do primeiro ponto de vista para o segundo ponto de vista.

[168] Uma situação similar ocorre em um sensor de profundidade estruturado à base de luz onde uma fonte de luz estruturada (infravermelho) e um sensor de imagem (infravermelho) têm uma localização espacial diferente no compartimento do sensor para formar uma linha de base para a captura de profundidade. Neste cenário, certas partes da cena que serão imagens pelo sensor não serão iluminadas pela fonte de luz (devido à linha de base entre estas duas). Também aqui, o conhecimento do vetor de direção 3D entre a fonte de luz estruturada e um sensor pode ser usado para formatar assimetricamente o núcleo de filtro.

[169] A descrição anterior deu enfoque a um cenário em que todas as imagens e mapas são considerados ter a mesma resolução. Entretanto, deve-se compreender que o princípio descrito é igualmente aplicável a cenários em que a



resolução pode ser diferente e, em particular, pode ser adequado para, por exemplo, o aumento artificial da taxa de amostragem de um mapa de propriedades de imagem com base em uma imagem de intensidade de luz de maior resolução.

[170] Nos exemplos, o mapa de profundidade/disparidade é acompanhado por um mapa de confiança associado que indica a confiança dos valores de profundidade. Por exemplo, o mapa de profundidade pode representar estimativas de disparidade e o mapa de confiança pode quantificar o quão confiável é a estimativa resultante. Outro exemplo é um sensor de profundidade à base de luz infravermelha que tipicamente também emite um mapa de confiança que quantifica a exatidão da medição.

[171] Geralmente, as imagens de intensidade de luz não têm mapas de confiança associados ou indicações de que tais imagens são tipicamente diretamente capturadas ou geradas, e são consideradas como sendo completamente confiáveis em todos os lugares. Entretanto, em algumas modalidades, a filtragem pode ser em resposta a uma terceira propriedade em adição aos valores de intensidade de luz e aos valores de confiança para o mapa de profundidade. Em tais exemplos, a terceira propriedade pode também ter valores de confiança associados. Por exemplo, a terceira propriedade pode ser uma propriedade de estimativa de movimento e, por exemplo, para cada pixel, podem ser fornecidos um vetor de movimento e um valor de confiança de movimento associado. A determinação dos pesos para os valores de profundidade vizinhos  $D_j$  pode ser adicionalmente realizada com base nestes valores.

[172] Especificamente, em algumas modalidades, o receptor pode, conseqüentemente, também receber um segundo

mapa de propriedades de imagem e um segundo mapa de confiança, sendo que o segundo mapa de confiança compreende segundos valores de confiança para os segundos valores de propriedade de imagem do segundo mapa de propriedades da imagem. Por exemplo, as imagens compostas podem compreender também um mapa de vetor de movimento e um mapa de confiança de movimento.

[173] Em tais modalidades, a determinação de um peso  $w_j$  para um valor de profundidade vizinho  $D_j$  pode também ser em resposta tanto ao segundo valor de propriedade de imagem, especificamente, o vetor de movimento, como ao segundo valor de confiança, especificamente, o valor de confiança de movimento, para a posição  $j$  (e, em algumas modalidades, também para a posição  $i$ ).

[174] Por exemplo, pode ser considerado um cenário em que um sensor de profundidade à base de infravermelho (luz estruturada ou tempo de voo) é usado para capturar um mapa de profundidade. Se o sensor também contém um sensor de cor de imagem, então este sensor de cor de imagem tipicamente terá uma resolução espacial muito maior. Além disso, se o sensor for usado ao longo do tempo (profundidade e imagem) para capturar uma sequência de, por exemplo, uma cena estática, ao se mover o sensor (varredura), o sensor de imagem pode ser usado para estimar um campo de movimento preciso de pixel  $v_{k,k+1}$  entre quadros  $k$  e  $k+1$ . Visto que as transições neste campo de movimento tipicamente mudam com as descontinuidades de profundidade (o sensor foi transladado), o campo de movimento pode ser usado para, por exemplo, ampliar/filtrar o mapa de profundidade do sensor de infravermelho. Entretanto, o campo de movimento é o resultado de um processo de estimativa de correspondência entre imagem

$\mathbf{I}_k$  e  $\mathbf{I}_{k+1}$  e, portanto, terá valores de confiança espacialmente variáveis.

[175] Este campo de movimento pode ser representado por um mapa de vetor de movimento e os valores de confiança de movimento podem ser representados por um mapa de confiança de movimento. O peso para um dado valor de profundidade vizinho  $D_j$  pode então ser dependente dos valores desses mapas na posição  $j$  e possivelmente também na posição  $i$ .

[176] Especificamente, o valor de profundidade vizinho combinado  $D_{i,vizinhos}$  pode ser determinado como:

$$[177] \quad D_{i,vizinhos} = \frac{\sum f_j g_j h_j D_j}{\sum f_j g_j h_j},$$

[178] Onde  $h_j$  representa um valor que é dependente tanto do valor de movimento na posição  $j$  como do nível de confiança na posição  $j$ . Tipicamente,  $h_j$  é dependente da diferença entre o valor de movimento na posição  $j$  e na posição  $i$ . Por exemplo,  $h_j$  pode ser dado como:

$$[179] \quad h_j \equiv \min.(C_{\mathbf{v}_i}, C_{\mathbf{v}_j}) e^{-\rho \|\mathbf{v}_i - \mathbf{v}_j\|}$$

[180] Onde  $\rho$  é um parâmetro de design que pode ser ajustado para variar o peso dos dados de movimento na filtração total,  $\|\mathbf{v}_i - \mathbf{v}_j\|$  reflete a diferença entre os valores de movimento nas posições  $i$  e  $j$  e  $C_{\mathbf{v}_i}$  e  $C_{\mathbf{v}_j}$  representa os valores de confiança de movimento na posição  $i$  e  $j$ , respectivamente.

[181] O uso de  $\min.(C_{\mathbf{v}_i}, C_{\mathbf{v}_j})$  para pesar este termo é uma solução lógica, visto que, quando o deslocamento é confiável tanto em posição central  $i$  como na posição vizinha  $j$ , então a diferença  $\|\mathbf{v}_i - \mathbf{v}_j\|$  é provavelmente confiável. Se  $i$  ou  $j$  tiver uma baixa confiança, então a diferença  $\|\mathbf{v}_i - \mathbf{v}_j\|$  provavelmente será menos confiável. Entretanto, deve-se

considerar que outras abordagens podem ser usadas em outras modalidades.

[182] Por exemplo, em tal cenário, a filtragem descrita pode ser realizada para cada posição no mapa de propriedades de imagem de saída, ou seja, para cada posição no mapa de profundidade de alta resolução. A vizinhança pode ser determinada nesta alta resolução e, para cada pixel nesta vizinhança de alta resolução, uma contribuição para o valor de propriedade de imagem da vizinhança combinado pode ser determinada. Em tal caso, o mesmo valor de propriedade de imagem de entrada e valor de confiança podem ser usados para todas as posições de alta resolução dentro do mesmo pixel de baixa resolução. No entanto, os valores de intensidade de luz diferirão, visto que são fornecidos na resolução mais alta e, dessa forma, pode ser obtido um aumento da taxa de amostragem eficaz do mapa de propriedades de imagem de baixa resolução com base na imagem de intensidade de luz de alta resolução.

[183] Deve-se considerar que, para fins de clareza, a descrição acima descreveu as modalidades da invenção com referência a diferentes circuitos, unidades e processadores funcionais. Entretanto, ficará evidente que qualquer distribuição adequada de funcionalidade entre os diferentes circuitos, unidades ou processadores funcionais pode ser usada sem se desviar da invenção. Por exemplo, a funcionalidade ilustrada a ser executada por processadores ou controladores separados pode ser executada pelo mesmo processador ou pelos mesmos controladores. Por isso, as referências a unidades ou circuitos funcionais específicos devem ser consideradas apenas como referências a meios adequados para fornecer a funcionalidade descrita e não como

indicativas de uma estrutura física ou uma organização lógica ou física estrita.

[184] A invenção pode ser implementada em qualquer forma adequada, incluindo hardware, software, firmware ou qualquer combinação dos mesmos. A invenção pode ser, opcionalmente, implementada, ao menos parcialmente, como software de computador que é executado em um ou mais processadores de dados e/ou processadores de sinal digital. Os elementos e componentes de uma modalidade da invenção podem ser implementados física, funcional e logicamente de qualquer forma adequada. De fato, a funcionalidade pode ser implementada em uma unidade única, em uma pluralidade de unidades ou como parte de outras unidades funcionais. Assim, a invenção pode ser implementada em uma unidade única ou pode ser distribuída física e funcionalmente entre diferentes unidades, circuitos e processadores.

[185] Embora a presente invenção tenha sido descrita em conexão com algumas modalidades, a mesma não está destinada a ser limitada à forma específica aqui apresentada. Em vez disso, o escopo da presente invenção é limitado apenas pelas reivindicações em anexo. Adicionalmente, embora possa parecer que um recurso é descrito em conexão com modalidades específicas, o elemento versado na técnica reconhecerá que vários recursos das modalidades descritas podem ser combinados de acordo com a invenção. Nas reivindicações, o termo "que compreende" não exclui a presença de outros elementos ou outras etapas.

[186] Um pixel pode ser considerado o que, na técnica anterior em geral, é denominado subpixel, ou seja, um único pixel colorido como, por exemplo, um pixel vermelho. No

entanto, na invenção descrição acima, o termo "pixel" pode também ser considerado compreender um grupo de subpixels, dessa forma, por exemplo, um pixel pode ser um pixel RGB, RGBY, entre outros.

[187] Além disso, embora individualmente mencionados, uma pluralidade de meios, elementos, circuitos ou etapas de métodos podem ser implementados, por exemplo por meio de um único circuito, uma única unidade ou um único processador. Adicionalmente, embora recursos individuais possam estar incluídos em reivindicações diferentes, eles podem ser vantajosamente combinados, e sua inclusão em reivindicações diferentes não implica que uma combinação de recursos não seja viável e/ou vantajosa. Além disso, a inclusão de um recurso em uma categoria de reivindicações não implica uma limitação a essa categoria, porém, em vez disso, indica que o recurso é igualmente aplicável a outras categorias das reivindicações, conforme for adequado. Além disso, a ordem dos recursos nas reivindicações não implica em nenhuma ordem específica na qual os recursos precisam ser trabalhados e, em particular, a ordem das etapas individuais em uma reivindicação de método não implica que as etapas precisam ser executadas nessa ordem. As etapas podem, na verdade, ser executadas em qualquer ordem adequada. Além disso, referências no singular não excluem uma pluralidade. Dessa forma, as referências a "um(a)", "uns/umas", "primeiro(a)", "segundo(a)" etc., não excluem uma pluralidade. Os sinais de referência nas reivindicações são fornecidos meramente como um exemplo de esclarecimento e não devem ser considerados como limitantes ao escopo das reivindicações.

### REIVINDICAÇÕES

1. APARELHO PARA CAPTURAR UM MAPA DE PROPRIEDADES DE IMAGEM, sendo o aparelho caracterizado por compreender:

um receptor (101) para receber uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem, sendo que a imagem de intensidade de luz compreende valores de intensidade de luz para pixels da imagem de intensidade de luz, sendo que o mapa de propriedades de imagem compreende valores de propriedade de imagem para os pixels da imagem de intensidade de luz e o mapa de confiança compreende valores de confiança para os valores de propriedade de imagem;

uma unidade de filtro (103) disposta para filtrar o mapa de propriedades de imagem em resposta à imagem de intensidade de luz e ao mapa de confiança, para gerar um mapa de propriedades de imagem filtrado, sendo que a unidade de filtro (103) é disposta de modo a, para uma primeira posição no mapa de propriedades de imagem filtrado:

determinar um valor de propriedade de imagem combinado de vizinhança em resposta ao fato de uma combinação ponderada de valores de propriedade de imagem ser valores de propriedades de imagem de vizinhança em uma vizinhança em torno da primeira posição, um peso de um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição ser dependente de um valor de confiança do primeiro valor de propriedade de imagem da vizinhança e de uma diferença entre um valor de intensidade de luz da primeira posição e um valor de intensidade de luz para a segunda posição; e

determinar o primeiro valor de propriedade de imagem filtrado para para a primeira posição como uma combinação de

um primeiro valor de propriedade de imagem na primeira posição no mapa de propriedades de imagem e o valor de propriedade de imagem da vizinhança combinado;

sendo que a unidade de filtro (103) é disposta de modo a determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a uma medida de diferença do valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o primeiro valor de propriedade de imagem, sendo que a determinação do peso é assimétrica em relação ao valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o valor de propriedade da primeira imagem, com o peso dependendo do sinal da medida de diferença.

2. APARELHO, de acordo com a reivindicação 1, caracterizado pela unidade de filtro (103) ser disposta para:

determinar um peso para o primeiro valor de propriedade de imagem em relação a um peso para o valor de propriedade de imagem da vizinhança combinado em resposta a um valor de confiança para o primeiro valor de propriedade de imagem.

3. APARELHO, de acordo com qualquer uma das reivindicações 1 ou 2, caracterizado pela unidade de filtro (103) ser disposta para determinar o valor de propriedade de imagem combinado substancialmente como:

$$D_{\text{vizinho}} = \frac{\sum \alpha_j f_j g_j D_j}{\sum \alpha_j f_j g_j}$$

onde  $f_j$  representa uma medida de diferença para um valor de intensidade de luz da segunda posição e um valor de intensidade de luz para um valor de propriedade de imagem da vizinhança na segunda posição;  $g_j$  representa uma medida de



diferença de confiança para o valor de confiança para o primeiro valor de profundidade e um valor de confiança para o valor de propriedade de imagem da vizinhança na segunda posição;  $D_j$  representa o valor de propriedade de imagem da vizinhança na segunda posição;  $\alpha_j$  representa um parâmetro de design, e as somas são sobre todas as posições  $j$  pertencentes à vizinhança.

4. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pela medida de diferença ser limitada a um valor mínimo quando o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança não exceder o limite do valor de confiança para o primeiro valor de propriedade de imagem por um certo limite, e pela medida da diferença exceder o valor mínimo se o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança exceder o limite do valor de confiança para o primeiro valor de propriedade de imagem por um certo limite.

5. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pela imagem de intensidade de luz, o mapa de confiança e o mapa de propriedades de imagem fazerem parte de uma sequência temporal de imagens de intensidade de luz, mapas de confiança e mapas de propriedade da imagem, e a vizinhança ter uma extensão temporal.

6. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pela unidade de filtro (103) ser disposta para aumentar um valor de confiança para um valor de propriedade de imagem se este for um valor de propriedade de imagem filtrado.

7. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pela vizinhança ser espacialmente assimétrica em relação à primeira posição.

8. APARELHO, de acordo com a reivindicação 7, caracterizado pelo mapa de propriedades de imagem ser um mapa de indicação de profundidade que compreende valores de profundidade determinados a partir de uma estimativa de disparidade em uma primeira imagem representando um primeiro ponto de vista e uma segunda imagem representando um segundo ponto de vista e tendo o primeiro ponto de vista como base para a estimativa de disparidade; e pela vizinhança se estender mais no sentido do segundo ponto de vista para o primeiro ponto de vista do que no sentido do primeiro ponto de vista para o segundo ponto de vista.

9. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo peso ser dependente adicionalmente de uma distância entre a primeira posição e a segunda posição.

10. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo receptor ser disposto adicionalmente para receber um segundo mapa de propriedades de imagem e um segundo mapa de confiança, sendo que o segundo mapa de confiança compreende segundos valores de confiança para segundos valores de propriedade de imagem do segundo mapa de propriedades de imagem; e pela unidade de filtro (103) ser disposta adicionalmente de modo a determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a um segundo valor de propriedade de imagem para a segunda posição e a um segundo valor de confiança para a segunda posição.

11. APARELHO, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo mapa de propriedades de imagem ser um dentre:

um mapa de profundidade;  
um mapa de disparidade;  
um mapa de estimativa de movimento; e  
um mapa de transparência.

12. MÉTODO DE PROCESSAMENTO DE UM MAPA DE PROPRIEDADES DE IMAGEM, sendo o método caracterizado por compreender:

receber uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem, sendo que a imagem de intensidade de luz compreende valores de intensidade de luz para pixels da imagem de intensidade de luz, o mapa de propriedades de imagem compreende valores de propriedade de imagem para os pixels da imagem de intensidade de luz e o mapa de confiança compreende valores de confiança para os valores de propriedade da imagem;

filtrar o mapa de propriedades da imagem em resposta à imagem de intensidade de luz e ao mapa de confiança, para gerar um mapa de propriedades de imagem filtrado, sendo que a filtragem compreende, para uma primeira posição no mapa de propriedades de imagem filtrado:

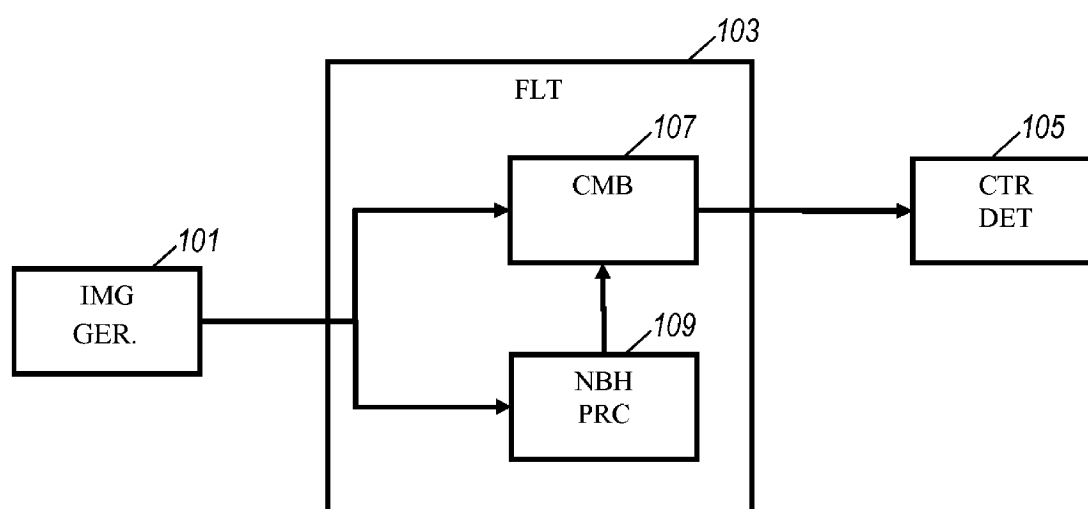
determinar um valor de propriedade de imagem combinado da vizinhança em resposta ao fato de uma combinação ponderada de valores de propriedade de imagem ser valores de propriedades de imagem de vizinhança em uma vizinhança em torno da primeira posição, um peso para um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição ser dependente de um valor de confiança para o primeiro valor de

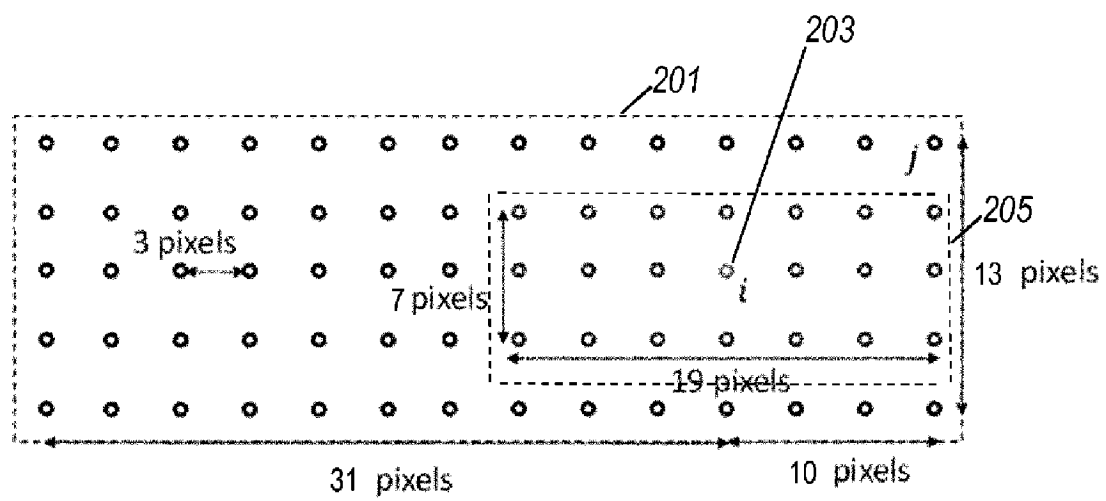
propriedade de imagem da vizinhança e de uma diferença entre um valor de intensidade de luz para a primeira posição e um valor de intensidade de luz para a segunda posição; e

determinar um primeiro valor de propriedade de imagem filtrado para a primeira posição como uma combinação de um primeiro valor de propriedade de imagem no mapa de propriedades de imagem na primeira posição e do valor de propriedade da imagem vizinho combinado;

sendo que a determinação do valor combinado de propriedade de imagem vizinho compreende determinar o peso para o primeiro valor de propriedade de imagem da vizinhança em resposta a uma medida de diferença para o valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e a um valor de confiança para o primeiro valor de propriedade de imagem, sendo que a determinação do peso é assimétrica em relação ao valor de confiança para o primeiro valor de propriedade de imagem da vizinhança e o valor de confiança para o valor de propriedade da primeira imagem, com o peso dependendo do sinal da medida de diferença.

13. PRODUTO DE PROGRAMA DE COMPUTADOR, caracterizado por compreender meios de código de programa de computador adaptados para executar todas as etapas conforme definidas na reivindicação 12, quando o dito programa é executado em um computador.

**FIG. 1**

**FIG. 2**

RESUMO

APARELHO PARA CAPTURAR UM MAPA DE PROPRIEDADES DE IMAGEM, MÉTODO DE PROCESSAMENTO DE UM MAPA DE PROPRIEDADES DE IMAGEM E PRODUTO DE PROGRAMA DE COMPUTADOR

A presente invenção se refere a um aparelho que compreende um receptor (101) que recebe uma imagem de intensidade de luz, um mapa de confiança e um mapa de propriedades de imagem. Uma unidade de filtro (103) é disposta para filtrar o mapa de propriedades de imagem em resposta à imagem de intensidade de luz e ao mapa de confiança. Especificamente, para uma primeira posição, a unidade de filtro (103) determina um valor de propriedade de imagem combinado da vizinhança em resposta a uma combinação ponderada de valores de propriedade de imagem vizinhos em uma vizinhança em torno da primeira posição, o peso para um primeiro valor de propriedade de imagem da vizinhança em uma segunda posição ser dependente de um valor de confiança para o primeiro valor de propriedade vizinho e de uma diferença entre valores de intensidade de luz para a primeira posição e a segunda posição; e determina um primeiro valor de propriedade de imagem filtrado para a primeira posição como uma combinação de um primeiro valor de propriedade de imagem na primeira posição no mapa de propriedades de imagem e o valor de propriedade de imagem combinado.